PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	A randomised trial assessing the impact of framing of fracture risk and osteoporosis treatment benefits in patients undergoing bone densitometry
AUTHORS	Kalluru, Ramanamma; Petrie, Keith; Grey, Andrew; Nisa, Zaynah; Horne, Anne; Gamble, Greg; Bolland, Mark

VERSION 1 - REVIEW

REVIEWER	Teppo Järvinen University of Helsinki Deparment of Orthopedics and Traumatology Helsinki FINLAND I have published quite extensively on osteoporosis, viability of a strategy to screen for patients "at risk" of fractures and the subsequent need for pharmacotherapy to boost bone mass. I am also a believer in shared decision making.
REVIEW RETURNED	30-Aug-2016

GENERAL COMMENTS	Thank you for this opportunity to review the manuscript entitled “Accuracy of perceptions of fracture risk and osteoporosis treatment benefits: a randomised trial”, submitted to the BMJ Open. This is a very interesting study from a group that is very well known in this field and who have contributed a number of pivotal papers in the recent history. Below please find my comments on the paper that I found of stellar quality; very thoroughly executed, clearly written, and of fundamental importance to the field. General comments Nowadays being at ‘high risk’ of having a disease has become a disease in and of itself. Sweeping educational programmes at all levels of health care now turn an otherwise healthy person’s ‘high’ blood pressure, elevated serum lipids/high risk of cardiovascular disease or a high risk of fracture into chronic conditions of increased risk of a potentially bad event.1 But what represents ‘high risk’? This question lies at the heart of modern medicine, particularly with respect to pharmacological primary prevention. Current debate on ‘High Risk as a Disease’ twirls exactly around the two topics of the paper under consideration, the ‘appropriate’ threshold for defining something as a disease and the appropriate benefits of the administered treatment (here, bone pharmacotherapy). Regarding the former, the medical experts of the National Osteoporosis Foundation (NOF, USA) introduced an osteoporosis

guideline that recommends osteoporosis medications if a person's 10-year probability of sustaining a hip fracture is 3% or over (more recently, this threshold has apparently been lowered to 2%). Applying these NOF recommendations to a large prospective cohort study, at least 72% of U.S. white women >65 years and 93% of those >75 years would be recommended for drug therapy.2 But who is the right party to determine the threshold for 'high risk'? Advocates of the hegemony of medical experts argue that doctors – as content experts – should be the responsible party.3 However, evidence from behavioural sciences suggests that we (people in general) are poor in making probability decisions:4 despite laudable efforts to improve the communication and comprehension of both the concept of risk and the anticipated treatment-benefit, risk-illiteracy of the gravest magnitude still affects both doctors and patients.5 The results of this paper, using top-notch methodology, show that the above noted NOF threshold is more than 15 times lower than what patients consider an appropriate 5-year fracture risk (to justify initiation of bone-targeted pharmacotherapy: 50-60%).

The other relevant question regarding preventive pharmacotherapy – as noted by the authors of this publication – is what represent 'effective treatment'. And as shown by this paper, there is a huge gap on what the doctors and lay public/patients consider 'effective treatment'; when a drug increased the probability of avoiding a hip fracture from 97.9% to 98.9% (1% absolute risk advantage, only framed as a 50% relative risk reduction) (Alendronate6), doctors world-wide began prescribing enthusiastically.

What if our patients do not agree with our perceptions on the two above noted issues, the threshold for 'high risk' and the appropriate efficacy of treatment offered? Without accurate and common comprehension on the threshold for 'high risk' (of a fracture), there is no basis for shared decisions. And without shared decision making, pharmacological prevention of fractures becomes eminence and evidence tyranny. This paper proves – in my opinion, quite conclusively – that we have neither accurate nor common comprehension on these two issues. I fully agree with the authors' interpretation of their main finding(s) that providing estimated absolute risks of fracture and benefits of treatment in different ways had little effect on participants' perception of their need to take treatment or their individual risk of fracture. As difficult as the interpretation of this study might be, should be accept the unpleasant truth and act accordingly?

This is – in brief – my understanding on the background and ramifications of this paper. Needless to say, I found the idea of extensively/comprehensively assessing the perceptions of fracture risk and osteoporosis treatment benefits through a randomized controlled trial amazingly innovative and vitally important. After a careful review of the paper, I have no comments (suggestions for edits) to make, as I feel that the reporting in the paper is both accurate and very balanced.

1. Moynihan R. Surrogates under scrutiny: fallible correlations, fatal consequences. Bmj 2011;343:d5160.
2. Donaldson MG, Cawthon PM, Lui LY, et al. Estimates of the proportion of older white women who would be recommended for pharmacologic treatment by the new U.S. National Osteoporosis Foundation Guidelines. Journal of bone and mineral research : the

official journal of the American Society for Bone and Mineral Research 2009;24(4):675-80.
3. Moynihan R. Medicalization. A new deal on disease definition. Bmj 2011;342:d2548.
4. Kahneman D, Tversky A. Prospect Theory - Analysis of Decision under Risk. Econometrica 1979;47(2):263-91.
5. Martyn C. Risky business: doctors' understanding of statistics. Bmj 2014;349:g5619.
6. Black DM, Cummings SR, Karpf DB, et al. Randomised trial of effect of alendronate on risk of fracture in women with existing vertebral fractures. Fracture Intervention Trial Research Group. Lancet 1996;348(9041):1535-41.

REVIEWER	Melissa Orlandin Premaor Federal University of Santa Maria, Brazil
REVIEW RETURNED	08-Oct-2016

GENERAL COMMENTS	This manuscript addresses a relevant subject that is the perceived risk of fractures. However, there are some issues with the design of the study that may prevent its generalization. Firstly, the study was carried out at bone density facility located at a public hospital clinic. All subjects had a bone scan. Sato et. al. have shown that the scan per se might affect the perception of osteoporosis which might have had some influence in the results. Moreover, only a minority of the population world wild might have access to these exams making the external validity of the study more difficult to interpret. Secondly, all the four study groups received the intervention. Although there were no differences between the groups, it is not possible to know if to do a presentation that framed the patient's absolute fracture risk is better than do no intervention. Finally, the authors use the visual analog scale to access the perceived risk of fractures. Its usefulness in assessing the perception of the risk of fractures, not this well established. Subjects tend to report their risk on this scale as 0, 50 and 100%, which may ultimately confound interpretation of the results found in this study. Minor comments Title: • This is not an accuracy study it is a clinical trial, please suppress "accuracy from the title. Methods: • The methods are incomplete and difficult to follow. A BMJ reader might not be able to follow it fully. Please be clearer and give more detail. Example: the authors never explained that they had used the analog scale in the methods it is only shown in the supplementary material. • Statistics: how the author handles the subjects who report their risk as 0, 50 and 100%? It is a well-known issue in the risk perception studies. • Page 8, last line. Why should the reader be worried about the patient's involvement in the design of the study? Discussion: • It is expected that the analog scale would overestimate the risk, but the authors never commented this instrument limitation at the discussion. • I do not agree with the author's comment on the ROSE study.

	Although there might be an anchoring bias into this study participant estimates its primary purpose was to classify then as a low, moderate and high risk. Regular people do not interpret risk the same way researches and physicians do. They usually perceive 50% risk as a moderate risk. (de Bruin WB Medical Decision Makings 32 (2) 232-36

REVIEWER	Prof Gonnelli Stefano, MD University of Siena 53100 Siena Italy
REVIEW RETURNED	10-Oct-2016

GENERAL COMMENTS	The manuscript “Accuracy of perceptions of fracture risk and osteoporosis treatment benefits: a randomised trial ” by Rama Kalluru et al aimed to investigate the effects of framing of the risk of fracture and benefits of osteoporosis treatments, and whether this influences patients’ beliefs about the need for osteoporosis. The conclusion was that providing estimated absolute risks of fracture and benefits of treatment in four different ways had little effect on participants’ perception of their need to take treatment or their individual risk of fracture. COMMENTS 1). The sections "Materials and Methods" and "Results" appear too long and difficult to understand. Authors should clarify whether the study participants after the DXA measurement received any anti-osteoporosis treatments and if this could affect the answers to the third questionnaire. 2) The study flow chart should be moved from the “Appendix” to the "Materials and Methods" section. 3). Table 2: the level of education (or scholarity) should be taken into consideration and reported in this table. 4) Did the Authors observe any differences in the answers to questionnaires between men and women? 5). The “Discussion” is too long and should focused on the results of this study.

REVIEWER	Isabel E. Allen University of California, San Francisco, USA
REVIEW RETURNED	31-Oct-2016

GENERAL COMMENTS	This is a well-designed study statistically but the design is not novel and the results need to be clarified and presented more clearly. The study is presented as a novel design but a pre-post knowledge/understanding design with 4 different presentations of risk/benefit and treatment options is not new in decision/counseling clinical literature. Decision-making and studies & understanding risk have been used in breast cancer counseling for many years (Lerman, 1994, JNCI). Beyond the design, there are several statistical areas of the paper that need tightening: (1) None of the tables present the p-values comparing the groups (either randomized or examined by needing treatment or not); (2) A p-value in the middle of page 10 [p>0.4] that refers to Table 3 or Figure 2 but with no attribution for a test or giving the exact p-value; (3) I assume

	that the p-values have been adjusted for multiple comparisons but this is not mentioned in the paper; (4) were there any differences when analyzing completers vs. all participants since some were lost to follow-up? I would recommend submitting to an osteoporosis journal where there would be more interest and the study design is novel.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Thank you for this opportunity to review the manuscript entitled “Accuracy of perceptions of fracture risk and osteoporosis treatment benefits: a randomised trial”, submitted to the BMJ Open. This is a very interesting study from a group that is very well known in this field and who have contributed a number of pivotal papers in the recent history. Below please find my comments on the paper that I found of stellar quality; very thoroughly executed, clearly written, and of fundamental importance to the field.

General comments
Nowadays being at ‘high risk’ …

This is – in brief – my understanding on the background and ramifications of this paper. Needless to say, I found the idea of extensively/comprehensively assessing the perceptions of fracture risk and osteoporosis treatment benefits through a randomized controlled trial amazingly innovative and vitally important. After a careful review of the paper, I have no comments (suggestions for edits) to make, as I feel that the reporting in the paper is both accurate and very balanced.

Response:
We thank the Reviewer for the supportive comments

Reviewer: 2
This manuscript addresses a relevant subject that is the perceived risk of fractures. However, there are some issues with the design of the study that may prevent its generalization.

Firstly, the study was carried out at bone density facility located at a public hospital clinic. All subjects had a bone scan. Sato et. al. have shown that the scan per se might affect the perception of osteoporosis which might have had some influence in the results. Moreover, only a minority of the population world wild might have access to these exams making the external validity of the study more difficult to interpret.

Response:
We have added this as a limitation in the relevant section of the Discussion.

“Whether the findings would be similar in cohorts at higher or lower risk of fracture, in cohorts who were not undergoing bone densitometry, or in other …”

Secondly, all the four study groups received the intervention. Although there were no differences between the groups, it is not possible to know if to do a presentation that framed the patient’s absolute fracture risk is better than do no intervention.

Response:
We asked the question ‘Did the different styles of presentation influence patients’ perception of their

fracture risk and treatment benefit?" This is a different question to whether giving patients information on their fracture risk and treatment benefit is better than not giving them this information, and would require a different study design, as the Reviewer highlights.

Finally, the authors use the visual analog scale to access the perceived risk of fractures. Its usefulness in assessing the perception of the risk of fractures, not this well established. Subjects tend to report their risk on this scale as 0, 50 and 100%, which may ultimately confound interpretation of the results found in this study.

Response:
We did not find this in our study. For example, the histogram of results of the participant estimates at baseline of their fracture risk are shown below.

Minor comments
Title:
• This is not an accuracy study it is a clinical trial, please suppress "accuracy from the title.

Response:
We have changed the title according to the Editor's guidance

Methods:
• The methods are incomplete and difficult to follow. A BMJ reader might not be able to follow it fully. Please be clearer and give more detail. Example: the authors never explained that they had used the analog scale in the methods it is only shown in the supplementary material.
Response
We have added to the relevant section of the Methods that we used a visual analogue scale. (Paragraph titled Questionnaires). We have carefully checked the Methods section- it contains all the information required in the Consort statement and we think it flows logically. We would be happy to make any other specific changes suggested.

• Statistics: how the author handles the subjects who report their risk as 0, 50 and 100%? It is a well-known issue in the risk perception studies.

Response
As indicated above, this was not an issue in our study. As described in the Methods, we used appropriate non-parametric tests where applicable throughout.

• Page 8, last line. Why should the reader be worried about the patient's involvement in the design of the study?

Response
We have removed the text from the manuscript as we believe this is not a requirement for submission to BMJ Open.

Discussion:
• It is expected that the analog scale would overestimate the risk, but the authors never commented

this instrument limitation at the discussion.

Response
Visual analogue scales are widely used in osteoporosis and other research fields. We were unaware that they systematically overestimate risk and undertook a brief literature search to find further information on this. We were not able to find information about this specifically for fracture risk. We also were unable to find reports of consistent overestimation of risk from visual analogue scales. We therefore have not changed the text but would be happy to consider this if the Reviewer provided relevant references.

• I do not agree with the author's comment on the ROSE study. Although there might be an anchoring bias into this study participant estimates its primary purpose was to classify then as a low, moderate and high risk. Regular people do not interpret risk the same way researches and physicians do. They usually perceive 50% risk as a moderate risk. (de Bruin WB Medical Decision Makings 32 (2) 232-36

Response
We are unsure what the Reviewer's concern is. We suggested that there may have been an anchoring bias introduced in this study because the investigators asked participants to classify their risk into one of five numeric categories, and that this may have contributed to the poor correlation between FRAX scores and their numeric estimates of risk. The Reviewer appears to agree that this may have been a possibility. Otherwise, we simply summarised the result of the ROSE study, and we specifically stated that 42% of women considered their risk below average, and 5% above average. The findings of our study seem consistent with the Reviewer's view "They usually perceive 50% risk …"- patients overestimated their risk of fracture but the majority considered themselves at low or very low risk of fracture.

Reviewer: 3
The manuscript "Accuracy of perceptions of fracture risk and osteoporosis treatment benefits: a randomised trial " by Rama Kalluru et al aimed to investigate the effects of framing of the risk of fracture and benefits of osteoporosis treatments, and whether this influences patients' beliefs about the need for osteoporosis. The conclusion was that providing estimated absolute risks of fracture and benefits of treatment in four different ways had little effect on participants' perception of their need to take treatment or their individual risk of fracture.

COMMENTS
1). The sections "Materials and Methods" and "Results" appear too long and difficult to understand. Authors should clarify whether the study participants after the DXA measurement received any anti-osteoporosis treatments and if this could affect the answers to the third questionnaire.

Response:
The Materials and Methods section contains 4 paragraphs and 674 words. It contains all the information required by the Consort statement. We have carefully checked the section and think it has a logical flow and does not contain extraneous information. The Results section has 6 paragraphs and we think it has a logical flow- baseline characteristics, primary endpoint, secondary endpoints. We would be happy to consider specific changes to either of these sections.

Table 4 shows the proportions who started or intended to start osteoporosis medication, and the relationship with the initial views about treatment. This is also discussed in the penultimate paragraph of the Results

2) The study flow chart should be moved from the “Appendix” to the "Materials and Methods" section.

Response:
We are conscious that the paper already has 4 tables and 2 figures. We note that BMJ Open recommends 5 or fewer figures and tables to improve readability. We therefore have retained this figure in the appendix as we think it is of lower importance than the other figures and tables, but would be happy to shift it into the manuscript if the Editor wishes.

3). Table 2: the level of education (or scholarity) should be taken into consideration and reported in this table.

Response:
We did not collect this information, although individuals who were unable to complete the questionnaires for language or cognitive reasons were excluded from participating, as outlined in the protocol.

4) Did the Authors observe any differences in the answers to questionnaires between men and women?

Response:
161/200 participants in the trial were female, which reflects the typical gender imbalance in osteoporosis studies. We think there are too few men in the study to perform analyses stratified by gender.

5). The “Discussion” is too long and should focused on the results of this study.

Response:
We followed the recommended structure for the Discussion, which includes comparison of the study findings to other studies. The Reviewer has commented that the Methods, Results and Discussion are all too long. However, the overall word count for the manuscript is ~3400 words, which falls within the guidance from BMJ Open, and is similar to other manuscripts of RCTs we have published previously. We would be happy to remove or shorten specific sections, if the Editor wishes.

Reviewer: 4
This is a well-designed study statistically but the design is not novel and the results need to be clarified and presented more clearly. The study is presented as a novel design but a pre-post knowledge/understanding design with 4 different presentations of risk/benefit and treatment options is not new in decision/counseling clinical literature. Decision-making and studies & understanding risk have been used in breast cancer counseling for many years (Lerman, 1994, JNCI).

Response:
We appreciate that there is an extensive literature about risk communication, and in the Introduction, paragraph 2, we succinctly summarise some of the relevant literature and provide references to some relevant systematic reviews.

However, as we highlight in the Introduction, osteoporosis management is now shifting towards a diagnosis or recommendation for treatment based on absolute risk of an event, just as has happened

with cardiovascular disease. As we indicated in the Introduction, few trials have explored risk communication in the setting of chronic conditions in which indications for treatment are based upon the absolute risk of an event within a set time frame.

We appreciate that similar studies have been carried out in other conditions previously, but as we state in the Discussion: “Although there is a large body of research into risk communication,5-11 we are not aware of similar trials that have explored the impact of risk framing on risk perception and treatment benefits, either where the absolute risk of an event forms the basis for treatment recommendations or in the field of osteoporosis.” If there are such studies, we would be happy to reference and discuss them.

Based on the Reviewer’s suggestion, we have removed the comments about novelty from the article summary and strengths and limitation section of the Discussion.

Beyond the design, there are several statistical areas of the paper that need tightening:
(1) None of the tables present the p-values comparing the groups (either randomized or examined by needing treatment or not);

Response:
Table 2 presents baseline characteristics. P-values for between-groups differences in baseline values are not recommended in the setting of a randomised trial. The upper 2/3 of Table 3 also presents baseline characteristics, and the relevant p-values for the lower 1/3 of Table 3 are presented in the text of the Results. Table 4 presents frequency data about responses and we do not think p-values are required.

(2) A p-value in the middle of page 10 [p>0.4] that refers to Table 3 or Figure 2 but with no attribution for a test or giving the exact p-value;

Response:
We have now reported the exact p-values for the between-groups differences in the difference in patient estimates of hip and total fracture risk. These were calculated using the Kruskal-Wallace test as indicated in the Statistics section.

(3) I assume that the p-values have been adjusted for multiple comparisons but this is not mentioned in the paper;

Response:
We have not adjusted the p-values for multiple comparisons, and have now indicated this in the Statistics section of the Methods. The study does not have the traditional hierarchical structure of clearcut primary, secondary, and tertiary endpoints/hypotheses, and relatively few statistical tests have been performed. Therefore, we considered the issue of multiple testing but felt that it was not necessary to adjust the p-values for multiple comparisons. If the Editor felt it necessary, we could add information about the threshold for statistical significance if a Bonferroni correction were applied.

(4) were there any differences when analyzing completers vs. all participants since some were lost to follow-up?

Response:

All participants completed the two questionnaires at their study visit, but 14 participants (7%) did not return their 3 month questionnaire. Thus, this comment is relevant to the data in Table 4. The only analyses presented in that data were frequency analyses which were presented using all available data. We did not perform a “completers analysis”, but we have done that now and there is very little change in the frequencies presented in Table 4. We therefore have not added this information to the text.

I would recommend submitting to an osteoporosis journal where there would be more interest and the study design is novel.

Response:
Osteoporosis management in many countries is largely conducted in primary care, and not by osteoporosis specialists. We therefore think this trial is relevant to a broad general audience rather than a subspecialty one, and therefore is suited to BMJ Open.

VERSION 2 – REVIEW

REVIEWER	Melissa Orlandin Premaor Universidade Federal de Santa Maria, Brazil
REVIEW RETURNED	09-Dec-2016

GENERAL COMMENTS	I am happy about the responses.

REVIEWER	Gonnelli Stefano University of Siena-ITALY
REVIEW RETURNED	25-Dec-2016

GENERAL COMMENTS	The revised version of the manuscript seems to be suitable for publication in BMJ Open

REVIEWER	Isabel E. Allen UCSF, USA
REVIEW RETURNED	12-Dec-2016

GENERAL COMMENTS	The authors have responded to all my questions and the manuscript is ready for acceptance. Thank you.