Reviewer's report

Title: Dairy product consumption and risk of hip fracture: A systematic review and meta analysis

Version: 0 Date: 25 Sep 2017

Reviewer: Liisa Byberg

Reviewer's report:

The submitted manuscript presents a thorough review and meta-analysis of cohort and case-control studies examining the association between milk and other dairy products and hip fracture through 17 April 2017. Several different analytical methods were applied. Meta analyses of dietary intakes is difficult when the range of intakes differs between studies, especially if small intakes may have beneficial effects in comparison to null intake, whereas higher intakes may have other effects. The difficulty of combining the estimates into a summary measure is also indicated by the heterogeneity estimates in the presented meta-analysis. The statistical analyses seem well-conducted but there are problems with exposure definition in the pooled estimates and the interpretation of the results rely heavily on statistical significance and do not discuss the direction and strength of the observed estimates in addition to the p values, for instance in the subgroup analyses.

Specific comments and questions for the authors

1. The main results, presented in the abstract, in the first paragraph and in the conclusion of the discussion, are based on a pooled estimate of a comparison between the highest vs the lowest intake of milk (and other milk products) in each study. This comparison renders a nonsense estimate due to mixing apples and oranges; an estimate comparing 7 milk servings or more per week (ie 1 serving per day) with < 1 serving per week (Sahni 2014) is pooled with estimates comparing several servings per day vs <1 serving per day (Meyer 1997, Michaëlsson 2014) (results presented in Figure 2). As stated above, a small intake may have beneficial effects whereas higher intakes may have other effects. The same reasoning applies to the pooling of estimates per 200 grams/day (Figure 4); extrapolation of an estimate based on a study where the highest intake is 1 serving (200 grams) per day is a very strong assumption. Further, there is evidence of strong heterogeneity in these analyses. However, the authors have made efforts circumventing these problems and present dose-response analyses (Figure 5). Due to the limitations and problems with the other analyses, it seems reasonable that the authors present the dose-response analysis as their main results and update the abstract, discussion, and conclusion accordingly. I also recommend that the authors discuss why the pooling of the estimates in Figure 2 and 4 is not appropriate. The authors should also in the results section (text) highlight that the analyses producing the pooled estimates indicate heterogeneity.
2. In Figure 5 it would be helpful if there was a line parallel to the x-axis indicating RR/OR=1. The results text describing the dose-response analysis for the case-control studies should indicate that the confidence intervals were wide.

3. I suspect that there has been a mistake in the labelling of the columns Case-control studies and Cohort studies in Table 2, and that Cohort studies are presented to the left. These analyses are interesting although I think that the description of these results on page 9 could be both more fair and more detailed. If no adjustment for confounders such as smoking, total energy intake, alcohol consumption, physical activity, or use of calcium and vitamin D supplements, the RRs were below 1 and in studies where adjustments were made, the RRs were above 1. Also, the larger the number of cases, the larger the RR.

4. Figure 3 shows funnel plots that evaluate publication bias. In the results section it is said that plot 3a is symmetric and that plot 3b is asymmetric but that both indicate presence of publication bias. Please explain to the reader how you came to this conclusion. On line 200, page 9, it is referred to "most of the hollow circles" in Figure 3b. I can only see one hollow circle in my copy of the manuscript. Also, the number of filled circles does not equal the number of studies in the meta analysis. Please revise or explain.

5. The trim and fill method was used to estimate the number of missing studies and to provide an adjusted estimate including the imputed (filled) studies. This estimate is presented for the pooled estimates of highest vs lowest category for both cohort and case-control studies. For reasons discussed above, this pooling provides a nonsense estimate and using imputation cannot fix that. Is it possible to apply the trim and fill method on the dose-response analysis? Furthermore, it is suggested that the trim and fill-estimates should be interpreted with great caution because of the limitations related with the methods. These limitations should be discussed.

6. On page 12, lines 267-268, it should be made clear that vitamin D supplementation, with or without calcium, is likely to have only minor effects on fracture risk among community-dwelling individuals (Bolland 2014). Nonetheless, Chapuy et al showed that hip fracture risk was reduced with vitamin D and calcium supplementation among institutionalized women (mean age 84 years) who had very low vitamin D levels and a simultaneous low dietary intake of calcium. High doses of vitamin D has in addition been associated with an increased risk of falls and the sentence on lines 268-270 could therefore be omitted. The discussion of vitamin D and calcium supplementation as a potential confounder (on lines 271-278) is somewhat confusing to me. If indeed supplement use is a confounder, what studies would you rather trust: those that adjust for
supplement use or those that do not? Furthermore, supplement use could also represent other aspects such as health seeking behaviour, which could be an important confounder.

7. Different intake ranges is discussed as one potential reason for heterogeneity between studies, different confounder control as another. Identification of fracture events may be an additional reason in cohort studies. Differential loss to follow-up and ascertainment of hip fracture cases is discussed on page 13 but it should also be noted that this is especially relevant to cohort studies. Please also report hip fracture ascertainment method (register, self-report or both) in Table 1.

8. Please include a more detailed discussion of how recall bias can influence the results. In a nested case-control study collecting dietary data both at cohort entry (baseline, pre-fracture) and post-fracture (at time of case ascertainment), the odds ratio for dairy intake on fracture was above 1 when the baseline exposure assessment was used and below 1 when the post-fracture exposure assessment was used (Michaëllsson Int J Epidemiol 1996;25(2):403-10). (There may be other studies examining the role of recall bias in case-control studies examining milk/dairy and fracture risk that could be cited.)

9. Table 1: was amount of intake not specified in the study by Kanis (1999)?

10. Figures 2, 3, and 5 say "odd ratio" instead of "odds ratio", please revise.

11. Why was not our study in the BMJ (Michaëllsson 2014) included in the analysis of yogurt and cheese?

Liisa Byberg

Associate professor of medical epidemiology

Uppsala University
Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

No

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

Unable to assess

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

No

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?
If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal.