

## 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

<b>TITLE (PROVISIONAL)</b>	Relations of magnesium intake to cognitive impairment and dementia among participants in the Women's Health Initiative Memory Study: a prospective cohort study
<b>AUTHORS</b>	Lo, Kenneth; Liu, Qing; Madsen, Tracy; Rapp, Steve; Chen, Jiu-Chiuan; Neuhouser, Marian; Shadyab, Aladdin; Pal, Lubna; Lin, Xiaochen; Shumaker, Sally; Manson, JoAnn; Liu, Prof. Simin

### VERSION 1 – REVIEW

<b>REVIEWER</b>	Peter Lee P.N.Lee Statistics and Computing Ltd., Sutton, Surrey, UK
<b>REVIEW RETURNED</b>	18-Mar-2019

<b>GENERAL COMMENTS</b>	<p>“Relations of Magnesium Intake to Cognitive Impairment and Dementia Among Women participated in Women’s Health Initiative Memory Study” Comments on the paper submitted to BMJ Open</p> <hr/> <p>Author : P.N. Lee Date : 18th March 2019</p> <p>The authors describe the results of a prospective study of women without mild cognitive impairment (MCI) or probable dementia (PD) at baseline, relating magnesium (Mg) intake at baseline to onset of MCI or PD during follow-up. I have been asked to comment on this paper mainly with respect to statistical issues. My first impressions when starting to read the title and abstract were that they were very poorly presented. The title clearly does not need to say women twice, and would be better phrased something like “Relationship of magnesium intake to cognitive impairment and dementia among participants in the Women’s Health Initiative Memory Study”. In the abstract, under “design, setting and participants” it should be made clear that WHIMS-ECHO involves a subset of the WHIMS women. At first reading it sounds as if they are two separate studies. It also should be “We analyzed data from the Women’s .....”.</p> <p>It was the first sentence of the results that made my eyebrows raise. We start with 7,479 women aged 65 to 79 at enrolment and after 20 years of follow-up there were still as many as 6,473 (despite various exclusions). They must be very healthy! Clearly this is not what was meant. Indeed an exclusion criterion at follow-up for less than one year hardly makes sense in a population stated to have over 20 years of follow-up. Presumably the true situation is that the study involved 7,479 women aged 65-79 years at baseline who had been followed for at least one year and were not excluded for the reasons shown in Figure 1. In any case neither the numbers</p>
-------------------------	---

in the study nor the exclusion criteria are “results”. The study description should be made much clearer, both in the abstract and in the methods section.

Other points evident in the abstract were using “C.I.” when “CI” is more normal, use of the word “fully” which with no explanation seems to imply adjustment for every conceivable variable one might think of, and misspelling “tertile” as “terile”.

There is also, here and in a number of places in the text, an unfortunate tendency to omit “the” and “a”. For example, the first sentence of the Conclusions section of the Abstract should state “Total Mg intake between the Estimated Average Requirement and the Recommended Dietary Allowances may be associated with a lower risk .....”. Or the first sentence in the Methods section should start “The WHI.....”. There are also other inaccuracies, e.g. line 4 of the data source “women that aged 65-79 years.....” has an unnecessary and incorrect “that”. Generally, parts of the paper read as if drafted by one of the authors who was not completely fluent in English. This needs attention.

Turning now to statistical issues, I will start with some points of detail.

Table 1 should give the number of women in each of T1, T2 and T3, and should also give the mean (SD) of magnesium intake in each group.

Table 2 should make it clear that the results presented with 95% CI are hazard ratios.

The footnotes in both Tables 2 and 3 should start “Model 1 adjustment:”.

Presenting smaller p values as  $p < 0.01$  is not very illuminating. One might at least separate out  $p < 0.001$  and  $0.001 \geq p > 0.01$ . Indeed one might use a system of labelling p values \*, \*\* or \*\*\* depending on significance of 0.05, 0.01 or 0.001

I note in Table 3 that a Model 1 p of 0.03 (for T3, MCI) is wrongly unlabelled.

The text describing the results in Table 2 is somewhat confusing. It is stated that “Total Mg intake per 100mg increment was not associated with risk of MCI and/or PD in any of the Cox models assuming a linear association”. This is not true. For Model 1, there is significance for both MCI/PD and for MCI. If the authors intend to draw inferences only from Model 2 then they should make this clear.

In the same paragraph the HR of MCI/PD for tertile 2 is stated to be 0.65 in the text but 0.67 in Table 2.

This brings me on to what I regard as the major limitation of the paper. This that it claims that there is a protective effect in the second but not the third tertile without properly testing whether the model fits a linear relationship. When considering a dose-response relationship, the usual procedure is to start by testing for trend and for departure from trend. If there is a significant trend but no significant departure, one can simply describe the results by the trend. If there is significant departure one can investigate the shape of it.

The analyses are also limited by use of tertiles. The authors claim that this was done “to preserve statistical power with regard to the sample size” but there were over 100 cases in each tertile in each analysis so one clearly could have split the scale more finely, perhaps into six levels. Then one could present the HRs graphically showing them at each level, and illustrating linear relationships and, if shown to be statistically significant, non-linear relationships. Such dose-response analyses could be conducted using either the

	<p>continuous variable or using the HRs and mean magnesium doses at each level.</p> <p>Though I could not investigate the results fully without the detailed data, I did run some test based on the HRs (CIs) in each tertile, assigned doses of 1, 2 and 3 units (i.e. assuming that the doses of magnesium in the tertiles were equally spaced). For the MCI/PD Model 2 analysis in Table 2 I got significant heterogeneity between groups (chisquared 8.61 on 2 d.f., p=0.014) and significant trend (chisquared 5.77 on 1 d.f., p=0.016) but the difference – the departure from trend of 2.84 on 1 d.f. is not significant at p&lt;0.05. Similarly for the MCI Model 2 analysis in Table 2 there was again significant heterogeneity (chisquared 8.19 on 2 d.f., p=0.017) and trend (chisquared 6.48 on 1 d.f., p=0.011) but no significant departure from trend (chisquared 1.72 on 1 d.f.). On this approximate interpretation of the results I would conclude simply that risk of MCI/PD and of MCI declines with increasing total magnesium intake, but not draw inferences that the decline is only between T1 and T2 (unless a more precise dose-response analysis shows this).</p>
--	--

<b>REVIEWER</b>	Emily Harville Tulane University, USA
<b>REVIEW RETURNED</b>	21-Mar-2019

<b>GENERAL COMMENTS</b>	<p>How do the included and excluded participants compare to each other?</p> <p>For the combined endpoint and MCI (which appears to be driving it), tertile 3 is not significant but the effect estimate is similar, so it can't be considered evidence of no effect or evidence of a quadratic effect. It might be appropriate to look at non-linear (quadratic, spline) effects for the continuous variable, given that the tertiles did not show a dose-response effect. Also, the study is not particularly small, so I don't find the rationale for tertiles all that convincing, which makes me wonder if the authors tried other categorizations which failed to reach statistical significance. For these reasons, I think the abstract could represent the results better.</p> <p>The discussion of proportional hazards is a little odd. They say that no interaction is found, but provide p-values that go down to 0.05, then offer another example where the p-value was even lower. How did the HR change over time for this interaction?</p>
-------------------------	--

### VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Peter Lee

Institution and Country: P.N.Lee Statistics and Computing Ltd., Sutton, Surrey, UK

Please state any competing interests or state 'None declared': None declared

Q1 "The authors describe the results of a prospective study of women without mild cognitive impairment (MCI) or probable dementia (PD) at baseline, relating magnesium (Mg) intake at baseline to onset of MCI or PD during follow-up...My first impressions when starting to read the title and abstract were that they were very poorly presented. The title clearly does not need to say women twice, and would be better phrased something like "Relationship of magnesium intake to cognitive impairment and dementia among participants in the Women's Health Initiative Memory Study".

Response: Done.

Q2: In the abstract, under “design, setting and participants” it should be made clear that WHIMS-ECHO involves a subset of the WHIMS women. At first reading it sounds as if they are two separate studies. It also should be “We analyzed data from the Women’s .....”.

Response: We have revised this part as “We analyzed data from participants in the Women’s Health Initiative Memory Study (WHIMS). WHIMS was subsequently extended and named WHIMS-Epidemiology of Cognitive Health (WHIMS-ECHO), which involved the same group of participants who remained in the cohort during the extension of followup to the next five years.

Q3 “...Presumably the true situation is that the study involved 7,479 women aged 65-79 years at baseline who had been followed for at least one year and were not excluded for the reasons shown in Figure 1. In any case neither the numbers in the study nor the exclusion criteria are “results”. The study description should be made much clearer, both in the abstract and in the methods section.

Response: Done. The number 6,473 refers to the number of participants at baseline after applying exclusion criteria. In this study, some participants lost to follow-up within a year, which might imply substantially health condition and retention rate when compared with other participants, so we have excluded them. We have revised method and abstract accordingly.

Q4: “Other points evident in the abstract were using “C.I.” when “CI” is more normal, use of the word “fully” which with no explanation seems to imply adjustment for every conceivable variable one might think of, and misspelling “tertile” as “terile”.

Response: We have made proper correction as requested.

Q5: “There is also, here and in a number of places in the text, an unfortunate tendency to omit “the” and “a”. For example, the first sentence of the Conclusions section of the Abstract should state “Total Mg intake between the Estimated Average Requirement and the Recommended Dietary Allowances may be associated with a lower risk .....”. Or the first sentence in the Methods section should start “The WHI.....”. There are also other inaccuracies, e.g. line 4 of the data source “women that aged 65-79 years.....” has an unnecessary and incorrect “that”. Generally, parts of the paper read as if drafted by one of the authors who was not completely fluent in English. This needs attention.

Response: We have corrected these grammatical errors.

Q6: “Turning now to statistical issues, I will start with some points of detail.

Table 1 should give the number of women in each of T1, T2 and T3, and should also give the mean (SD) of magnesium intake in each group.”

Response: Done.

Q7: “Table 2 should make it clear that the results presented with 95% CI are hazard ratios. The footnotes in both Tables 2 and 3 should start “Model 1 adjustment:”.

Response: Done.

Q8: “Presenting smaller p values as  $p < 0.01$  is not very illuminating. One might at least separate out  $p < 0.001$  and  $0.001 \geq 0.01$ . Indeed, one might use a system of labelling p values \*, \*\* or \*\*\* depending on significance of 0.05, 0.01 or 0.001

Response: As recommended, we have now indicated \*/\*\*/\*\* according to the level of significance.

Q9: “I note in Table 3 that a Model 1 p of 0.03 (for T3, MCI) is wrongly unlabelled.

The text describing the results in Table 2 is somewhat confusing. It is stated that “Total Mg intake per 100mg increment was not associated with risk of MCI and/or PD in any of the Cox models assuming a linear association”. This is not true. For Model 1, there is significance for both MCI/PD and for MCI. If the authors intend to draw inferences only from Model 2 then they should make this clear.

Response: We intend to draw inferences from fully adjusted models only, and we have revised this part.

Q10: "In the same paragraph the HR of MCI/PD for tertile 2 is stated to be 0.65 in the text but 0.67 in Table 2. "

Response: We have made the correction.

Q11: "This brings me on to what I regard as the major limitation of the paper. It claims that there is a protective effect in the second but not the third tertile without properly testing whether the model fits a linear relationship. When considering a dose-response relationship, the usual procedure is to start by testing for trend and for departure from trend. If there is a significant trend but no significant departure, one can simply describe the results by the trend. If there is significant departure one can investigate the shape of it."

Response: We appreciate your comments. Indeed, we have tested for linear relation by assigning median values for quintiles, then treated it as a continuous variable in regression model. To examine non-linearity, we conducted a likelihood ratio test to compare the fit of continuous models with or without quadratic terms of Mg intake. A likelihood test with  $p < 0.05$  would suggest better fit regression model by including quadratic term.

Q12: "The analyses are also limited by use of tertiles. The authors claim that this was done "to preserve statistical power with regard to the sample size" but there were over 100 cases in each tertile in each analysis so one clearly could have split the scale more finely, perhaps into six levels.

Response: We have now used quintiles to split the scale finely as recommended.

Then one could present the HRs graphically showing them at each level, and illustrating linear relationships and, if shown to be statistically significant, non-linear relationships. Such dose-response analyses could be conducted using either the continuous variable or using the HRs and mean magnesium doses at each level.

Though I could not investigate the results fully without the detailed data, I did run some test based on the HRs (CIs) in each tertile, assigned doses of 1, 2 and 3 units (i.e. assuming that the doses of magnesium in the tertiles were equally spaced). For the MCI/PD Model 2 analysis in Table 2 I got significant heterogeneity between groups (chisquared 8.61 on 2 d.f.,  $p=0.014$ ) and significant trend (chisquared 5.77 on 1 d.f.,  $p=0.016$ ) but the difference – the departure from trend of 2.84 on 1 d.f. is not significant at  $p < 0.05$ . Similarly for the MCI Model 2 analysis in Table 2 there was again significant heterogeneity (chisquared 8.19 on 2 d.f.,  $p=0.017$ ) and trend (chisquared 6.48 on 1 d.f.,  $p=0.011$ ) but no significant departure from trend (chisquared 1.72 on 1 d.f.). On this approximate interpretation of the results I would conclude simply that risk of MCI/PD and of MCI declines with increasing total magnesium intake, but not draw inferences that the decline is only between T1 and T2 (unless a more precise dose-response analysis shows this).

Response: Thank you for your comment. We now categorized exposure into quintiles for analysis. We firstly tested for linear relationship by conducted by assigning median values for quintiles, then treated it as a continuous variable in regression model. To examine non-linearity of relationship, we conducted a likelihood ratio test to compare the fit of continuous models with or without quadratic terms of Mg intake. A likelihood test with  $p < 0.05$  would suggest better fit regression model by including quadratic term. The results from our analysis showed non-linear relationship between total Mg intake and the risk of MCI/PD and MCI.

Reviewer Name: Emily Harville

Institution and Country: Tulane University, USA

Please state any competing interests or state 'None declared': None declared

Q1: "How do the included and excluded participants compare to each other?"

Response: This is shown in Supplementary Table 1S. There were no significant differences in age, recreational physical activity, total B6 and B12 intake, prevalent of cancer, use of hormonal replacement therapy, treated high cholesterol, and family history of diabetes/heart attack/stroke.

Q2: "For the combined endpoint and MCI (which appears to be driving it), tertile 3 is not significant but the effect estimate is similar, so it can't be considered evidence of no effect or evidence of a quadratic effect. It might be appropriate to look at non-linear (quadratic, spline) effects for the continuous variable, given that the tertiles did not show a dose-response effect. Also, the study is not particularly small, so I don't find the rationale for tertiles all that convincing, which makes me wonder if the authors tried other categorizations which failed to reach statistical significance. For these reasons, I think the abstract could represent the results better."

Response: Thank you for your comment. We firstly tested for linear relationship by conducted by assigning median values for quintiles, then treated it as a continuous variable in regression model. To examine non-linearity of relationship, we conducted a likelihood ratio test to compare the fit of continuous models with or without quadratic terms of Mg intake. A likelihood test with  $p < 0.05$  would suggest better fit regression model by including quadratic term. The results from our analysis showed non-linear relationship between total Mg intake and the risk of MCI/PD and MCI

Q3: "The discussion of proportional hazards is a little odd. They say that no interaction is found, but provide p-values that go down to 0.05, then offer another example where the p-value was even lower. How did the HR change over time for this interaction?"

Response: The assumption of Cox regression models was examined via testing the interaction between exposure and survival term, and we have now examined all models using the Schoenfeld residual test but found no violation in all of our models.

## VERSION 2 – REVIEW

<b>REVIEWER</b>	Peter Lee P.N.Lee Statistics and Computing Ltd Sutton Surrey UK
<b>REVIEW RETURNED</b>	03-Jul-2019

<b>GENERAL COMMENTS</b>	<p>I thank the authors for the many changes made to the paper which have improved it very much. However, there remain a few points which require attention. A number are minor and are listed at the end of this note. The main points are as follows:</p> <p>According to the data analysis section statistical tests have been made to compare each of quintiles Q2 to Q5 with Q1, to test for linear trend (using median values for the quintiles) and to test for a non-linear relationship comparing the fit of models with or without quadratic terms). That is fine, but there are two points here.</p> <p>First Tables 2 and 3 give the results by quintile OR, and they also present another p value (e.g. 0.01 Model 1 and 0.42 Model 2 in Table 2 for MCI/PD) without saying what it is. I assume it is the p value for the linear relationship. This should be made clear in these tables, and it would also seem sensible to give the p-values for the non-linear relationship.</p> <p>Second, I would have thought that the test for linear trend is the most important test, but when describing the results in the abstract the non-significance of the linearity is not mentioned. In describing the results from Table 2 in the abstract I would say something like</p>
-------------------------	---

	<p>“For MCI/PD and MCI risks tended to be lower for quintiles Q2 to Q5 than for Q1. Though the linear trend was not significant, there was a significantly lower risk at Q3 for both MCI/PD (HR = 0.69, 95% CI 0.53-0.91) and MCI (0.63, 0.45-0.87) with similar significant reductions at Q4 and Q5 for MCI after adjustment for demographics, diet, life style, medication use and medical history. There was no association of PD with total magnesium intake.”</p> <p>Note that I would not say “every conceivable” – one can always conceive of others!</p> <p>In Table 1 it is not stated what the p value related to. Presumably this is for linear trend, but say so. Why not code the p values *, **, *** as in Tables 2 and 3?</p> <p>In the limitations, residual confounding due to inaccuracy of measurement of some adjustment variables might be mentioned. It might also be noted somewhere that for MCI/PD and MCI the Model 1 and Model 2 OR estimates were similar in magnitude, the lesser significance seen in Model 2 being due to the increased width of the 95% CI following adjustment.</p> <p>Other minor points, mainly relating to English are as follows:</p> <p>Abstract – objective Better phrased as “To examine the associations of dietary and supplemental magnesium (Mg) as assessed by food frequency questionnaire with cognitive outcomes in aging women”.</p> <p>Abstract – results – first sentence Add “1,006” after “excluding”, and delete “of them” line 3.</p> <p>Abstract – results – second sentence “During over 20 .....” not “With over 20 .....”.</p> <p>Background – para 1 – line 4 “..... with MCI, which would impair their .....”.</p> <p>Methods – data source – line 4/5 “..... was sent to women in the WHI Hormone Therapy Trial who were aged 65 to 79 years and without dementia at enrolment”.</p> <p>Methods – data source – line 9 “..... approval of all protocols was obtained .....” (not “were”).</p> <p>Methods – data source – line 5 of last paragraph Change “women with prevalent MCI/PD .....” to “had prevalent MCI/PD at baseline”, and change “and received” to “or received”.</p> <p>Methods – data analysis – para – line 15 “The test for linear relationship was conducted .....”.</p> <p>Methods – data analysis – para – line 18 “To examine the potential .....”.</p> <p>Methods – data analysis – para – line 22 “..... hence a non-linear ....”.</p> <p>Results – participant characteristics – line 3 “..... on average, a longer time.....”.</p> <p>Results – participant characteristics – line 6 “..... by the chisquared test .....”.</p> <p>Results – participant characteristics – line 6/7/8 “..... enrolled in the control group of the Estrogen+Progestin trial, with ≥ 7 alcohol drinks per week, with a history .....”.</p> <p>Results – participant characteristics – line 9 “..... education were more likely to have a higher level of Mg intake. The .....”.</p> <p>Results – participant characteristics – line 11 “..... significant for the majority of variables except for baseline age .....”.</p> <p>In Table 2 the note c should be against the heading p values not against the linear trend results.</p> <p>Para starting Table 2 – line 12 “not significant” rather than “insignificant”.</p> <p>Para starting Table 2 – last line “..... assumption of the Cox proportional .....”.</p>
--	---

	<p>Para starting Table 3 – last sentence “..... assumption of the Cox proportional .....”.</p> <p>Discussion – summary – line 10/11 “Although Mg intake only from the dietary source did not significantly associate with MCI/PD, this may be because the levels .....”.</p> <p>Discussion – comparison – line 1 “Our findings are consistent .....”.</p> <p>Conclusions – line 3 “and MCI but not in .....”.</p>
--	---

<b>REVIEWER</b>	Emily W Harville Tulane University, USA
<b>REVIEW RETURNED</b>	03-Jul-2019

<b>GENERAL COMMENTS</b>	<p>Abstract: I am somewhat amused by the idea that the authors adjusted for every conceivable variable. I am sure I could conceive of another (phase of the moon?). I suggest rewording.</p> <p>I still think there is some concern that the authors are over-interpreting statistical significance for analysis across quantiles. There is a large amount of overlap in Q3,4,5 in model 1 for table 2, for instance. While a quadratic term may have added to the predictive power, it doesn't follow that the effect estimate fell to null in the highest quintile. So it is not necessarily the case that moderate levels of Mg intake are ideal, as indicated in the discussion.</p>
-------------------------	--

### VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

According to the data analysis section statistical tests have been made to compare each of quintiles Q2 to Q5 with Q1, to test for linear trend (using median values for the quintiles) and to test for a non-linear relationship comparing the fit of models with or without quadratic terms). That is fine, but there are two points here.

First Tables 2 and 3 give the results by quintile OR, and they also present another p value (e.g. 0.01 Model 1 and 0.42 Model 2 in Table 2 for MCI/PD) without saying what it is. I assume it is the p value for the linear relationship. This should be made clear in these tables, and it would also seem sensible to give the p-values for the non-linear relationship.

Response: We appreciated for the reviewer’s suggestions. We have removed note c in Table 2 and 3. Also, we added two rows to present the p-value for trend and the p-value for non-linearity.

Second, I would have thought that the test for linear trend is the most important test, but when describing the results in the abstract the non-significance of the linearity is not mentioned. In describing the results from Table 2 in the abstract I would say something like “For MCI/PD and MCI risks tended to be lower for quintiles Q2 to Q5 than for Q1. Though the linear trend was not significant, there was a significantly lower risk at Q3 for both MCI/PD (HR = 0.69, 95% CI 0.53-0.91) and MCI (0.63, 0.45-0.87) with similar significant reductions at Q4 and Q5 for MCI after adjustment for demographics, diet, life style, medication use and medical history. There was no association of PD with total magnesium intake.” Note that I would not say “every conceivable” – one can always conceive of others!

Response: We have revised the abstract as suggested and added the result on the test for linearity.



In Table 1 it is not stated what the p value related to. Presumably this is for linear trend, but say so. Why not code the p values \*, \*\*, \*\*\* as in Tables 2 and 3?

Response: We have coded the P values as suggested and added a footnote describing the P values in Table 1.

In the limitations, residual confounding due to inaccuracy of measurement of some adjustment variables might be mentioned. It might also be noted somewhere that for MCI/PD and MCI the Model 1 and Model 2 OR estimates were similar in magnitude, the lesser significance seen in Model 2 being due to the increased width of the 95% CI following adjustment.

Response: We have added the discussion on residual confounding in limitation and mentioned the issue of wider 95%CI following adjustment in the first paragraph of discussion.

Other minor points, mainly relating to English are as follows:

Abstract – objective Better phrased as “To examine the associations of dietary and supplemental magnesium (Mg) as assessed by food frequency questionnaire with cognitive outcomes in aging women”.

Abstract – results – first sentence Add “1,006” after “excluding”, and delete “of them” line 3.

Abstract – results – second sentence “During over 20 .....” not “With over 20 .....”.

Background – para 1 – line 4 “..... with MCI, which would impair their .....”.

Methods – data source – line 4/5 “..... was sent to women in the WHI Hormone Therapy Trial who were aged 65 to 79 years and without dementia at enrolment”.

Methods – data source – line 9 “..... approval of all protocols was obtained .....” (not “were”).

Methods – data source – line 5 of last paragraph Change “women with prevalent MCI/PD .....” to “had prevalent MCI/PD at baseline”, and change “and received” to “or received”.

Methods – data analysis – para – line 15 “The test for linear relationship was conducted .....”.

Methods – data analysis – para – line 18 “To examine the potential .....”.

Methods – data analysis – para – line 22 “..... hence a non-linear ....”.

Results – participant characteristics – line 3 “..... on average, a longer time.....”.

Results – participant characteristics – line 6 “..... by the chisquared test .....”.

Results – participant characteristics – line 6/7/8 “..... enrolled in the control group of the Estrogen+Progestin trial, with  $\geq 7$  alcohol drinks per week, with a history .....”.

Results – participant characteristics – line 9 “..... education were more likely to have a higher level of Mg intake. The .....”.

Results – participant characteristics – line 11 “..... significant for the majority of variables except for baseline age .....”.

In Table 2 the note c should be against the heading p values not against the linear trend results.

Para starting Table 2 – line 12 “not significant” rather than “insignificant”.

Para starting Table 2 – last line “..... assumption of the Cox proportional .....”.

Para starting Table 3 – last sentence “..... assumption of the Cox proportional .....”.

Discussion – summary – line 10/11 “Although Mg intake only from the dietary source did not significantly associate with MCI/PD, this may be because the levels .....”.

Discussion – comparison – line 1 “Our findings are consistent .....”.

Conclusions – line 3 “and MCI but not in .....”.

Response: Thank you so much for the detailed suggestions regarding English language. We have addressed all the issues that you pointed out. In addition, we have removed note c in Table 2 and 3 because we have added a new line to present p-value for trend.

Reviewer: 2

Abstract: I am somewhat amused by the idea that the authors adjusted for every conceivable variable. I am sure I could conceive of another (phase of the moon?). I suggest rewording.

Response: We have revised the wording as ‘adjustment for demographic characteristics, diet, lifestyle, medication use and medical history’.

I still think there is some concern that the authors are over-interpreting statistical significance for analysis across quantiles. There is a large amount of overlap in Q3,4,5 in model 1 for table 2, for instance. While a quadratic term may have added to the predictive power, it doesn't follow that the effect estimate fell to null in the highest quintile. So it is not necessarily the case that moderate levels of Mg intake are ideal, as indicated in the discussion.

Response: Thank you for the comment. Although further increment of Mg intake did not provide additional benefit for preventing MCI/PD, comparing with the lowest quintile, the fourth and fifth quintile of total Mg intake associated with lower risk of MCI. We have added this point in the section of ‘Comparison with previous literature’.

### VERSION 3 – REVIEW

<b>REVIEWER</b>	Peter Lee P N Lee Statistics and Computing Ltd Sutton, Surrey, UK
<b>REVIEW RETURNED</b>	31-Jul-2019
<b>GENERAL COMMENTS</b>	I am happy with the scientific content of the paper. I only have a very few comments on the English and presentation. 1. Abstract line 4 - change “.. which extended and named WHIMS..” to “.. subsequently extended and named the WHIMS..”. 2. Discussion First paragraph line 8 - change “..Model 2 being might due to..” to “..Model 2 possibly being due to..”

	<p>3. Discussion paragraph 1 last sentence - change ".. because the levels of Mg intake from dietary source was lower than the sum of dietary and supplemental source." to "..because the level of Mg intake from diet was lower than that from diet and supplementary sources combined."</p> <p>4. It seems strange in the tables to have so many cases of "&lt;0.01*** " where *** means &lt;0.001". Perhaps change these to "&lt;0.001*** " or alternatively simple give p values as *, **, *** or NS (not significant = <math>p \geq 0.05</math>), without actually giving the exact p value.</p>
--	---

### VERSION 3 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Peter Lee

Institution and Country: P N Lee Statistics and Computing Ltd, Sutton, Surrey, UK

Please state any competing interests or state 'None declared': None declared

I am happy with the scientific content of the paper. I only have a very few comments on the English and presentation.

1. Abstract line 4 - change ".. which extended and named WHIMS.." to ".. subsequently extended and named the WHIMS..".
2. Discussion First paragraph line 8 - change "..Model 2 being might due to.." to "..Model 2 possibly being due to..".
3. Discussion paragraph 1 last sentence - change ".. because the levels of Mg intake from dietary source was lower than the sum of dietary and supplemental source." to "..because the level of Mg intake from diet was lower than that from diet and supplementary sources combined."

Response: We have revised the parts that you have mentioned.

4. It seems strange in the tables to have so many cases of "<0.01\*\*\* " where \*\*\* means <0.001". Perhaps change these to "<0.001\*\*\* " or alternatively simple give p values as \*, \*\*, \*\*\* or NS (not significant =  $p \geq 0.05$ ), without actually giving the exact p value.

Response: We have revised the Tables as requested.