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)	Vitamin D status in an Australian patient population – a large retrospective case series focusing on factors associated with variations in serum 25(OH)D
AUTHORS	Voo, Veronica Tsin Fong; Stankovich, Jim; O'Brien, Terence; Butzkueven, Helmut; Monif, Mastura

VERSION 1 – REVIEW

REVIEWER	Tom Hill Newcastle University
REVIEW RETURNED	22-Aug-2019

GENERAL COMMENTS	This is a powerful study which reports vitamin D status data over a 5 year period in a clinical setting. The robust sampling frame is a strength of the study. Overall it appears that vitamin D deficiency is low but vitamin D insufficiency as defined by 25-50 nmol/L is significant (>20%). There is no doubt that the wealth of available 25OHD data adds hugely to the uniqueness of the study but I felt more could have been achieved in the analysis. Has vitamin D status changed over time? What are the public health significance of the findings in Australia? Did anyone exceed the IOM upper level for 25OHD of 125 nmol/L? Please see the attached paper from Ireland which might be of interest McKenna et al 2015. Rising trend in vitamin D status from 1993 to 2013: dual concerns for the future. Endocr Connect. 2015 Sep;4(3):163-71. doi: 10.1530/EC-15-0037. Epub 2015 Jun 1. No information is provided on the lab assay for 25OHD. Was it standardized? Was there external QC with DEQAS? How confident are the authors with the performance of their assay? Assay variation is measuring 25OHD is a major problem in vitamin D science. A discussion of these points are essential. See Binkley and Carter, 2017. Toward Clarity in Clinical Vitamin D Status Assessment: 25(OH)D Assay Standardization. Endocrinol Metab Clin North Am. 2017 Dec;46(4):885-89 How does one interpret the terms 'deficiency', insufficiency', sub-optimal' and 'sufficiency'? What is the (I) clinical and (ii) biological meaning/implications for each state? Why was the data not assessed for normality? Please add P values to table 3.

	Mush of the literature in the introduction is old eg references 2,3 and 4 are over 12 years old and there are more recent references available. More should be made of the vitamin D RCTs over the last 10-15 years.

REVIEWER	Davoud Vahabzadeh Ilam University of medical sciences, Iran
REVIEW RETURNED	03-Sep-2019

GENERAL COMMENTS	1- The title is too long and required to be edited. 2- Line 13 in abstract required to be edited. 3- In line 14 in abstract results CI with P value has not presented well. 4- "In males were lower than in females, increased with age, higher in neurologic patients"?!?!. Likely due to higher supplementation. Not enough justification and scientific rationale has presented for some of these results. 5- Some results seem quietly different, and it should be considered that some results are in counteracting with most previous studies because of some possibly biases. 6- Some references are too old. (refs no 12, 25, 34, 35, 36 and 43)

REVIEWER	Rosemary L Schleicher, PhD Centers for Disease Control and Prevention Atlanta, Georgia USA
REVIEW RETURNED	05-Sep-2019

GENERAL COMMENTS	Overall, this is an interesting paper. Reviewer's comments follow: **General Comments** 1. It seems strange that no mention of race-ethnicity is made in this paper. I expected to see 25OHD concentrations stratified by race-ethnicity. 2. "Monthly" doesn't seem to add any additional information to the analysis. Season seems adequate. 3. Are there any differences between inpatient and outpatient vitamin D status? Some mention should be made about this variable. 4. Children are mentioned in the Conclusions, but they are never defined (age range) in the study population. Perhaps consider dropping this small (n=379), potentially diverse (infants, toddlers, preschool, school-age, pre-teen, teenage) group of <20y olds. 5. Laboratory methods for 25-hydroxyvitamin D should be mentioned in the Methods section. Were any reference materials for 25OHD tested during 2013-2017? Did the laboratory participate in any 25OHD proficiency testing during this time? If so, assay bias information and/or proficiency testing information (such as "passed") should be provided. 6. Too much is made of the medical specialty that handled (admitted or treated as an outpatient) these patients. I

	don't think that the data are interesting enough to include them in each and every table. Especially as the connection between vitamin D and multiple sclerosis is speculative. For the amount of attention given to the neurology patients in the results, it not given adequate depth in the Discussion. 7. The comment in the Discussion that inter-assay variability from the two assays used in the lab was "minimal" is not well-supported. One would need to know the number of samples tested using the two methods, and the degree of scatter of the data to draw this conclusion. **Specific Comments** 1. It would help the reviewers to number the pages. The line numbers don't line-up with the text. 2. "Vitamin D level" should be changed to "25-hydroxyvitamin D concentration" or an abbreviated form of this such as "25OHD concentration" or "vitamin D status." Vitamin D, *per se*, was not measured. 3. Abstract Results (line 15): *p*-value is given. What hypothesis is being tested? 4. Strengths and limitations of this study (line 10): "geographic locations" seems like an over-reach. Aren't all the data from one location? 5. Introduction (lines 10-14): Holick's paper (reference 5) seems to use <20 ng/mL as the cutoff for risk of deficiency. This follows your sentence that "vitamin D deficiency is generally defined as <25 nmol/L," which is <10 ng/mL. You might not want to juxtapose these two sentences. 6. Introduction (line 26): the sentence about vitamin D-containing foods seems out of place. 7. Results (lines 6-8): *p*-values are given. What hypotheses are being tested? 8. It would be interesting to see Table 3 stratified by age. 9. Conclusion (line 51): "Our study …previously thought." The meaning of this sentence as it relates to this study is unclear.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Tom Hill

1. **Please add P values to table 3.**

   P-values have been added to table 3 in the revised version.

2. **No information is provided on the lab assay for 25OHD. Was it standardized? Was there external QC with DEQAS? How confident are the authors with the performance of their assay? Assay variation is measuring 25OHD is a major problem in vitamin D science.**

   We take this comment on board. Information about the lab assay used for 25(OH)D measurements has been included in the revised version (page 6, line 9-17).

3. **How does one interpret the terms 'deficiency', insufficiency', sub-optimal' and 'sufficiency'? What is the (I) clinical and (ii) biological meaning/implications for each state?**

   Currently there is no complete consensus on 25(OH)D values to define these clinical terms, hence the interpretation of the terms vitamin D 'deficiency', 'insufficiency', 'sub-optimal' and 'sufficiency' varies and depends on the clinical context. However, vitamin D deficiency (which is generally defined as levels <25nmol/L) is well known to increase the risk of rickets and osteoporosis. The optimal level of vitamin D for non-skeletal chronic diseases such as autoimmunity, cardiovascular diseases, multiple sclerosis, etc. is yet to be defined, and there are various ranges/levels to define 'sufficiency' in the literature. Future research may elucidate different requirement levels for vitamin D in relation to different non-skeletal outcomes. We have discussed this issue in the revised version (page 10, line 34-38).

4. **Much of the literature in the introduction is old eg references 2,3 and 4 are over 12 years old and there are more recent references available. More should be made of the vitamin D RCTs over the last 10-15 years.**

   More recent references have been added to the revised version.

5. **Why was the data not assessed for normality?**

   The histogram in Figure 1 shows that vitamin D measurements are approximately normally distributed. Though the measurements are slightly skewed to the right, log transformation resulted in greater skew to the left. Thus, we decided to proceed with an analysis of untransformed vitamin D measurements, as transformations other than a log transformation would make the results harder to interpret.

6. **Has vitamin D status changed over time? What are the public health significance of the findings in Australia? Did anyone exceed the IOM upper level for 25OHD of 125 nmol/L? Please see the attached paper from Ireland which might be of interest McKenna et al 2015. Rising trend in vitamin D status from 1993 to 2013: dual concerns for the future. Endocr Connect. 2015 Sep;4(3):163-71. doi: 10.1530/EC-15-0037. Epub 2015 Jun 1.**

   According to our findings, there was a significant increase of serum 25(OH)D levels from 2013 compared to subsequent years (2014-2017). Although 25(OH)D levels were higher in 2014-2017 compared to 2013, levels were found to slightly reduce each subsequent year from 2014-2017, suggesting a downward trend.

As for public health significance, our findings suggest that other interventions are required to boost vitamin D status during winter as our results showed that despite higher prescription of supplementation during winter compared to other seasons, 25(OH)D remained the lowest during winter. Our findings also suggest that those who are below 20 years could be more at risk of vitamin D deficiency, hence interventions to boost vitamin D status should focus on this sub-group.

A small number of patients had 25(OH)D levels that exceeded 125nmol/L. Patients who had levels higher than 200nmol/L were excluded during analysis.

Thank you for the paper, we have included it in the revised version.

7. **A discussion of these points are essential. See Binkley and Carter, 2017. Toward Clarity in Clinical Vitamin D Status Assessment: 25(OH)D Assay Standardization. Endocrinol Metab Clin North Am. 2017 Dec;46(4):885-89**

   Discussion of these points has been added to the revised version (page 10, line 34-38).

Reviewer: 2
Reviewer Name: Davoud Vahabzadeh

1. **The title is too long and required to be edited.**

   We take this comment on board. The new title in the revised version is “Vitamin D status in an Australian patient population – a large retrospective case series focusing on factors that influence variations in serum 25(OH)D”.

2. **Line 13 in abstract required to be edited.**

   Line 13 in abstract has been changed to “Main outcome measures: Serum 25(OH)D levels stratified according to patients’ age, gender and medical specialty admitted to, as well as month, season and year (2013-2015) 25(OH)D was measured” in the revised version.

3. **In line 14 in abstract results CI with P value has not presented well.**

   We take this comment on board. The insertion of this P-value was a mistake and has been removed in the revised version.

4. **"In males were lower than in females, increased with age, higher in neurologic patients"?!?!. Likely due to higher supplementation. Not enough justification and scientific rationale has presented for some of these results.**

   We take this comment on board and have extended our discussion on this in the revised version. The conclusion that serum 25(OH)D levels were higher in female patients, older age and in Neurology patients is based on the results of our analysis of 38,385 patients. We suggested that this could be due to supplementation as our data showed prescription of vitamin D supplementation increased with increasing age, hence is likely to result in higher serum 25(OH)D levels measured in patients with older age. As for female patients and Neurology patients, because vitamin D supplements could be purchased without a prescription and in various sources, it is plausible that patients with high 25(OH)D levels are on supplementation that is not included in the prescription records. Of course there are other

factors that could influence serum 25(OH)D levels, but due to limitation of the data collected, we could only speculate that the higher serum 25(OH)D was influenced by unrecorded supplementation. We hope this clarifies your question.

8. **Some results seem quietly different, and it should be considered that some results are in counteracting with most previous studies because of some possibly biases.**

   We take this comment on board. We are aware that due to our patient population, our findings might not represent the general population. Moreover, our findings could be due to particular characteristics of patients who attend a tertiary health care centre, who are also likely to have comorbidities. These factors are possible biases that could influence the findings presented in our study.

9. **Some references are too old. (refs no 12, 25, 34, 35, 36 and 43)**
   New references have been added to the revised version.

Reviewer: 3
Reviewer Name: Rosemary L Schleicher

1. **Co-author's name mismatch: The author "Butzkueven, Helmut" in your main document is registered as "Butkueven, Helmut" in ScholarOne. Please ensure that the author has same registered name.**

   The name 'Helmut Butzkueven' as stated in the manuscript is the correct spelling. We have notified this co-author of the misspelling of his name in ScholarOne.

2. **Co-author's complete affiliations in SC1: Please provide complete affiliations (institutions and department) for all the authors both in the manuscript and in the submission system (ScholarOne).**

   We have added street addresses for all the author affiliations in the manuscript. We have also notified all authors to alter SC1 affiliations accordingly to match what is on the manuscript.

We have referred to the STROBE statement checklist of items that should be included in reports of **cross-sectional studies** to make sure all items are included in our manuscript.

VERSION 2 – REVIEW

REVIEWER	Rosemary L Schleicher, PhD Centers for Disease Control and Prevention Atlanta, Georgia USA
REVIEW RETURNED	25-Sep-2019

GENERAL COMMENTS	Reviewer's reply to General Comments on original manuscript 1. I suggest mentioning earlier in the paper (Methods) that race-ethnicity information is unavailable. 2. I don't see that monthly information adds to the seasonal information. Authors did not specify anything gained by display of monthly data. Perhaps a supplemental on-line figure could show this.

	3. I suggest mentioning earlier in the paper (Methods) that inpatient/outpatient information is unavailable. 4. Children should be deleted from the study because they are only 1% of the participants (n=379 in Table 1), they have the highest rate of deficiency (n=40 children in Table 3), and only 2 children were prescribed supplements (Table 4). Conclusions (p. 11) were drawn about the risk for vitamin D deficiency for those not taking vitamin D supplements, especially those below 20 years of age (again, n=2 in Table 4). This conclusion is not warranted, considering how few children were prescribed supplements. 5. Still no mention of using/not using reference materials or participating in proficiency testing. 6. Medical specialty data are not that interesting, other than in passing. They don't prove anything considering that real supplementation usage is not known for any group. According to Table 3, vitamin D supplementation information, i.e., prescription dosing, is missing for 94% of the study participants. 7. In the methods, I suggest removing an item for discussion, namely, "Hence, inter-assay … level measured." Reviewer's New Comments on revised manuscript 1. The new title is now suggesting cause and effect when only associations were noted. 2. The discussion of patients with neurological conditions seems muddy (pp. 9-10). It seems that the only conclusions that make sense are higher 25(OH)D compared to other patients and lower rate of prescribed vitamin D compared to other patients. 3. Under Limitations on p. 10, it is incorrect to indicate that vitamin D assays are not standardized. Standardization activities were ongoing 2013-2017. Reference materials were available from NIST (https://www-s.nist.gov/srmors/view_detail.cfm?srm=972A) and certification of 25(OH)D results was possible through participation in a world-wide program (https://www.cdc.gov/labstandards/pdf/hs/CDC_Certified_Vitamin_D_Procedures-508.pdf). 4. "Admitted" seems to be used for all patients. Is this appropriate? 5. The last item in the "Strengths and Limitations of this study" (p. 4) should include not only that the quantity of prescription vitamin D was not recorded but that information about usage of prescription and non-prescription vitamin D was largely unavailable.

REVIEWER	Joe Nolan Northern Kentucky University USA
REVIEW RETURNED	02-Nov-2019

GENERAL COMMENTS	I am classifying this as major revision for a singular reason (it would otherwise be minor). That reason is the assay change between Feb. 2014 and Mar 2014. While the authors have suggested they believe the assays to be identical, I am not at all convinced of this. While the authors have proposed potential reasons for variation across years, the numbers provided do not convince me that these could explain why the 25(OH)D levels from 2013 were so much lower. The change in assay absolutely could explain the difference, and I remain suspicious that it probably explains much of it. My recommendation is that analysis be conducted using only data from March 2014 through 2017 (earlier data omitted). This should still maintain sufficient sample size, and may well result in any number of changes to the rest of the analysis (such changes in any analyses that did not originally

	account for "year" would essentially confirm that this is an issue). The other alternative would be to provide more concrete proof that the assays are identical, but this would take a comparison sample of much larger than n=80 to do. Please note: I have been asked to review the statistical methods and presentation, and my comments will be limited to that scope. My comments are as follows: 1. In the abstract, i assume these are 95% CI's. This should be stated (e.g. (95% CI: 5.02 to 6.33)). As CI's are included and far more valuable than p-values, the p-values provide no additional information and should be omitted. 2. Abstract line 28 and Line 32: What is the relevance of the observation that 2013 was significantly lower? Year to year variability doesn't seem important enough to address in the abstract. Follow-up (Page 6, lines 16-28) - in reading about the laboratory assessments I note a process change in March 2014. While you have some assessment of the bias present, it is difficult to tell what you did and I don't know that I buy it as evidence that "interassay variability is small and the change in assay is unlikely to cause a significant shift in the 25(OH)D level measured". Perhaps a better alternative would be to exclude the 2013 and early 2014 data. A third alternative would be to apply some sort of correction factor to the early data, based on the relationship noted from the 80 duplicate samples. Treating the data as if they are the same assay does not seem supported here. **It seems unlikely to me that your 2013 difference is associated to anything other than the change in the assay.** (If there is some other explanation, it should be clearly provided). More follow-up (page 10 line 36-39): "However comparison of the new assay to the previous one demonstrated a slope = 1 and intercept = 1nmol/L, i.e. inter-assay variability is likely to be minimal, hence the change in assay is unlikely to cause a shift in the vitamin D level measured." Such cannot be demonstrated without providing a confidence interval for the slope. It seems likely that the standard error for the slope estimate will be large enough that one cannot exclude the possibility of differences in assay. A less than 4% increase in supplements seems unlikely to raise the average serum 25(OH)D by the amount that seems to be indicated. **My recommendation remains to use only data from March 2014 on.** This should still yield a very large retrospective sample. Doing otherwise may be confounding the study in unknown ways. If you do this I would expect "year" to not be statistically significant (and if it is, then you have a better set of evidence from which to speak about other possible contributing factors. 3. (p6 line 56 and following): Confidence intervals are far more useful than P-values. I'm happy to see you including

	them. When confidence intervals are given, p-values may be omitted as they do not provide any additional information. Conversely, confidence intervals provide more (useful) information than p-values and so should be included whenever possible. 4. (p7 lines 20 and following): It seems that you have constructed several univariable analyses. This opens the possibility that some of your conclusions are confounded with one another. To assess this, compare the univariable results with your multiple regression analyses. Look at 95% confidence intervals for slopes. If those intervals are similar across models, then you have a valid result. If they vary substantially, this is evidence that some of your variables are confounded.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 3
Reviewer Name: Rosemary L Schleicher

Reviewer's reply to General Comments on original manuscript:
1. I suggest mentioning earlier in the paper (Methods) that race-ethnicity information is unavailable.

We take this comment on board. The information has been included in the Methods section in the revised version (page 6, line 24-26).

2. I don't see that monthly information adds to the seasonal information. Authors did not specify anything gained by display of monthly data. Perhaps a supplemental on-line figure could show this.

We take this comment on board. Monthly information has been removed throughout the revised version.

3. I suggest mentioning earlier in the paper (Methods) that inpatient/outpatient information is unavailable.

Thank you for your suggestion. The information has been included in the Methods section in the revised version (page 6, line 26-27).

4. Children should be deleted from the study because they are only 1% of the participants (n=379 in Table 1), they have the highest rate of deficiency (n=40 children in Table 3), and only

2 children were prescribed supplements (Table 4). Conclusions (p. 11) were drawn about the risk for vitamin D deficiency for those not taking vitamin D supplements, especially those below 20 years of age (again, n=2 in Table 4). This conclusion is not warranted, considering how few children were prescribed supplements.

We take this comment on board. Information about children has been removed throughout the revised version, and conclusions for our study has been modified to reflect this change.

5. Still no mention of using/not using reference materials or participating in proficiency testing.

We appreciate your suggestion. We have included information on reference materials and proficiency testing in the revised version (page 6, line 10-14).

6. Medical specialty data are not that interesting, other than in passing. They don't prove anything considering that real supplementation usage is not known for any group. According to Table 3, vitamin D supplementation information, i.e., prescription dosing, is missing for 94% of the study participants.

Thank you for your suggestion. The lack of supplementation data applies to the entire cohort and irrespective of supplementation, it would still be relevant to see if certain medical specialties have significantly higher values of serum 25(OH)D. Given the wider interest of various medical communities in vitamin D status, it would still be relevant to have the current data as it is with plan for future prospective studies to look into supplementation in more detail.

7. In the methods, I suggest removing an item for discussion, namely, "Hence, inter-assay … level measured."

We take this comment on board. The item has been removed in the revised version.

Reviewer's New Comments on revised manuscript:

1. The new title is now suggesting cause and effect when only associations were noted.

The title has been modified in the revised version. The new title is "Vitamin D status in an Australian patient population – a large retrospective case series focusing on factors associated with variations in serum 25(OH)D".

2. The discussion of patients with neurological conditions seems muddy (pp. 9-10). It seems that the only conclusions that make sense are higher 25(OH)D compared to other patients and lower rate of prescribed vitamin D compared to other patients.

Thank you for your comment. We found that patients from Neurology had higher vitamin D values compared to other specialties. We have attempted to address the potential reasons for this in our discussion. Ultimately, as our large set of data was entirely retrospective, certain assertions might seem less plausible than others but future prospective studies could address this more accurately.

3. Under Limitations on p. 10, it is incorrect to indicate that vitamin D assays are not standardized. Standardization activities were ongoing 2013-2017. Reference materials were available from NIST (https://www-s.nist.gov/srmors/view_detail.cfm?srm=972A) and certification of 25(OH)D results was possible through participation in a world-wide program (https://www.cdc.gov/labstandards/pdf/hs/CDC_Certified_Vitamin_D_Procedures-508.pdf).

We take this comment on board. Information regarding vitamin D assay standardisation has been removed in the revised version.

4. "Admitted" seems to be used for all patients. Is this appropriate?

We take this comment on board. With the word "admitted", it does not specify only inpatients but describe inpatients and outpatients. We have now clarify this in the revised version (page 6, line 28-29).

5. The last item in the "Strengths and Limitations of this study" (p. 4) should include not only that the quantity of prescription vitamin D was not recorded but that information about usage of prescription and non-prescription vitamin D was largely unavailable.

Thank you for your suggestion. We have included unavailable information on total vitamin D intake of patients as one of the limitations of this study in the revised version (page 4, line 11-12).

Reviewer: 4
Reviewer Name: Joe Nolan

Please leave your comments for the authors below:
I am classifying this as major revision for a singular reason (it would otherwise be minor). That reason is the assay change between Feb. 2014 and Mar 2014. While the authors have suggested they believe the assays to be identical, I am not at all convinced of this. While the authors have proposed potential reasons for variation across years, the numbers provided do not convince me that these could explain why the 25(OH)D levels from 2013 were so much lower. The change in assay absolutely could explain the difference, and I remain suspicious

that it probably explains much of it. My recommendation is that analysis be conducted using only data from March 2014 through 2017 (earlier data omitted). This should still maintain sufficient sample size, and may well result in any number of changes to the rest of the analysis (such changes in any analyses that did not originally account for "year" would essentially confirm that this is an issue). The other alternative would be to provide more concrete proof that the assays are identical, but this would take a comparison sample of much larger than n=80 to do. Please see attached file for additional details and comments.

We take this comment on board. As suggested by the reviewer, we now only include data from March 2014 onwards (earlier data omitted) in the revised version.

Please note: I have been asked to review the statistical methods and presentation, and my comments will be limited to that scope. My comments are as follows:

1. In the abstract, i assume these are 95% CI's. This should be stated (e.g. (95% CI: 5.02 to 6.33)). As CI's are included and far more valuable than p-values, the p-values provide no additional information and should be omitted.

Thank you for your suggestion. We have removed the p-values and stated the 95% confidence intervals as “95% CI” in the revised version (page 3, line 11-18).

2. Abstract line 28 and Line 32: What is the relevance of the observation that 2013 was significantly lower? Year to year variability doesn't seem important enough to address in the abstract. Follow-up (Page 6, lines 16-28) - in reading about the laboratory assessments I note a process change in March 2014. While you have some assessment of the bias present, it is difficult to tell what you did and I don't know that I buy it as evidence that "inter-assay variability is small and the change in assay is unlikely to cause a significant shift in the 25(OH)D level measured". Perhaps a better alternative would be to exclude the 2013 and early 2014 data. A third alternative would be to apply some sort of correction factor to the early data, based on the relationship noted from the 80 duplicate samples. Treating the data as if they are the same assay does not seem supported here. It seems unlikely to me that your 2013 difference is associated to anything other than the change in the assay. (If there is some other explanation, it should be clearly provided). More follow-up (page 10 line 36-39): “However comparison of the new assay to the previous one demonstrated a slope = 1 and intercept = 1nmol/L, i.e. inter-assay variability is likely to be minimal, hence the change in assay is unlikely to cause a shift in the vitamin D level measured.” Such cannot be demonstrated without providing a confidence interval for the slope. It seems likely that the standard error for the slope estimate will be large enough that one cannot exclude the possibility of differences in assay. A less than 4% increase in supplements seems unlikely to raise the average serum 25(OH)D by the amount that seems to be indicated. My recommendation remains to use only data from March 2014 on. This should still yield a very large retrospective sample. Doing otherwise may be confounding the study in unknown ways. If you do this I would expect “year” to not be statistically significant (and if it is, then you have a better set of evidence from which to speak about other possible contributing factors.

Thank you for your suggestion, we take your comment on board. We have omitted the data from 2013 to February 2014 in the revised version. Given the wider interest of various medical communities in vitamin D status, it would be relevant to include the yearly variability in vitamin D

levels in the abstract as it is. As for the statistical significance, after omitting the data from 2013 to February 2014 as suggested, there is still significant variation by year of measurement, though not as large as before. Our results now show that vitamin D levels in 2016 and 2017 were significantly lower than in 2014 (page 7, line 37-38, Figure 5, see also the table below).

3. (p6 line 56 and following): Confidence intervals are far more useful than P-values. I'm happy to see you including them. When confidence intervals are given, p-values may be omitted as they do not provide any additional information. Conversely, confidence intervals provide more (useful) information than p-values and so should be included whenever possible.

We take this comment on board. P-values are removed when confidence intervals are provided in the revised version.

4. (p7 lines 20 and following): It seems that you have constructed several univariable analyses. This opens the possibility that some of your conclusions are confounded with one another. To assess this, compare the univariable results with your multiple regression analyses. Look at 95% confidence intervals for slopes. If those intervals are similar across models, then you have a valid result. If they vary substantially, this is evidence that some of your variables are confounded.

Thank you for your suggestion. We have compared both models and included the table below for your reference. Because Table A (below) did not take into account vitamin D supplementation, the values of the multivariable analyses presented differ from Table 2 (manuscript). We did not include vitamin D supplementation in Table A because the univariable analyses did not account for supplementation. While there are some differences in point estimates and confidence intervals between the univariable and multivariable analyses, particularly for age and medical specialty, there are strong associations with age and specialty in both analyses.

Table A

Variable	Univariable analysis Difference (95% CI)	Multivariable analysis Difference (95% CI)
Sex		
Female	Reference	Reference
Male	-6.1 (-6.9, -5.4)	-6.0 (-6.7, -5.2)
Age		
20-29	Reference	Reference
30-39	3.0 (1.4, 4.6)	1.5 (-0.1, 3.0)
40-49	3.4 (1.8, 4.9)	2.3 (0.7, 3.8)
50-59	6.2 (4.7, 7.8)	5.5 (4.0, 7.0)

60-69	8.4 (6.9, 9.9)	8.7 (7.2, 10.2)
>70	9.9 (8.6, 11.2)	11.9 (10.5, 13.2)
Medical specialty		
NEUR	Reference	Reference
AMU	-7.8 (-10.7, -4.8)	-12.0 (-15.0, -9.0)
BOE	-5.5 (-8.8, -2.3)	-9.8 (-13.0, -6.7)
EMER	-7.1 (-10.1, -4.1)	-12.0 (-15.0, -9.1)
ENDO	-2.8 (-5.9, 0.3)	-4.5 (-7.5, -1.5)
GAST	-9.9 (-12.9, -6.9)	-9.5 (-12.5, -6.6)
NEPH	-12.6 (-15.7, -9.5)	-13.4 (-16.5, -10.4)
ORTH	-8.9 (-12.0, -5.7)	-14.8 (-18.0, -11.6)
OTHER	-8.8 (-11.4, -6.2)	-10.6 (-13.2, -8.1)
PRIV	-1.3 (-3.9, 1.4)	-2.4 (-5.0, 0.2)
Season		
Winter	Reference	Reference
Spring	1.3 (0.3, 2.3)	1.2 (0.2, 2.2)
Summer	11.2 (10.1, 12.3)	11.4 (10.3, 12.5)
Autumn	9.1 (8.1, 10.1)	8.8 (7.8, 9.8)
Year		
2014	Reference	Reference
2015	-0.2 (-1.3, 0.8)	-1.3 (-2.4, -0.3)
2016	-2.9 (-3.9, -1.8)	-3.6 (-4.6, -2.5)
2017	-3.8 (-4.8, -2.7)	-4.0 (-5.1, -3.0)

VERSION 3 – REVIEW

REVIEWER	Joe Nolan Northern Kentucky University, USA
REVIEW RETURNED	24-Dec-2019

GENERAL COMMENTS	Returning a response to reviewers would have been desirable as it saves reviewer time. Based on my review of the resubmitted manuscript, it appears that my primary concern in the initial review, namely the change in assay / measurement system mid-stream, has been satisfied by using only data from 2014 on.

VERSION 3 – AUTHOR RESPONSE

Reviewer: 4
Reviewer Name: Joe Nolan

Please leave your comments for the authors below:

That reason is the assay change between Feb. 2014 and Mar 2014. While the authors have suggested they believe the assays to be identical, I am not at all convinced of this. While the authors have proposed potential reasons for variation across years, the numbers provided do not convince me that these could explain why the 25(OH)D levels from 2013 were so much lower. The change in assay absolutely could explain the difference, and I remain suspicious that it probably explains much of it. My recommendation is that analysis be conducted using only data from March 2014 through 2017 (earlier data omitted). This should still maintain sufficient sample size, and may well result in any number of changes to the rest of the analysis (such changes in any analyses that did not originally account for "year" would essentially confirm that this is an issue). The other alternative would be to provide more concrete proof that the assays are identical, but this would take a comparison sample of much larger than n=80 to do. Please see attached file for additional details and comments.

We take this comment on board. As suggested by the reviewer, we now only include data from March 2014 onwards (earlier data omitted) in the revised version. "