PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Leisure-time physical activity and circulating 25-hydroxyvitamin D levels in cancer survivors, a cross-sectional analysis using data from the US National Health and Nutrition Examination Survey
AUTHORS	Yang, Lin; Toriola, Adetunji

VERSION 1 - REVIEW

REVIEWER	Anna-Maija Tolppanen University of Eastern Finland, Finland
REVIEW RETURNED	06-Feb-2017

GENERAL COMMENTS	The manuscript describes the association between leisure-time physical activity and 25OHD-levelsamong the cancer survivors of NHANES participants from 2001-2010. In general, the manuscript is easy to follow and the statistical methods are mainly appropriate but the following points should be taken into account: 1. Abstract& methods: please change multiple linear regression to multivariable 2. Abstract: please indicate clearly what the numbers of the result section indicate (differences between LTPA categories). 3. Is it necessary to use the term statistically significant? The results demonstrate when the 95% Cis indicate a difference, and if the authors want to bring in the significance, it would be much more helpful to indicate clinical significance. 4. To continue with my previous comment, the authors could add some discussion about the clinical significance of results. Is approximately 10nmol/L difference clinically meaningful? Any studies showing that this kind of increase would manifest as health benefits? How much of it is perhaps due to confounding (ability to conduct more intensive LTPA could be a marker for overall better health status/socioeconomic position or something else)? It would be good to discuss these alternative explanations. 5. The background section: the authors relyr heavily to their own publication on 25OHD levels and cancer prognosis as evidence on 25OHD levels being associated with better prognosis and survival, and 25OHD levels being the best indicator of vitamin status. Although this is a systematic review, it would be good to include more recent individual studies (this also may not be the best possible reference for vitD status biomarkers). 6.It is unclear how “understanding the association between physical activity and vitamin D could inform cancer survivorship care strategies ”. This isbetter suited to discussion than the aims section where it is currently located. In addition, the authors may want to clarify why this is the case. 7. Exclusion criteria (p.6) “excluded those who were never diagnosed withcancer, and were pregnant.” should “or” be used instead of “and”.

	8. The authors state that participants who received chemotherapy within last month were excluded from blood collection. The implications of this should be discussed, as well as the time since cancer diagnosis and the resulting consequences to the generalizability and implications of the results. Is it so that the 25OHD levels promote better survival, or is it so that the results are more indicative of better physical condition and health status (ability to perform more intensive LTPA), which is proxied by higher 25OHD levels. 9. To continue, it would be good to draw a flow chart how the study population weas arrived at. 10. The explanation of continuous& categorical analyses should be moved from p.7 to Statistical analysis section. 11. on page 9, the authors mention that they classified activities that could be indoor/outdoor to indoor to ensure more conservative estimation. It would be interesting to see how robust the results were, i.e. how different the results would be if they were classified as outdoor, or 50/50. 12. Please add the numbers to “sufficiently active (≥MET-min/week” 13. It is not evident why the authors doubled the time spent in vigorous activities. If the intensity impacts 25OHD levels, I would prefer to see an explanation for this in the introduction. Is it more likely that the sunlight exposure, defined more by time spent outdoors than intensity matters? If they wanted to assess the impact of LTPA, why not conduct the analyses separately for vigorous and moderate instead of doubling the time? If they prefer the approach they chose, it should be justified. 14. I did not quite understand the sentence “In the multiple linear regression models, we simultaneously adjusted for both indoor and outdoor activities, provided they were significantly different (P value<0.001).” Does it mean that they were added to the same model and if P<0.001, then they were both kept in? I would expect these variables to be highly correlated, and including both of them in the same model would likely lead to a situation when one of them no longer is related to outcome (and this also has implications on interpreting the individual coefficients –they are no longer interpretable). Please clarify what was actually done and how collinearity was accounted for. This collinearity may also explain why the association of indoor 25OHD levels disappeared in the mutually adjusted model. 15. Why LTPA was analysed as categorized variable, instead of continuous one? 16. The authors have made a good start in discussing the possible limitations. However, they should discuss the implications of these to results (especially unmeasured confounding, and time since diagnosis). 17. The authors say that bias arising from the lack of time since diagnosis is likely nondifferential. Wouldn’t this be actually likely differential between LTPA categories? Wouldn’t it be more likely that in general, those with longer time since dg are able to participate in more demanding LTPA 18. “Our findings suggest that 25-OHD might be a surrogate marker of physical activity that accounts for the direct and indirect effects of LTPA, particularly outdoor.” please add references to previous papers that have reported the same. 19. While I wholeheartedly agree that physical activity among cancer survivors should be promoted, I don’t interpret the findings of this study as evidence on prioritizing outdoor activities with potential benefits of sunlight exposure. The authors may want to reconsider this, as their findings did not actually account for “sun time”, but

	emphasized vigorous activity (and thus the possibility that 25OHD is, to some extent, a general health status indicator). 20. Typo in Scargg (should be Scragg). 21. PLease clarify which weights were used in the analyses.

REVIEWER	Alison Mondul University of Michigan, Ann Arbor, USA
REVIEW RETURNED	08-Feb-2017

GENERAL COMMENTS	This study used data from the National Health and Nutrition Examination Survey (NHANES) to investigate the association between leisure-time physical activity (LTPA) and circulating 25-hydroxyvitamin D (OHD) levels in cancer survivors. The authors found that overall physical activity is associated with higher levels of 25-OHD. In addition, after stratification by indoor/outdoor activity the association between LTPA and 25-OHD remained statistically significant for outdoor LTPA. According to the authors, this is the first study to look at the association between vitamin D and physical activity among cancer survivors. Therefore, the research question posed by the authors is novel with an important public health significance for this population. Comments: 1) I'm very concerned by the lack of thorough consideration of BMI. Body fatness is strongly associated with vitamin D level such that obese people are more likely to have vitamin D deficiency. There is evidence that this is because the large adipose mass in obese individuals is a storage depot for vitamin D and the large amount of vitamin D necessary to saturate the adipose tissue leads to lower circulating levels. The authors need to read and cite this literature. In addition, clearly physical activity and BMI will be associated. Thus, the use of broad categories of BMI, which isn't even a specific measure of body fat, is likely insufficient to control for this important source of confounding. Further, there may be a different relationship between vitamin D and physical activity in obese and non-obese individuals. Body fat percentage by bioimpedance is available from multiple survey cycles, and DEXA scans are available from multiple cycles, as well. Use of these measures will provide much more specific estimates of body fatness, which is the important measure for this analysis. More careful consideration of BMI as a continuous variable or using finer cutpoints should also be undertaken. This analysis needs to conduct several additional sensitivity analyses adjusting much more carefully for BMI and body fatness. It also needs to stratify by body fat and/or BMI as this may be an important effect modifier. 2) Please also indicate why you use two different cutoffs in your analysis for overall LTPA and indoor/outdoor LTPA. Also, please include the cutoff value you used for sufficiently active in both indoor/outdoor activity. 3) It seems that high levels of 25-OHD (>=50 nmol/L) is prevalent in this population (67% for 2001-2006, and 72% for 2007-2010, contrary to what you expected. Is there any reason for these results? This could be because vitamin D supplement use is common and may be recommended for cancer survivors by their physicians. NHANES has detailed information on supplement use and dietary intake of vitamin D. Adjustment for or stratification by use of

	individual vitamin D supplements or total vitamin D intake might be useful. Physically active individuals are likely to engage in other healthy behaviors so this would help address your hypothesis that physical activity is related to higher vitamin D through sun exposure. 4) In addition, if your outcome is common in the population. A log-binomial regression analysis should have been used to obtain prevalence ratios instead of the logistic regression. 5) Configuration of table 1 and 2, the columns should add to 100%. It would be much more informative to know the proportion of those with high vitamin D who were obese, for example. I also recommend showing broader categories of vitamin D in this table. Perhaps quintiles.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Anna-Maija Tolppanen
Institution and Country: University of Eastern Finland, Finland
Please state any competing interests: None declared

Please leave your comments for the authors below

The manuscript describes the association between leisure-time physical activity and 25OHD-levelsamong the cancer survivors of NHANES participants from 2001-2010. In general, the manuscript is easy to follow and the statistical methods are mainly appropriate but the following points should be taken into account:

Response: We thank the reviewer for their comments. Below are our point by point reponse to each comment.

1. Abstract& methods: please change multiple linear regression to multivariable

Response: We have changed “multiple linear regression” to “multivariable linear regression”. (Multiple locations where appropriate)

2. Abstract: please indicate clearly what the numbers of the result section indicate (differences between LTPA categories).
Response: We have indicated the comparison group. (Pages 2-3)

3. Is it necessary to use the term statistically significant? The results demonstrate when the 95% Cis indicate a difference, and if the authors want to bring in the significance, it would be much more helpful to indicate clinical significance.

Response: We have reworded the conlusion. Please see response to comment 4.

4. To continue with my previous comment, the authors could add some discussion about the clinical significance of results. Is approximately 10nmol/L difference clinically meaningful? Any studies showing that this kind of increase would manifest as health benefits? How much of it is perhaps due to confounding (ability to conduct more intensive LTPA could be a marker for overall better health status/socioeconomic position or something else)? It would be good to discuss these alternative explanations.

Response: We have discussed the clinical significance of the findings. (Pages 16-17) We have also added discussions on the alternative explanation on 25-OHD as an indicator for overall better health status. (Page 19)

5. The background section: the authors relyr heavily to their own publication on 25OHD levels and cancer prognosis as evidence on 25OHD levels being associated with better prognosis and survival, and 25OHD levels being the best indicator of vitamin status. Although this is a systematic review, it would be good to include more recent individual studies (this also may not be the best possible reference for vitD status biomarkers).

Response: We have added more recent individual studies 25-OHD levels and cancer survival. We also added reference for vitamin D status biomarkers. (Page 4)

6.It is unclear how “understanding the association between physical activity and vitamin D could inform cancer survivorship care strategies ”. This isbetter suited to discussion than the aims section where it is currently located. In addition, the authors may want to clarify why this is the case.

Response: We have edited this statement to provide clarity, moved it from the aims section to the discussion. (Pages 20-21)

7. Exclusion criteria (p.6) “excluded those who were never diagnosed withcancer, and were pregnant.” should “or” be used instead of “and”.

Response: We have changed “and” to “or’. (Page 6)

8. The authors state that participants who received chemotherapy within last month were excluded from blood collection. The implications of this should be discussed, as well as the time since cancer diagnosis and the resulting consequences to the generalizability and implications of the results. Is it so that the 25OHD levels promote better survival, or is it so that the results are more indicative of better physical condition and health status (ability to perform more intensive LTPA), which is proxied by higher 25OHD levels.

Response: We have provided additional references and discussion on the implication of excluding participants who receive chemotherapy within last month from blood collection. We acknowledged that this is a procedure carried out by the NHANES study. (Pages 7, 19, 20)
With regards to the time since cancer diagnosis, we have included in the study limitation that “we were not able to conduct analyses stratified by cancer type or time since diagnosis”. (Page 19) We further expanded the discussion on 25-OHD as an indicator of better physical condition and health status, and the limitation of a cross-sectional study to determine the direction of this association. (Page 19)

9. To continue, it would be good to draw a flow chart how the study population weas arrived at.

Response: We have included flow charts as suggested. As we described in the “Statistical Analysis session”, “Information on socio-demographic characteristics, weight, height, season of blood draw, and self-reported LTPA was complete among cancer survivors who had available data on circulating 25-OHD levels.” As a result, we have included all cancer survivors who had available data on circulating 25-OHD levels. (Figure 1, 2)

Figure 1. Participants flow chart - cancer survivors aged 20 years or older from the National Health and Nutrition Examination Survey (2001 - 2006)

Figure 2. Participants flow chart – cancer survivros aged 20 years or older from the National Health and Nutrition Examination Survey (2007-2010).

10. The explanation of continuous& categorical analyses should be moved from p.7 to Statistical analysis section.

Response: We have moved this part to the statistical analysis section. (Page 11)

11. on page 9, the authors mention that they classified activities that could be indoor/outdoor to indoor to ensure more conservative estimation. It would be interesting to see how robust the results were, i.e. how different the results would be if they were classified as outdoor, or 50/50.

Response: Among 48 activities, we considered basketball, bicycling, skating, stretching, swimming, martial arts, rope jumping and “others” could be indoor or outdoor activities and classified them to indoor to ensure more conservative estimation. As suggested, we alternatively classified them as outdoor and 50/50, respectively, for sensitivity analyses and observed similar results in the multivariable adjusted models.

When we classified these activities as outdoor activities, higher circulating 25-OHD level was seen among cancer survivors who were insufficiently active outdoor (5.19 nmol/L, 95% CI 0.52-9.85), and sufficiently active outdoor (6.39 nmol/L, 95% CI 2.85-9.94) (p-trend <.001). When we classified these activities as 50/50 (calculated based on MET-min/week) to indoor and outdoor activities, higher circulating 25-OHD level was seen among cancer survivors who were insufficiently active outdoor (5.65 nmol/L, 95% CI 0.6.55-10.65), and sufficiently active outdoor (7.26 nmol/L, 95% CI 2.88-11.64) (p-trend =0.002).

We have added the description of these analyses in the methods (page 12) and results sections. (Page 14)

12. Please add the numbers to “sufficiently active (≥MET-min/week”

Response: We have added the number, “sufficiently active (450 ≥MET-min/week)” (Page 10)

13. It is not evident why the authors doubled the time spent in vigorous activities. If the intensity impacts 25OHD levels, I would prefer to see an explanation for this in the introduction. Is it more likely that the sunlight exposure, defined more by time spent outdoors than intensity matters? If they wanted to assess the impact of LTPA, why not conduct the analyses separately for vigorous and moderate instead of doubling the time? If they prefer the approach they chose, it should be justified.

Response: We set out to examine the associations of the amount of LTPA with circulating 25-OHD levels. The data handling strategy of 2007-2010 LTPA was chosen because of two reasons, 1) the change of LTPA measure between 2001-2006 data and 2007-2010 data in the NHANES study, 2) prior published study. First, MET values were used to indicate the energy expenditure associated with the amount of physical activity. In the 2001-2006 data, MET values of LTPA were assessed by summarizing MET-minutes per week of each reported specific activity from a list of 48 activities, where the specific MET score for each activity was provided in the NHANES study. Therefore the overall METs have accounted for the frequency, intensity and duration of the LTPA. Whilst in the 2007-2010 data, information on specific activities was not available, thus the specific MET scores were not available. However, given the defined MET score cut-off for moderate (3.0-6.0 METs) and vigorous (>6.0 METs) physical activity, we doubled the amount of vigorous activity so that it is comparable to the amount of moderate activity. Please see reference number 24 for prior published study where the same assessment was used.

14. I did not quite understand the sentence “In the multiple linear regression models, we simultaneously adjusted for both indoor and outdoor activities, provided they were significantly different (P value<0.001).” Does it mean that they were added to the same model and if P<0.001, then they were both kept in? I would expect these variables to be highly correlated, and including both of them in the same model would likely lead to a situation when one of them no longer is related to outcome (and this also has implications on interpreting the individual coefficients –they are no longer interpretable). Please clarify what was actually done and how collinearity was accounted for. This collinearity may also explain why the association of indoor 25OHD levels disappeared in the mutually adjusted model.

Response: The correlation between indoor and outdoor activities was not significant using chi-square test (P value<0.001). We further tested the Pearson’s correlation of indoor and outdoor activity in continuous forms, and got a P-value=0.17; therefore they were simultaneously adjusted for. We have rephrased this sentence for clarification. (Page 11)

15. Why LTPA was analysed as categorized variable, instead of continuous one?

Response: Questionnaire assessment is known to over estimate the amount and the energy expenditure associated with physical acitivty. To provide public health implications to cancer survivors, using categorized variables could serve the purpose of this analysis without providing biased information on the association between specific amount of LTPA and circulating 25-OHD levels.

16. The authors have made a good start in discussing the possible limitations. However, they should discuss the implications of these to results (especially unmeasured confounding, and time since diagnosis).

Response: We have provided further details on the implications of these limitations. We have also included dietary vitamin D supplement in our analyses and have expanded our discussion as needed. (Pages 8-9, 19)

17. The authors say that bias arising from the lack of time since diagnosis is likely nondifferential. Wouldn’t this be actually likely differential between LTPA categories? Wouldn’t it be more likely that in general, those with longer time since dg are able to participate in more demanding LTPA

Response: Thank you for this comment. We have removed the statement in the text to reflect the correction.

18. “Our findings suggest that 25-OHD might be a surrogate marker of physical activity that accounts for the direct and indirect effects of LTPA, particularly outdoor.” please add references to previous papers that have reported the same.

Response: We have added references to previous papers. (Page 20)

19. While I wholeheartedly agree that physical activity among cancer survivors should be promoted, I don’t interpret the findings of this study as evidence on prioritizing outdoor activities with potential benefits of sunlight exposure. The authors may want to reconsider this, as their findings did not actually account for “sun time”, but emphasized vigorous activity (and thus the possibility that 25OHD is, to some extent, a general health status indicator).

Response: We have tempered our conclusions and also indicated that the higher 25-OHD levels

among those who were physically active could be a reflection of better overall health status. (Pages 19, 21)

20. Typo in Scargg (should be Scragg).

Response: We have changed “Scargg” to “Scragg”. (Page 16)

21. PLease clarify which weights were used in the analyses.

Response: We have clarified the weights used in the analyses, which was Medical Examination Center (MEC) exam weight. (Pages 8, 11)

Reviewer: 2
Reviewer Name: Alison Mondul
Institution and Country: University of Michigan, Ann Arbor, USA
Please state any competing interests: None Declared

Please leave your comments for the authors below

This study used data from the National Health and Nutrition Examination Survey (NHANES) to investigate the association between leisure-time physical activity (LTPA) and circulating 25-hydroxyvitamin D (OHD) levels in cancer survivors. The authors found that overall physical activity is associated with higher levels of 25-OHD. In addition, after stratification by indoor/outdoor activity the association between LTPA and 25-OHD remained statistically significant for outdoor LTPA.

According to the authors, this is the first study to look at the association between vitamin D and physical activity among cancer survivors. Therefore, the research question posed by the authors is novel with an important public health significance for this population.

Response: We thank the reviewer for their comments. Below are our point by point reponse to each comment.

Comments:
1) I’m very concerned by the lack of thorough consideration of BMI. Body fatness is strongly associated with vitamin D level such that obese people are more likely to have vitamin D deficiency. There is evidence that this is because the large adipose mass in obese individuals is a storage depot for vitamin D and the large amount of vitamin D necessary to saturate the adipose tissue leads to lower circulating levels. The authors need to read and cite this literature. In addition, clearly physical activity and BMI will be associated. Thus, the use of broad categories of BMI, which isn’t even a specific measure of body fat, is likely insufficient to control for this important source of confounding. Further, there may be a different relationship between vitamin D and physical activity in obese and non-obese individuals. Body fat percentage by bioimpedance is available from multiple survey cycles, and DEXA scans are available from multiple cycles, as well. Use of these measures will provide much more specific estimates of body fatness, which is the important measure for this analysis. More careful consideration of BMI as a continuous variable or using finer cutpoints should also be undertaken.

This analysis needs to conduct several additional sensitivity analyses adjusting much more carefully for BMI and body fatness. It also needs to stratify by body fat and/or BMI as this may be an important effect modifier.

Response: We have added the literatures on body fatness and vitamin D level. (Page 18) We agree that body fat percentage by bioimpedance and DEXA would add more information on the measure of body fatness. Unfortunately, the NHANES study has limited data on bioelectrical impedance analysis (2001-2002, 2003-2004) and DEXA measured body composition (2001-2002, 2003-2004 and 2005-2006). Furthermore, the age range of the sample was restricted to 8-49 years for bioelectrical impedance analysis data in all waves, and 6-69 years for DEXA scan in 2005-2006. Hence, we were not able to achieve a sizeable sample of cancer survivors with such measures that would have provided robust analyses.

There are ongoing debates on the association between physical activity and BMI. Physical activity is beneficial for initial weight loss at a small magnitude, and habitual physical activity is required for long-term weight maintenance (to avoid weight again). Please see Luke & Cooper, Int J Epidemiol (2013). Also, there is less data on cancer survivors to examine this association in longitudinal or experimental design, although the mechanism might differ between cancer and cancer-free people given the altered biological and physiological response in cancer survivors due to cancer and its treatment.

We have carried out additional sensitivity analyses by – 1) treating BMI as a continuous variable; 2) stratified analyses by BMI categories. Further, we adjusted for vitamin D supplement as suggested in comment 3.

We obtained similar results when we adjusted for BMI as a continuous variable - higher 25-OHD among cancer survivors who were active outdoor (5.82 nmol/L, 95% CI 1.69-9.95) in 2001-2006 data, non-significant findings in 2007-2010 data.

In analyses stratified by BMI, obese cancer survivors who were sufficiently active (7.10 nmol/L, 95% CI: 2.51-11.70), insufficiently active (6.47 nmol/L, 95% CI: 0.65-12.30) had higher circulating 25-OHD levels; p-trend=0.002 in the 2001-2006 data. Findings were similar in the 2007-2010 data. There were, however, no clear patterns observed in the normal weight and overweight groups. We have updated our manuscript to reflect these findings.

We have included the description and interpretation of the sensitivity analyses in the method (Page 12), results(Page 14), and discussion session (Page 18).

2) Please also indicate why you use two different cutoffs in your analysis for overall LTPA and indoor/outdoor LTPA. Also, please include the cutoff value you used for sufficiently active in both indoor/outdoor activity.

Response: The cutoff used for LTPA corresponds to the standard definition of physical activity guideline, 750 MET-min/week. A lower cutoff was used for indoor and outdoor activities, given they are sub-sets of overall LTPA. Therefore, we used 450 MET-min/week as the cut-off given is the minimal goal of weekly LTPA.

3) It seems that high levels of 25-OHD (>=50 nmol/L) is prevalent in this population (67% for 2001-2006, and 72% for 2007-2010, contrary to what you expected. Is there any reason for these results? This could be because vitamin D supplement use is common and may be recommended for cancer survivors by their physicians. NHANES has detailed information on supplement use and dietary intake of vitamin D. Adjustment for or stratification by use of individual vitamin D supplements or total vitamin D intake might be useful. Physically active individuals are likely to engage in other healthy behaviors so this would help address your hypothesis that physical activity is related to higher vitamin D through sun exposure.

Response: As requested, we have updated our manuscript to include data and analyses on

supplement use. In the 2007-2010 data, total dietary supplement information for each participant was provided by NHANES. In 2001-2006 data, we obtained information on individual supplement product in the "30-days dietary supplement dataset" for each participant, and linked to the "Ingredient Information" from the "Dietary Supplement Database" (Updated December, 2016 by NHANES). We then aggregated all the supplement use data for each participant if the ingredient contains Vitamin D. (Page 8, 9)

We created a binary variable for dietary vitamin D supplement use (yes/no), reported the mean circulating 25-OHD levels by dietary vitamin D supplement use (Page 13). We included this variable in tables 1 and 2 for descriptive statistics, and in the multivariable adjusted models (Tables 3 and 4). We observed that circulating 25-OHD levels were signifincalty higher among those who reported dietary vitamin D supplement use than those who did not in both 2001-2006 (68.82 vs 56.74 nmol/L, p<.001) and 2007-2010 data (83.73 vs 60.88 nmol/L, p<.001) data.

The association between LTPA and circulating 25-OHD level was retained in the 2001-2006 data but was attenuated in the 2007-2010 data. (Tables 3 and 4).

We observed no associations between Vitamin D supplement use and LTPA (p=0.19), indoor LTPA (p=0.32), or outdoor LTPA (0.83) in 2001-2006 data. Yet this association was significant in 2007-2010 LTPA data (p=0.03). Relevant discussion was added. (Page 18).

4) In addition, if your outcome is common in the population. A log-binomial regression analysis should have been used to obtain prevalence ratios instead of the logistic regression.

Response: We have corrected the odd ratios to prevalence ratios using method proposed by Zhang & Yu (JAMA, 1998) and updated our tables. We have described the procedure in the methods section and added the reference where appropriate. (Page 12 and table 4)

5) Configuration of table 1 and 2, the columns should add to 100%. It would be much more informative to know the proportion of those with high vitamin D who were obese, for example. I also recommend showing broader categories of vitamin D in this table. Perhaps quintiles.

Response: We have reformatted tables 1 and 2, with columns adding up to 100%, showing quintiles of vitamin D as suggested. (Tables 1 and 2) We also revised the statistical analyses (Page 11) and result (Page 13) section for table 1.

VERSION 2 – REVIEW

REVIEWER	Alison Mondul University of Michigan, USA
REVIEW RETURNED	12-Apr-2017

GENERAL COMMENTS	Comments to authors: The authors have been very responsive to previous comments and the manuscript is much improved. This is a very interesting study. Comments: - The BMI stratified analysis results are quite interesting. It appears that the results were different by BMI and this is highlighted in the Discussion section. Thus, it would be useful to have a BMI stratified table with a p for interaction.

	- Thank you for the additional analyses of supplemental vitamin D. In the Methods you mention that “total” supplement information was available for 2007-2010, suggesting that the information for the earlier period was less comprehensive. This wasn’t clear to me, though. Then in the Discussion you state: The association between LTPA and dietary vitamin D supplement use appeared to differ between 2001-2006 data (p=0.19) and 2007-2010 (p=0.03) data, although the prevalence of dietary vitamin D supplement use were similar in two study phases (51.4% vs. 51.5%). In the 2007-2010 data, active cancer survivors are more likely to report dietary vitamin D supplement use compared to inactive ones. Thus, the non-significant findings of LTPA and circulating 25-OHD levels could arise from the change in self-reported LTPA measures from 2001-2006 to 2007-2010 data. This was unclear to me. Did you mean that the non-significant findings in 2007-2010 could be due to differences in the collection of self-reported supplemental vitamin D measures in 2007-2010? If so, and the vitamin D measures are more comprehensive in 2007-2010, perhaps there is residual confounding in the earlier period, and the results from the more recent period are more correct? Please clarify this point. - If space allows, in the abstract, the authors should mention briefly about the findings from the 2007-2010 data for the associations between LTPA and 25-OHD. - In the methods, racial group should be “racial/ethnic” groups (page 7). - Thank you for your clarification about why there are two different cut offs from the overall and indoor/outdoor physical activity.Including this explanation briefly in the Methods would be helpful. - There is a typo in the first paragraph (5 line) of the results, it should say “significantly”

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2
Reviewer Name: Alison Mondul
Institution and Country: University of Michigan, USA Please state any competing interests: None declared

Please leave your comments for the authors below

Comments to authors:

The authors have been very responsive to previous comments and the manuscript is much improved. This is a very interesting study.

Response: We thank the reviewer for their comments. Below are our point by point reponse to each comment.

Comments:
- The BMI stratified analysis results are quite interesting. It appears that the results were different by BMI and this is highlighted in the Discussion section. Thus, it would be useful to have a BMI stratified table with a p for interaction.

Response: We agree that the BMI stratified analyses are interesting and offer additional information, given the association of adiposity with circulating vitamin D levels. We, however, suggest that stratified analyses should be interpreted cautiously because the relatively small number of participants in the different strata may not allow for very robust effect estimates. We have included the analyses as supplemental tables.

- Thank you for the additional analyses of supplemental vitamin D. In the Methods you mention that "total" supplement information was available for 2007-2010, suggesting that the information for the earlier period was less comprehensive. This wasn't clear to me, though. Then in the Discussion you state:
The association between LTPA and dietary vitamin D supplement use appeared to differ between 2001-2006 data (p=0.19) and 2007-2010 (p=0.03) data, although the prevalence of dietary vitamin D supplement use were similar in two study phases (51.4% vs. 51.5%). In the 2007-2010 data, active cancer survivors are more likely to report dietary vitamin D supplement use compared to inactive ones. Thus, the non-significant findings of LTPA and circulating 25-OHD levels could arise from the change in self-reported LTPA measures from 2001-2006 to 2007-2010 data.

This was unclear to me. Did you mean that the non-significant findings in 2007-2010 could be due to differences in the collection of self-reported supplemental vitamin D measures in 2007-2010? If so, and the vitamin D measures are more comprehensive in 2007-2010, perhaps there is residual confounding in the earlier period, and the results from the more recent period are more correct? Please clarify this point.
Response: Available information on dietary vitamin D supplement differ between 2001-2006 and 2007-2010 data. In the 2001-2006 data, only indivudal product level information were available, which was then linked to the Dietary Supplements Ingredient Database. We identified and aggregated ingredients that contained Vitamin D from all indivudal products for each cancer survivor. In the 2007-2010 data, aggregated data was available, which has taken into account all intakes from each individual product that icontains vitamin D. We have rephrased the text to clarify. (Page 9)

The vitamin D meaures are not neccesarily more comprehensive in the 2007-2010 data, than the 2001-2006 data. Rather, the physical acitivty measures differ between the 2001-2006 and 2007-2010 data. As described in the method session, this was why we analyzed data for 2001-2006 and 2007-2010 phases separately.

- If space allows, in the abstract, the authors should mention briefly about the findings from the 2007-2010 data for the associations between LTPA and 25-OHD.
Response: We have added a sentence in the abstract and edited accordingly to the word limit. (Page 3)

- In the methods, racial group should be "racial/ethnic" groups (page 7).
Response: We have corrected it to "racial/ethnic groups". (Page 7, 22)

- Thank you for your clarification about why there are two different cut offs from the overall and indoor/outdoor physical activity.Including this explanation briefly in the Methods would be helpful.
Response: We have included this explanation in the method. (Page 10)

- There is a typo in the first paragraph (5 line) of the results, it should say "significantly"
Response: We have corrected the typo. (Page 14)