Title: Association between serum 25-hydroxyvitamin D and prostate cancer in middle-aged and elderly Americans: A national population-based analysis of NHANES 2001–2018

Version: 1 **Date:** 09 Apr 2025

Reviewer's report:

Thank you for the opportunity to review this manuscript which explores the relationship between vitamin D and prostate cancer using data from NHANES. The results of this analysis suggested a positive association between increasing serum 25(OH)D levels and risk of pancreatic cancer among American adults over 40. The study used robust, nationally-representative survey data from NHANES and addresses a research question where reported associations are currently inconsistent. Suggested edits and comments are broken down by section below.

Article Information & Abstract
Abstract
- Line 17: The word controversial suggests that you are going to present both sides to this argument. Suggestion that the authors should rephrase the first sentence of the background
- Line 29: Can the n for number of prostate cancers identified in the cohort be provided? If not, 3.3% of the population suffices.
- Line 30 – 32: Only present result from the upper group with the p trend reported. No need to report the lower vitamin D group odds ratios.
- Line 35: Please specify what sensitivity analyses were completed. As written “several sensitivity analyses”, is vague.
- Line 39: Suggestion to authors to rephrase this sentence to include study minimum age – i.e., Among Americans over 40.

Keywords
- Please add the following: biomarkers, NHANES

Introduction
- Line 60: Include a comment on the recency of these referenced studies to help contextualize the research landscape.
- Line 63: As written, “black” is not clear what population group the referent study explored. Please ensure to indicate the country.
- Line 70: The authors should revise the study objective to clearly state the exposure of interest (vitamin D) and outcome (prostate cancer), even if it is repetitive.

Methods
Study Population
- Line 73: Methods should be capitalized
- Line 79: Instead of “which falls under”, suggest that the authors rephrase to “within”
- Line 80: Specify what was sanctioned by the RERC – is it NHANES itself or the research study. Suggestion to reorder the second and last sentence of this paragraph to improve flow.
- Line 82-87: Thank you to the authors for including a detailed overview of the sample size following each exclusion criteria. This is very helpful. Suggestion to revise the last sentence to remove the phrase “was included” and incorporate the language, the analytic dataset for this …. Study was 17,989.

Figure 1 – Caption should be revised to state “study inclusion/exclusion criteria”

Assessment of serum 25(OH)D levels and PCa
- Line 91: Missing citation from the sentence. Indicate that serum 25(OH)D is the gold standard biomarker for quantifying vitamin D status.
- As written, the authors do not comment on if NHANES blood collection was done in a fasted or non-fasted setting. Suggestion to revise this paragraph to include a short sentence on the biospecimen collection process for the survey.
- Missing comment on the missingness of the 25(OH)D biomarker in the analytic cohort and how samples with levels under the assay’s limit of detection (LOD) were handled. The LOD should also be reported, if it differs across cycles, it would be helpful to have this information in a supplement table.
Covariates
- Please include a citation or reference for the clinical definitions of hypertension, diabetes and CVD that are used in this analysis.
- Line 127: Please indicate the recency of the alcohol variable used to derive the drinkers and non drinkers classification (i.e., lifetime history of alcohol use or weekly).
Statistical Analysis
- Line 130: Move this statement to the end of the section to improve flow.
- Line 132: Include a comment about why the weights were used – i.e., to generate nationally representative estimates.
- Line 135: What test was used to compare the baseline characteristics across the 25-(OH)D status groups?
- Line 151: Suggestion to authors to remove word “initially” and proceed with the rest of the sentence. While the words initial, secondly and thirdly tell that three sensitivity analyses were conducted, it is not helpful for flow.
- Missing a description is what was done to create table 2

Results
- While the subheadings in this section are helpful to guide the reader through the results, it is suggested that the authors remove this as it is not needed since the analysis plan builds on each step and tells a cohesive story as is.
Characteristics of Participants
- Line 158: Rephrase this sentence and remove “ This research”. Suggested that the authors use the following language… “this study analyzed…”
- Line 159- 162: Please include the mean and % for each covariate referenced. Additionally, this would be strengthened by reporting the relevant p-value.
- Consider adding the range of the 25(OH)D biomarker in the cohort.
Table 2
- Correct table spelling. Prostate cancer is spelled incorrectly as “prastate”
- In the corresponding text in the manuscript, this is a univariate analysis. As written, the table caption says association between vitamin D status and PCa. I believe this is an error and should be re-titled for accuracy of this analysis.
- Please add the word “Abbreviation” before BMI. Missing abbreviation definition for WC, BMI.
Association between serum 25(OH)D and PCa
- Line 166: The statistical approach used in Table 2 is unclear. Please see comment above related to table caption for Figure 2. Please clarify and indicate effect estimates following each of the variables listed in the text.

- Line 178: Please provide the p-trend estimates to support this sentence.

Table 3

- Correct spelling in table caption (currently –"prastate")
- Add abbreviations to the footnote

Figure 2

- Update figure caption to specify that this is the results of restricted cubic spline regression
- Please specify why the curve does not intersect with the y axis. Is this due to the LOD of the assay? If so, please indicate the value and specify this as a footnote.

Figure 3

- High school should be capitalized in all categories
- Please use a combination of upper and lowercase levels for NO and YES
- Please reiterate how each health outcome is defined in a footnote below the figure and only report 2 decimal places for P-interaction.

Sensitivity Analysis

- Line 227: Rephrase sentence as a subgroup instead of re-obtaining data.
- Line 232: How was "extreme" serum 25(OH)D defined?
- Line 238: This sentence is unclear as written. If the authors are trying to suggest that the OR are similar in this sensitivity analysis and the full cohort analysis, this should be more explicitly stated.

Table 4

- Please include the n for each sensitivity analysis within the table
- Please only report 2 decimal places for the p values unless < 0.001
- Please add abbreviations to bottom of the table as a footnote

Discussion

- Line 240: Suggested that the authors revise the first sentence of the discussion to highlight the main findings of the study (with an effect estimate interpretation) and the sample size.
- Line 246: Please indicate the research type that is being referred to in this sentence to improve clarity.
- Line 248: Suggestion to specify the types of extraneous factors the authors are referring to in this sentence to improve clarity.
- Line 249: Please include an effect estimate from the original paper.
- Line 255: Remove extra "the"
- Line 257: While this is an interesting point the authors are making, suggestion to remove it as it is not well referenced and does not improve the validity of the associations observed in the study at present. These comments can be incorporated into the previous sentence.
- Line 260: Please clarify what was found in this meta-analysis. If the results were null, then that should be stated.
- Line 265: Suggestion to remove "consistent with former findings", as it is unclear which former findings the authors bare referring to given the inconclusive relationship that is highlighted in the paragraph above.
- Line 288: Suggestion to rephrase from "notable" to plausible
- Line 294: Suggestion to not start sentence with because
- This reviewer is appreciative of the thoughtful discussion of mechanistic plausibility of this pathway. The strengths and limitations sections of this manuscript are also an honest evaluation of the work. Addition considerations to further strengthen this paragraph include:

o Discussion of seasonality of 25(OH)D biomarker and potential for bias introduced by this
o Bias from imputation and handling of missing information

o Please revise the type of bias suggested through self-report of PCa. At present, memory bias is cited; however, this reviewer does not find this compelling as cancer status is not often misreported. If so, please provide an appropriate reference.

Conclusion
- Line 310: revise sentence to say “serum 25(OH)D” for clarity