1.a. Full Title: Influence of prevalent disease on associations between 25-hydroxyvitamin D and CVD, cancer and mortality: The Atherosclerosis Risk in Communities Study (ARIC)

b. Abbreviated Title (Length 26 characters): Vitamin D & all-cause mortality

2. Writing Group: Christopher Sempos, Pamela L Lutsey, Jeffrey R Misialek, Ramon Durazo-Arvizu, Joyce Merkel, Erin D Michos, Aaron Folsom, and Paul Coates. Others welcome.

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. __X__

First author: Christopher T. Sempos  
Address: 520 Ferdinand Dr.  
Havre de Grace, MD 21078  
Phone: (716) 400-3925  
E-mail: semposch@nih.gov

ARIC author to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (this must be an ARIC investigator).

Name: Pamela L. Lutsey  
Address: 1300 South 2nd St, Suite 300  
Minneapolis, MN 55126  
Phone: (612) 624-5812  
Fax: (612) 624-0315  
E-mail: Lutsey@umn.edu

3. Timeline: Data analyses to begin immediately. Goal completion is 9/1/2016.

4. Rationale: Vitamin D status is traditionally assessed by measuring and then evaluating the serum concentration of total 25-hydroxyvitamin D [25(OH)D]^1. Serum total 25(OH)D is the sum of the concentrations of serum 25(OH)D$_2$ and 25(OH)D$_3$. Serum total 25(OH)D is also known as serum 25(OH)D and, often, just as 25(OH)D.

Currently, there is thought to be a J-shaped association between serum 25(OH)D concentration and risk of death from all-causes$^{2-4}$. In order to understand what causes of death are producing the J-shaped curve, researchers have evaluated the association between serum 25(OH)D level and risk of death from three broad causes of death: (1) Cardiovascular Diseases; (2) Cancer and (3) All Other Causes of Death. To date, an inverse association has been found for 25(OH)D and risk of CVD death while no
association has been found with cancer deaths. Finally, there appears to be a j-shaped association between 25(OH)D and risk of death from “All Other Causes”.

In analyses where the outcome is all-causes mortality, typically there are no exclusions for preexisting disease, e.g. CVD and/or Cancer. As a result, a larger and much more difficult problem is to try and understand the meaning of results based on the typical design of studies evaluating associations with all-cause mortality. Thus, it is hard to know if the associations between 25(OH)D and risk of death from CVD and Cancer are the same as when participants with preexisting CVD and Cancer are excluded from the analyses. Moreover, it is generally unclear if the associations between 25(OH)D and death from CVD and/or Cancer – after excluding cases of baseline cases of CVD and Cancer – will resemble the associations between serum 25(OH)D and the incidence of CVD and Cancer. One tantalizing result from Framingham was a j-shaped association between serum 25(OH)D and incident CVD.$^5$

To get a better understanding of the meaning of the association between serum 25(OH)D and death from all-causes, we propose three sets of analyses in the ARIC Cohort to assess the association between serum 25(OH)D and:

1. All causes mortality, and death from CVD, Cancer and Other Causes \textit{without} excluding cases of baseline preexisting cases of CVD or Cancer;
2. All causes mortality, and death from CVD, Cancer and Other Causes \textit{after} excluding cases of baseline preexisting cases of CVD or Cancer;
3. Incidence of CVD and Cancer.

In other words, the two key questions to be answered by this study after excluding baseline (prevalent) cases of CVD and cancer are: (1) Is the pattern of the association of serum total 25(OH)D with all-causes mortality similar to that where prevalent cases are not excluded? and (2) Does the association between serum total 25(OH)D and CVD and cancer death reflect the association of serum 25(OH)D with incident CVD and cancer.

The analyses will be structured so as to be similar to prior analyses on the association between serum 25(OH)D and All-Causes mortality in the NHANES III (1988-1994) cohort with follow-up through 2006.$^2$

5. \textbf{Main Hypothesis:}

The pattern for the association between serum total 25(OH)D and the risk of death from CVD and Cancer will be similar to that with serum total 25(OH)D and the incidence of CVD and Cancer, i.e. a hypothesis of no difference. If there is no reason to reject this hypothesis then it can be stated that the association of vitamin D with CVD and Cancer mortality mirrors that for incidence.
6. Design and analysis (study design, inclusion/exclusion, outcome and other variables of interest with specific reference to the time of their collection, summary of data analysis, and any anticipated methodologic limitations or challenges if present).

Study Design

Prospective cohort from visit 2 through December 31, 2013

Inclusion/Exclusion

Analysis 1: All participants missing information of serum 25(OH)D, age, sex, race, month of exam, Education, Sports Index, Cigarette Smoking, Diabetes status at baseline, eGFR, SBP, High BP meds, Lipid Lowering Meds, HDL-C, LDL-C and BMI will be excluded.

Analyses 2 and 3: Participants with prevalent CVD and Cancer at visit 2 will be excluded. In addition, all participants missing information on variables for Analysis 1 will be excluded.

Variables

Exposures:

Primary: Serum 25(OH)D (measured in visit 2 serum; sum of 25(OH)D$_2$ + 25(OH)D$_3$). Serum 25(OH)D, initially, will be divided up into the following 9 groups in units of nmol/L: < 20, 20-30, 30-40, 40-50, 50-60, 60-75, 75-100 (Reference), 100-120 and ≥ 120. If the distribution of deaths is not adequate to support analyses stratified by the 9 pre-specified categories, categories will be collapsed.

Covariates: Age (years), Sex (1=male, 0=female), Race (B=Black, W=White)*, and Season** (Winter = 1, Summer = 0), Education (< High =1, Others = 0; High = 1, Others = 0); Cigarette Smoking (Current = 1, Others = 0, Former = 1, Others = 0); Sports Index; eGFR, SBP, HBP meds (1=yes, 0=no), Lipid Lowering Meds (1=yes, 0=no) HDL-c, LDL-c, BMI, and diabetes status (1=yes, 0=no).

*As usual, we will exclude those who are neither black nor white, as well as blacks from the MD and MN centers.

**Season will be defined as: (a) Winter when the month of exam was November – April; and (b) Summer when the month of exam was May-October.

Outcome: They are: (a) Death from All Causes, CVD, Cancer and Other Causes†; and (b) incident CVD and Cancer‡.
Table. ICD 9 and ICD 10 Classification Codes*

<table>
<thead>
<tr>
<th>Cause List</th>
<th>ICD Code Range</th>
<th>Comparability**</th>
</tr>
</thead>
<tbody>
<tr>
<td>113 Cause List #</td>
<td>ICD 9</td>
<td>ICD 10</td>
</tr>
<tr>
<td>Malignant Neoplasms*</td>
<td>19</td>
<td>140-208</td>
</tr>
<tr>
<td>Major CVD *</td>
<td>53</td>
<td>390-436-436-448</td>
</tr>
<tr>
<td>Deaths from all other causes Residuum</td>
<td>Residuum</td>
<td>Residuum</td>
</tr>
</tbody>
</table>

* Cancer and CVD, respectively.
**Comparability Ratio: Coding of a 80% of deaths from 1996 by ICD 10 and ICD 9 Rules
\[
\text{C}_{\text{Ratio}} = \frac{\text{ICD10 Deaths}}{\text{ICD 9 Deaths}}
\]

**Incident CVD will be defined as CHD, CHD mortality, stroke, heart failure, atrial fibrillation and sudden cardiac death.

*Time to Event:* Person-years of follow-up to event or to censoring.

*Potential effect modifiers:* None.

**Data analysis**
Visit 2 will serve as baseline for the current analysis. Visit 2 participant characteristics will be described using means and proportions stratified by: Vital Status at the end of follow-up and levels of vitamin D – all as separate tables.

A. **Analyses including** baseline cases of CVD and Cancer:

**Events:** (a.) All-Cause Mortality; (b.) CVD, Cancer, Other Deaths; and (c.) CVD and Cancer Incidence.
1. Determine the total sample size and the total numbers of events overall and within each event in each of the nine (9) 25(OH)D groups with exclusions by Model.
2. Calculate means and SE for: Follow-up Time; Age, Sex (%), Black(%), White(%), Winter(%), Summer (%) for: Entire sample, Alive, Dead, CVD Death, Cancer Death and Other Death.
3. Cox proportional hazards and Poisson categorical regression models of the form:

Model 1: Event = Age + Sex + Race + Season + VD<20 + VD20-30 + VD30-40 + VD40-50 + VD50-60 + VD60-75 + VD100-120 + VD≥120.

Model 2: Variables in Model 1 + <High School + High School + Current Smokers + Former Smokers + Sports Index
Model 3: Variables in Model 2 + eGFR + SBP + HBP Meds + Lipid Lowering Meds + HDL-c + LDL-c + BMI + Diabetes

4. Plot for each event: Relative Risk or Risk Ratio (RR)±SE by 25(OH)D category (Please keep scale of Y-axis constant and do not connect points with a line)

B. Analyses excluding baseline cases of CVD and Cancer:

Events: (a.) All-Cause Mortality; (b.) CVD, Cancer, Other Deaths; and (c.) CVD and Cancer Incidence.

1. Determine the total sample size and the numbers of events in each of the nine (9) 25(OH)D groups.
2. Calculate means and SE for: Follow-up Time; Age, Sex (%), Black(%), White(%), Winter(%), Summer (%) for: Entire sample, Alive, Dead, CVD Death, Cancer Death, Other Death, No CVD CVD Incidence, No Cancer Incidence.
3. Cox proportional hazards and Poisson categorical regression models of the form: Event = Age + Sex + Race + Season + VD<20 + VD20-30 + VD30-40 + VD40-50 + VD50-60 + VD60-75 + VD100-120 + VD≥120.
4. Plot for each event: RR±SE by 25(OH)D category (Please keep scale of Y-axis constant and do not connect points with a line).

C. Sensitivity Analyses: Use Analytical sample from Analyses B.

1. Run Models 1-3 before and after excluding cancer deaths which occurred in participants who – at the time of death are CVD free.
2. Run Models 1 -3 before and after excluding CVD deaths which occurred in participants who – at the time of death were cancer free.
3. Adjust 25(OH)D for Season separately in Blacks and Whites using the ARIC residual approach. Then please repeat analyses C1 and C2.

7.a. Will the data be used for non-CVD analysis in this manuscript? _X_ Yes ___ No

b. If Yes, is the author aware that the file ICTDER03 must be used to exclude persons with a value RES_OTH = “CVD Research” for non-DNA analysis, and for DNA analysis RES_DNA = “CVD Research” would be used? ____ Yes ____ No
(This file ICTDER has been distributed to ARIC PIs, and contains the responses to consent updates related to stored sample use for research.)

8.a. Will the DNA data be used in this manuscript? 
___ Yes _X_ No

9. The lead author of this manuscript proposal has reviewed the list of existing ARIC Study manuscript proposals and has found no overlap between this
proposal and previously approved manuscript proposals either published or still in active status. ARIC Investigators have access to the publications lists under the Study Members Area of the web site at: http://www.cscc.unc.edu/ARIC/search.php

___X___ Yes _______ No

10. What are the most related manuscript proposals in ARIC (authors are encouraged to contact lead authors of these proposals for comments on the new proposal or collaboration)?

No prior proposals have looked at 25(OH)D and mortality. The most similar proposals are those that look at 25(OH)D concentrations and incident outcomes:

<table>
<thead>
<tr>
<th>MS #</th>
<th>Outcome</th>
<th>First author (and Senior, if student first author)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019</td>
<td>Stroke</td>
<td>Schneider (Michos)</td>
</tr>
<tr>
<td>2377</td>
<td>CHD</td>
<td>Michos</td>
</tr>
<tr>
<td>2224</td>
<td>HF</td>
<td>Lutsey</td>
</tr>
<tr>
<td>2425</td>
<td>AF</td>
<td>Alonso</td>
</tr>
<tr>
<td>2479</td>
<td>PAD</td>
<td>Rapson (Lutsey)</td>
</tr>
</tbody>
</table>

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? ___X__ Yes _____ No

11.b. If yes, is the proposal

___X__ A. primarily the result of an ancillary study 2009.17 (Lutsey PI)
- “Serum vitamin D and cardiovascular disease risk in the biethnic ARIC cohort”

B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* __________ __________ __________)

*ancillary studies are listed by number at http://www.cscc.unc.edu/aric/forms/

12a. Manuscript preparation is expected to be completed in one to three years. If a manuscript is not submitted for ARIC review at the end of the 3-years from the date of the approval, the manuscript proposal will expire.

12b. The NIH instituted a Public Access Policy in April, 2008 which ensures that the public has access to the published results of NIH funded research. It is your responsibility to upload manuscripts to PubMed Central whenever the journal does not and be in compliance with this policy. Four files about the public access policy from http://publicaccess.nih.gov/ are posted in http://www.cscc.unc.edu/aric/index.php, under Publications, Policies & Forms. http://publicaccess.nih.gov/submit_process_journals.htm shows you which journals automatically upload articles to PubMed central.
References


