ARIC Manuscript Proposal #3482

PC Reviewed: 10/8/19  Status: ____  Priority: 2
SC Reviewed: _________  Status: _____  Priority: ____

1.a. Full Title: Plasma Proteins and All-Cause Mortality in Cancer Survivors in ARIC

b. Abbreviated Title (Length 26 characters): Proteins in cancer survivors

2. Writing Group: Platz, Barber, Couper, Meeker, Coresh, Ugoji, Prizment, Joshu, Tin and other interested ARIC investigators

We have invited Cristian Tomasetti from the Department of Biostatistics to be a co-author. He is a collaborator of Bert Vogelstein’s. Their group is interested in proteomics for the early detection of cancer.

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. EAP (ARIC author) [please confirm with your initials electronically or in writing]

First author: Elizabeth A. Platz, ScD, MPH
Address: Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health
615 N. Wolfe St., Room E6132
Baltimore, MD 21205
Phone: 410-614-9674
Fax: 410-614-2632
E-mail: eplatz1@jhu.edu

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: Adrienne Tin
Address: Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health
615 N. Wolfe St., Room W6017
Baltimore, MD 21205
Phone: 443-287-4740
E-mail: atin1@jhu.edu

3. Timeline: Expect to complete this work by December 2020

4. Rationale: Cancer survivors are purported to experience accelerated aging due the biological influences of shared cancer and aging risk factors (e.g., cigarette smoking), the cancer
itself and endogenous response to the cancer, and cancer therapies (e.g., chemotherapy, radiation therapy, hormonal therapy, immunotherapy). Accelerated aging phenotypes in cancer survivors have been described to include increased fatigue/fatigability, frailty, poor physical function, cognitive decline, osteoporosis, susceptibility to second primary cancers, cardiovascular events, and premature mortality, among other aging-related states (1). For example, in ARIC, compared with those who do not have a cancer history, cancer survivors have a higher risk of CVD (HR=1.43, 95% CI 1.34-1.58), which does not appear to be explained simply by shared risk factors for cancer and CVD (Florido et al. in process). However, in that analysis, CVD risk differed by cancer site, and was strong for lung, intermediate for breast, colorectal, and bladder, and not present for prostate cancer survivors.

A candidate biological mechanism underlying accelerated aging phenotypes in cancer survivors relative to similarly aged persons without a history of cancer is cellular senescence and associated (or consequent) “inflammaging”. Research on these candidate pathways (e.g., inflammaging) in cancer survivors in ARIC (Ugoji et al. ancillary study 2019.06 and associated manuscript proposals to be submitted on cognitive function and premature mortality, and on comparing inflammatory proteins between cancer survivors and persons without cancer) and in ARIC irrespective of cancer status (Walker et al. manuscript proposal #3327 on cognitive function) is already underway.

As a step toward discovering other mechanisms of accelerated aging in cancer survivors relative to persons who never had cancer, we propose to investigate whether cancer survivors – shorter (5-10 years) and longer term (>10 years), and specifically common survivable cancers – breast, prostate, colorectal, manifest different plasma protein levels than otherwise similar individuals without a cancer history. Then, we will determine whether these proteins are associated with mortality from all causes other than their cancer diagnosis as the ultimate indicator of accelerated aging. While these proteins may differ in concentration between cancer survivors and participants without a cancer history, the proteins may be uniquely related to premature mortality in cancer survivors (e.g., in those without a cancer history protein level is uniformly low or is not associated without premature mortality irrespective of level) or alternatively, may be similarly associated with premature mortality in both cancer survivors and those without a cancer history (e.g., protein level is higher in cancer survivors on average, but the RR is the same for the same protein level in the two groups). Thus, we will compare these findings to those conducted by other ARIC investigators in participants irrespective of cancer status to determine whether the same proteins emerge and for the same proteins, whether the associations with premature mortality in cancer survivors are unique to cancer survivors or a common experience of aging.

Any proteins that emerge and that are unique to cancer survivors later could be investigated in detail in relation to accelerated aging phenotypes in cancer survivors in ARIC, given that the experience of accelerated aging phenotypes is variable among cancer survivors. Ultimately, if associations between these proteins and accelerated aging phenotypes, including premature mortality from causes other than their cancers are identified and confirmed, they could be used to monitor cancer patients and survivors and to implement preventive (e.g., intervene on continued exposure to shared risk factors for their cancer and other cancers/chronic diseases) and supportive care to optimize their well-being.

Key points:
In the cross-sectional analysis, our primary goal is to identify proteins common across cancer sites in cancer survivors that differ in plasma level from participants without a cancer history. The sample size may not be sufficient to discover proteins unique to specific cancer sites, but we will assess whether the unique proteins identified for cancer survivors overall also differ among cancer sites.

In the prospective analysis, our primary goal is to determine whether the unique proteins common in cancer survivors, if variable in plasma level among them, inform their cancer survivors’ increased risk of accelerated aging measured as premature mortality. This project is focusing on discovering of etiology; we are not aiming to create a parsimonious proteomic risk score for use clinically in cancer survivors, which would require a different conceptual approach (e.g., scan through 5,000 in cancer survivors for those that are variable and then relate to outcomes using standard stepwise regression of LASSO [least absolute shrinkage and selection operator]).

We do not propose to investigate whether protein profiles can be used for cancer risk stratification, early detection of cancer, or cancer prognosis as these will be the subject of a separate ancillary study that this team will propose. Our focus here is on the etiology of premature mortality, that is, death from causes other than their cancer that occurs earlier than if the person had never had cancer, as the ultimate indicator of accelerated aging.

5. Main Hypothesis/Study Questions:

Among ARIC participants diagnosed with a first primary cancer after Visit 1 and who have survived at least 5 years after their most recent primary cancer by the time of the protein scan (Visit 5) and matched participants without a cancer history, evaluate whether:

Q1. Plasma protein levels differ between cancer survivors and participants without a cancer history who are matched by propensity score* on demographics (age, sex, race, field center) and socioeconomic factors (life course SES, access to and uptake of healthcare).
   a. Overall, and then evaluate whether differences in levels of these proteins are found by:
   b. Cancer site – especially common survival cancers – breast (post-menopausal female), prostate, and colorectum
   c. Survival time before Visit 5 – 5-10 years, >10 years
   d. Stage at cancer diagnosis – early, more advanced stage (e.g., T4 or N1 or M1)

Q2. Plasma proteins levels differ between cancer survivors and participants without a cancer history who are matched by propensity score* on demographics, socioeconomic factors, major cancer risk and protective factors that are shared with all-cause mortality (smoking, obesity, inactivity, alcohol drinking, diabetes, family history).
   a. Overall, and then evaluate whether differences in levels of these proteins are found by:
   b. Cancer site – especially common survival cancers – breast (post-menopausal female), prostate, and colorectum
   c. Survival time before Visit 5 – 5-10 years, >10 years
   d. Stage at cancer diagnosis – early, more advanced stage (e.g., T4 or N1 or M1)
Matching additionally on cancer risk/protective factors that are shared with all-cause mortality via a propensity score will allow us to home in on biological mechanisms related to the cancer itself, endogenous response to the cancer, and cancer therapies, rather than related to the biological influences of shared cancer and aging risk factors.

*If individual matching on propensity score proves inefficient, we will instead restrict the distribution of propensity score to the overlap between the cancer survivors and participants without a cancer history and weight by propensity score in the analysis.

In an exploratory aim, among ARIC participants diagnosed with a first primary cancer after Visit 1 and who have survived at least 5 years after their most recent primary cancer by the time of protein scan (Visit 5), evaluate for proteins identified in Q1 or Q2 whether:

Q3. Plasma protein levels are associated with risk of death from all causes (other than from their cancer(s) diagnosed before Visit 5) adjusting for demographics, socioeconomic factors, and major cancer risk and protective factors that are shared with all-cause mortality.
   a. Overall, and then in very exploratory analysis, evaluate whether differences in associations are found by:
      b. Cancer site – especially common survival cancers – breast (post-menopausal female), prostate, and colorectum
      c. Survival time before Visit 5 – 5-10 years, >10 years
      d. Stage at cancer diagnosis – early, more advanced stage (e.g., T4 or N1 or M1)

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:

   Q1 and Q2 – Propensity score matched cross-sectional study at Visit 5 (cancer survivors and participants without a cancer history)

   Q3 - Prospective cohort study of cancer survivors (follow from Visit 5 [baseline] to 2018)

   Analytic population:
   Q1 and Q2: Men and women who have Visit 5 protein scan data that passed quality control checks, who did not have prevalent cancer at Visit 1, who survived at least 5 years before Visit 5 after their latest cancer diagnosis or never had an incident cancer diagnosis between Visit 1 and Visit 5, and who consented to studies on chronic diseases including cancer are eligible. We will exclude participants who are not White or Black and participants who are Black from the Washington County and suburban Minneapolis (small numbers).
From these participants, we will select all eligible cancer survivors, and 1 to 5 propensity score matched participants without a cancer history for each survivor, using these criteria (Figure, Table):

- **Cancer survivors** - participants who had one or more first primary invasive cancers (other than non-melanoma skin) diagnosed after Visit 1 and by 5 years before Visit 5.
- **Participants without a cancer history** - participants who never had a cancer diagnosis by Visit 5. Participants with a diagnosis of non-melanoma skin cancer or in situ malignancies are eligible.

Some cancer survivors will not have a good match with a participant without a cancer history. One option is to exclude them from the analysis.

- **We will exclude from the analytic population participants** (Figure): 1) diagnosed with invasive cancer at least 5 years before Visit 5 and who have a new invasive cancer diagnosis between 5 years before Visit 5 and Visit 5, and 2) without an invasive cancer diagnosis at least 5 years before Visit 5 and who have a new invasive cancer diagnosis between 5 years before Visit 5 and Visit 5.

### Figure. Participants to be included and excluded from the analysis

<table>
<thead>
<tr>
<th>Analytic Cohort</th>
<th>5 years before Visit 5</th>
<th>Alive at Visit 5</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Row 1: Included</td>
<td>Cancer(s)</td>
<td>No new cancer</td>
<td>Cancer survivor</td>
</tr>
<tr>
<td>Row 2: Included</td>
<td>No cancer</td>
<td>No cancer</td>
<td>No cancer history</td>
</tr>
<tr>
<td>Row 3: Excluded</td>
<td>Cancer(s)</td>
<td>New cancer(s)</td>
<td>-</td>
</tr>
<tr>
<td>Row 4: Excluded</td>
<td>No cancer</td>
<td>Cancer(s)</td>
<td>-</td>
</tr>
</tbody>
</table>

### Table 1. Numbers of eligible cancer survivors and participants without a cancer history and deaths from all causes and causes other than their cancer

<table>
<thead>
<tr>
<th>Participants without a history of cancer (Row 2 in Figure)</th>
<th>Estimated through 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIC Cancer eligible, no cancer before Visit 1 and consented</td>
<td>5,158</td>
</tr>
<tr>
<td>Cancer Survivor Visit 1 to Visit 5</td>
<td>948  247   165</td>
</tr>
<tr>
<td>Cancer Survivor ≥ 5 yrs before Visit 5* (Row 1 in Figure)</td>
<td>598  148   114</td>
</tr>
</tbody>
</table>

*Row 1 in Figure represents cancer survivors who had one or more first primary invasive cancers (other than non-melanoma skin) diagnosed after Visit 1 and by 5 years before Visit 5.
Mean cancer survival 12 years (min 5, max 26 years)

Q3: Cancer survivors only from Q1 and Q2.
   o In a sensitivity analysis, we will include any cancer survivors who were excluded in Q1 and Q2 because of the inability to propensity score match to a participant without a cancer history.
   o In a sensitivity analysis, we will include any cancer survivor irrespective of with the diagnosis was ≥ 5 yrs before Visit 5 (N=948).

Measurements:

Protein scan – We will use protein levels, measured as relative fluorescence units [RFU] for a standard plasma volume per participant using aptamer-based profiling in blood collected at Visit 5. Some proteins were FLAGGED on one or more plates. Several options are available for handling such proteins: exclude protein (conservative), include protein (not conservative), exclude only if the number of plates with the flagged protein is larger than expected by chance alone (optimal, but need to define an arbitrary expected)

Cancer – Existing 2015 ARIC cancer case file, which captures cases diagnosed between Visit 1 and 5 years before Visit 5

Death – 2018 ARIC mortality file (Q3)

Additional variables (either covariates or for use in sensitivity analyses) – Visit 5 values, unless otherwise noted:

Demographics: age, sex (Visit 1), race (Visit 1), field center (Visit 1)

Socioeconomic factors: Lifecourse SES (Visit 4), neighborhood income, health insurance status, frequency of routine physical examination, having a dentist, frequency of routine dental visit, last time of dental visit with cancer survivor status (Visit 4)

Cancer risk/protective factors shared with all-cause mortality: Body mass index, waist circumference, height (Visit 1), cigarette smoking status and pack years (accumulated to Visit 5), alcohol consumption (Visit 5), meeting physical activity guidelines (Visits 1 and 3), diabetes status (Visit 5), aspirin use (Visit 5), statin drug use (Visit 5), hormone replacement therapy use (women, Visit 5)

Factors that influence plasma protein levels or are indicative of acutely distorted protein levels (for sensitivity analyses): In a subanalysis, we will exclude participants using a diuretic. Key proteins that mark kidney and liver function and acute phase inflammation were measured currently in the protein scan. In the sensitivity analysis, we will exclude individuals at the extremes of the distributions of these: cystatin-C; aspartate aminotransferase [AST] and alanine aminotransferase [AAT]; and C-reactive protein, respectively. As an alternative, we could use indicators previously measured, albeit, not concurrently with the protein scan, that have
clinically relevant cutpoints: eGFRcr-cys (15 mL/min/1.73m² [stage 5]), hsC-reactive protein concentration (acute inflammation, 10 mg/L), liver function (>3 times the reference ranges: ALT 7-56 IU/L, AST 0-35 IU/L).

**Propensity score**
For Q1, we will model the association of age (continuous), sex, race, field center, lifecourse SES, neighborhood income, health insurance status, frequency of routine physical examination, having a dentist, frequency of routine dental visit, last time of dental visit with cancer survivor status using logistic regression to predict the propensity score for each participant (2). We will confirm the positivity assumption of the use of the propensity score (i.e., no/negligible number of participants have a probability of 0 or 1 of being a cancer survivor) and determine whether the overlap in scores between the cancer survivors and those without a cancer history is sufficient. Then, for each cancer survivor, we will select up to 5 participants without a cancer history with the closest propensity scores (nearest neighbor; we may impose a maximum difference after reviewing the distribution) (3). For Q2, we will repeat these steps with the expanded set of cancer risk/protective factors that are shared with all-cause mortality (smoking, obesity, inactivity, alcohol drinking, diabetes, family history).

**Statistical analysis**
We will generally follow statistical analysis approaches described previously for proteomics data in epidemiologic studies (4, 5). At this time, we are not seeking to develop an optimized classifier based on these proteins, and thus have not proposed to use such methods (e.g., (6)). We will transform protein RFUs to achieve normality based on the participants without a cancer history. We expect overlap in the proteins identified in Q1 and Q2 with those that will be investigated as candidates for inflammation pathways in Ancillary Study 2019.06. We will not investigate these in Q3.

Q1a and Q2a. We will use linear regression to determine whether level of each protein in the scan differs between the cancer survivors and matched participants without a cancer history adjusting for the factors in the propensity score. We will use the Holm-Bonferroni method to deal with multiple testing, which is less conservative than the Bonferroni method (https://en.wikipedia.org/wiki/Holm%E2%80%93Bonferroni_method). In Q1b-1d and Q2b-2d, we will investigate whether levels of the proteins that emerge from Q1a and Q1b differ between cancer survivors and their matched participants within subgroups defined by cancer site (e.g., breast, prostate, colorectum), survival time before Visit 5 (5-10 years, >10 years), and stage at cancer diagnosis (early, more advanced stage). We will test for interaction by including in the model main effects terms and a cross-product term(s), the coefficient for which will be evaluated by the Wald test.

We will determine the correlations among the proteins that emerge separately among the cancer survivors and participants without cancer. Post-hoc, we will determine whether the proteins that differ in level cluster in particular pathways (e.g., using propriety software such as Ingenuity https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/, or use publicly available databases such as STRING https://string-db.org/cgi/input.pl?sessionId=MsFSINgwAW97&input_page_show_search=on, which includes KEGG https://www.genome.jp/kegg/brite.html#gene and Reactome
Q3a. We will first determine whether levels of the proteins identified in Q1 and Q2 are variable among the cancer survivors, and determine whether transformation is needed to normalize the distribution. We will move all proteins from Q1 and Q2 that are variable among cancer survivors to the next step (i.e., will not use a top hit approach). Restricting to cancer survivors, we will use Cox proportional hazards regression to estimate the association of each of these proteins with all-cause mortality (other than from their cancer(s) diagnosed before Visit 5), expressed as hazard ratios (HRs) and 95% confidence intervals (CIs). Start of follow will be date of blood draw at Visit 5, end is date of death from any cause except their cancer(s), date of death from their cancer (censor time, non-event), or end of 2018. We will adjust for demographics, socioeconomic factors, and major cancer risk and protective factors that are shared with all-cause mortality. We will separately enter each protein into model as a series indicator variables for quantiles of level based on the analytic cohort of cancer survivors (e.g., quartiles, tertiles, binary at median). We will test for trend by entering protein level as a continuous variable, with appropriate transformations, and testing its coefficient using the Wald test. In very exploratory analysis, we will repeat these steps separately by cancer site (especially common survival cancers – breast, prostate, and colorectum with adjustment for cancer-site specific covariates), survival time before Visit 5 (5-10 years, >10 years), and stage at cancer diagnosis (early, more advanced).

**Minimum detectable association:**
Q3. Based on conventional p-values uncorrected for multiple testing and using currently available sample size estimates: for total cancer, with the sample size in the Table, for a 2-sided test with alpha=0.05 and a power of 80%, comparing extreme categories, we can detect as statistically significant HRs of: quartiles 2.41, tertiles 2.22, and binary at median 1.92 or higher. The number of outcomes will be larger after the death file is updated. The number of cancer survivors (analytic cohort) would be larger if we incorporated data for 1) cases with diagnoses post Visit 1 who at Visit 3 had survived at least 5 years and had proteomics data at Visit 3, and 2) cases not eligible at Visit 5 because they had not yet survived 5 years and who have proteomic profiling at Visit 6 and have survived 5 years.

Points of clarification and challenges:
- We propose to use Visit 5 protein scan data cross-sectionally to compare cancer survivors and those without a cancer history, and with follow-up of cancer survivors, to compare levels of the unique proteins in relation to mortality through 2018. If Visit 6 data become available, we will model protein levels as time-varying in the mortality analysis.
- Persons with and without a cancer history may profoundly differ on demographics (due to differences in cancer risk profiles and cancer survivorship profiles) and on social factors (e.g., socioeconomic status, access to and uptake of care) that are also related to risk of premature mortality, thus we will propensity score match cancer survivors and those without a history of cancer as Visit 5.
- Given the large number of proteins in the scan, we will use a straightforward, efficient method of assessing differences between cancer survivors and matched participants simply by comparing adjusted geometric means and standard deviations and will determine statistical
significance taking into account multiple testing. Future refinements may include assessment of differences in protein detectability (e.g., present versus absent after defining detectability [versus noise such as non-specific aptamer binding by use of data from the negative controls]), differences in the spread or shape of the distribution including thresholds, and or patterns of proteins (e.g., network analysis).

- Several approaches may be used to deal with multiple testing and false positives. Possible approaches are p-value based while others incorporate information about the magnitude of the differences between groups. Subsets of the 5,000 proteins in the scan are correlated (the tests we will perform will not be independent), thus corrections for multiple testing may be too conservative and result in false negatives for individual proteins.

- We will restrict to participants with cancer who survived at least 5 years to reduce the likelihood that they are still in treatment (minimize effects of current treatment on the proteins) and because these are the individuals who have the opportunity to experience accelerated aging (they did not die of their cancer in the near term, and so are at risk for late effects). For Q3, in a sensitivity analysis, we will include these participants.

- Some participants have been diagnosed with multiple first primaries. Second and subsequent primaries can be the result of or otherwise related to the first primary (e.g., treatment, genetics, shared risk factors). Subsequent cancers that precede 5 years before Visit 5 and that later result in death will not be included as death from all causes, whereas subsequent cancers diagnosed after Visit 5 and that result in death will be included as a death from all causes.

- Past cancer treatment could influence current protein levels, e.g., some systemically delivered chemotherapies can cause acute damage to the kidneys. However, this damage tends to occur close in time to treatment and is more often seen in those with advanced stage cancer (given the nature of therapies used) and those with pre-existing chronic kidney disease (https://www.uptodate.com/contents/cisplatin-nephrotoxicity; (7)). From linkages with cancer registries, we have first course of treatment, but not details on type, dose, frequency for chemotherapy, radiation therapy, immunotherapy or other biological treatments, or history of subsequent treatments for recurrence/progression. In this analysis, we will restrict to longer-term cancer survivors, who are less likely to be those with advanced stage disease, and in the near term, we will stratify by first course of treatment. We expect that surgery (e.g., for early prostate cancer) would not likely influence protein patterns or levels. While not in the scope of this manuscript proposal, ARIC Cancer investigators are currently determining how the already available CMS data can be used to supplement the treatment information from the cancer registries. If we had complete treatment data, we would exclude participant who received nephro- or hepatotoxic chemotherapies, for example, in a sensitivity analysis.

- We expect overlap in the proteins identified in Q1 and Q2 with those that will be investigated as candidates for inflammation pathways in Ancillary Study 2019.06. We will not investigate these in Q3. Again, the goal of this study is to discover new pathways, rather than to study candidate pathways.

- While we expect to identify proteins and pathways that systematic differ between cancer survivors and those without a cancer history we cannot rule out:
  - Abundant proteins that are perturbed in many acute and chronic disease states, such as albumin and C-reactive protein (would be reported by Ugoji et al.), may predominate the proteins that differ between cancer survivors and those without a cancer history, if the former participants are, on average, less healthy at Visit 5 than the latter participants (e.g.,
whether due residual differences after matching in a higher prevalence of past and current shared risk factors, or consequences of cancer or its treatment).

o Alternatively, given their longer-term survival, the cancer survivors (mean of 12 years, minimum of 5 years, maximum of 26 years) in this study may be more like those without a cancer history, and thus, no protein differences may be found overall (Q1 and Q2). However, this observation would not rule out effect modification by cancer survivorship status, and thus, we would proceed with Q3 (scan through all 5,000 proteins [minus those in Ugoji et al.] for their associations with premature mortality among cancer survivors only).

• We recognize that the sample size is small for ‘omics research and in a setting in which the disease – cancer (site, aggressiveness, treatment) – is heterogeneous. Again, in Q1 and Q2, our goal is to describe important (big) differences between those with and with a cancer history, given that cancer survivors, on average, tend to experience premature aging. And, given that not all cancer survivors experience premature aging, in Q3, our goal is to describe the variability of those proteins among the cancer survivors and relate them to the ultimate indicator of premature aging, death from causes other than their cancer. We intend this proposal to provide preliminary data for a grant application that would take advantage of the repeated protein measures in ARIC and complement other ARIC proposals using a candidate pathway approach to understand cancer survivorship in older adults.

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? ___X_ 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

8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the file ICTDER03 must be used to exclude those with value RES_DNA = “No use/storage DNA”? _____ Yes  _____ 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)?
MS # 3057: Repeatability and Longitudinal Variability of the Plasma Proteome (Tin, Coresh et al.)

MS # 3415: Characterizing prostate specific antigen (PSA) measured by SOMAscan and its correlates in men and women in ARIC (Platz, Tin, Coresh et al.)

Ancillary study: 2019.06 Systemic Inflammation, Aging phenotypes and Mortality in Cancer survivors and associated manuscript proposals to be submitted (Ugoji et al.)

MS # 3327 A proteomic analysis of incident dementia: The ARIC Study (Walker et al.)

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
   ___  A. primarily the result of an ancillary study (list number*)
   2017.27 Proteomic longitudinal ARIC study: SOMAscan of multiple visits
   2011.07 Enhancing ARIC Infrastructure to Yield a New Cancer Epidemiology Cohort
   1995.04 Cancer Study
   ___  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.cscu.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.cscu.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.

13. Per Data Use Agreement Addendum, approved manuscripts using CMS data shall be submitted by the Coordinating Center to CMS for informational purposes prior to publication. Approved manuscripts should be sent to Pingping Wu at CC, at pingping_wu@unc.edu. I will be using CMS data in my manuscript ____ Yes __X__ No.

References