1.a. Full Title: Preliminary Analysis of Obesity-related DNA Methylation Markers and Risk of Post-menopausal Breast and Endometrial Cancer

b. Abbreviated Title (Length 26 characters): DNA Methylation and Cancer

2. Writing Group:
   Writing group members: Ellen Demerath, Weihua Guan, Steven Nguyen, Eric Boerwinkle, Corinne Joshu, Elizabeth Platz, Heather Nelson, plus others

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

First author: Mark E. Sherman
Professor of Epidemiology and Pathology
Address: Mayo Clinic, 4500 San Pablo Road, Jacksonville, FL 32224
Phone: 904-953-0771 Fax:
E-mail: Sherman.Mark@Mayo.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: Ellen Demerath
Professor, Epidemiology and Community Health

Address: Department of Epidemiology and Community Health
University of Minnesota School of Public Health
1300 S. Second Street, Suite 300
Minneapolis, MN 55454
Phone: 612-624-8231 E-mail: ewd@umn.edu

Other co-author email address:
Corinne Joshu: Cjoshu1@jhu.edu
Elizabeth Platz: eplatz1@jhu.edu
3. **Timeline:** Complete analysis: June 1, 2018 - December 31, 2018

4. **Rationale:**

**Obesity and Cancer Risk**

The prevalence of obesity has increased dramatically in the U.S. since the 1970s and has emerged as a leading risk factor for diabetes, cardiovascular disease and cancer. A recent umbrella analysis identified obesity as a risk factor for at least 11 types of cancer (Kyrgiou *BMJ* 2017) and obesity is associated with a worse prognosis for many cancer types (apart from lung cancer) (Goodwin *Ann Rev Med* 2015). In particular, obesity is associated with increased risks of postmenopausal breast and endometrial cancer, and consequently obesity-related cancers disproportionately impact women, accounting for 20% of cancer deaths as compared with 14% among men (Calle *N Engl J Med* 2003). Further, obesity prevalence is higher in groups affected by health disparities, including African Americans and members of lower socioeconomic groups, which may exacerbate the impact of obesity-related diseases.

While obesity, defined as a body mass index (BMI) >30 Kg/m$^2$, is strongly linked to adverse health outcomes among groups of women, the pathophysiological consequences of excess body mass among individual women is highly variable. Data suggest that accumulation of visceral fat in particular is associated with adverse metabolic alterations and increased risk of cancer and cancer predisposing conditions, such as diabetes and metabolic syndrome. However, BMI and visceral adiposity are not perfectly correlated. Consequently, improved methods of stratifying individuals with respect to cancer risk are needed, such as alternative anthropometric measures that may better reflect visceral adiposity (e.g. waist-to-hip ratio) or circulating biomarkers, which reflect its pathophysiologic consequences. Further, adult weight gain, which may better reflect increasing adiposity than BMI, has been linked to increased risk, particularly for endometrial cancer and breast cancer (Giovannucci *CA Cancer J Clin* 2010).

Several proposed mechanisms may account for the relationships between obesity and cancer risk, including increased production of circulating estrogens and relative increases in its bioavailability (lower sex-hormone-binding globulin), chronic inflammation with increases in circulating inflammatory cytokines, increased leptin to adiponectin ratio, hyperinsulinemia and elevated insulin-like growth factors (IGFs), immune suppression, changes in adipose tissue derived exosome secretion and related miRNA cargo and DNA methylation (reviewed in Allott *Endocr Rel Cancer* 2015, Azvolinsky *J Natl Cancer Inst* 2016 and Khandekar *Nat Rev Cancer* 2001; Thomou *Nature* 2017). Most, if not all of these factors have been associated with increased risk of women's cancers and their causal role is supported by mechanistic evidence. Mendelian Randomization analyses support a cause link between obesity and several cancers, including endometrial cancer (Painter *Cancer Epidemiol Biomarkers Prev* 2016, Jarvis *Br J Cancer* 2016).

Thus, we hypothesize that the identification of biomarkers that predict cancer risk among obese women would have value for selecting candidates for cancer prevention interventions. Further, we suggest that DNA methylation markers are particularly promising candidates for risk stratification because epigenetic modifications are mechanistically linked to
carcinogenesis through transcription silencing of tumor suppressor genes and may represent modifiable intermediate endpoints that can be monitored to assess weight change in relation to cancer risk.

**Obesity and DNA methylation**

Epigenetic Wide Association Studies (EWAS) have demonstrated associations of BMI and related adiposity metrics (e.g. waist-to-hip ratio) with DNA methylation in circulating leukocytes (Wahl et al *Nature* 2017, Geurts *Int J Obesity* 2017), including prior research conducted in ARIC (Demerath *Hum Mol Genet* 2017). Limited within person analyses have also revealed concordant methylation marks in leukocytes and somatic tissues, including adipose depots, muscle, liver, spleen and pancreas. Despite potential challenges related to population stratification and shifts in blood leukocyte differential counts, many loci discovered in EWAS have been independently validated. Most sites identified by EWAS in relation to BMI reflect hypermethylation of CpG islands, many of which occur in tumor promoter regions of genes, and when methylated are associated with transcription silencing. Methylation silencing of tumor suppressor genes is an established carcinogenic mechanism. Data from the EWAS reported by Wahl et al support a mechanistic relationship between DNA methylation and reduced gene expression; increased levels of DNA methylation in CpG islands near promoter regions were linearly associated with reduced levels of corresponding gene transcripts. DNA methylation of tumor suppressor genes is a common mechanism in obesity-related cancers, including postmenopausal breast and endometrial cancer. Further, several obesity-related methylated markers found in EWAS are linked to purported carcinogenic mechanisms, including insulin resistance and diabetes, and some of these methylation markers have been identified in breast cancer tissues (W. Taylor, Mayo; personal communication).

**Prior EWAS in ARIC**

A prior EWAS among African American participants in ARIC from the Mississippi and North Carolina centers identified methylation markers associated with BMI (76 probes) and waist circumference (167 probes), which overlapped considerably; 8 of these probes were also related to BMI change (Demerath *Hum Mol Genet* 2017). Independent replication of these associations in blood was achieved for 37 probes related to BMI and one related to waist circumference; an additional 16 were replicated in adipose tissue. Given that obesity and diabetes are important established risk factors for postmenopausal breast and endometrial cancers, we propose to analyze the relationships between previously identified adiposity-associated DNA methylation markers and risk of postmenopausal breast and endometrial cancers among women with available EWAS data in ARIC, including published data for 1,673 post-menopausal African American women enrolled at the Mississippi and North Carolina centers, of whom 124 were diagnosed with incident breast cancer and 16 with incident endometrial cancer (Demerath *Hum Mol Genet* 2017). ARIC also has unpublished EWAS data for 577 White women of whom 32 developed incident breast and 11 incident endometrial cancers, which we plan to include in this analysis of cancer risk. Given uncertainties in the sequence and causal relationships of obesity, waist-to-hip circumference, diabetes and specific DNA methylation markers, we will develop preliminary statistical models to explore these relationships, albeit with modest power. Further, we recognize that some women who are not obese may have metabolic disturbances comparable
to heavier women, which could be linked to methylation and cancer risk. Results of this analysis may provide preliminary data to support the development of an expanded study in ARIC of obesity-related DNA methylation and cancer risk.

5. **Main Hypothesis/Study Questions:** We hypothesize that BMI-associated DNA methylation markers are associated with increased risk of postmenopausal breast and endometrial cancers. Thus, we propose the following specific study questions:

1) We will assess associations for methylation markers found in the EWAS of African American participants with BMI, waist circumference and BMI change among 577 White women with existing unpublished DNA methylation data.

2) We will assess whether adiposity-related DNA methylation markers discovered in a prior EWAS in ARIC (Demerath *Hum Mol Genet* 2017) are associated with increased risk of post-menopausal breast (n=156) and endometrial cancers (n=27) using Cox regression. We will analyze cancer risks associated with individual DNA methylation markers, assess correlations among the markers, and develop a multivariable model to assess the most predictive combination of markers that independently predict cancer risk. We will assess the attributable risk related to these markers and assess risk of postmenopausal breast cancer alone. We will assess multivariate models both including and excluding BMI, waist circumference, weight gain and diabetes. We will also perform analyses stratified by BMI and diabetes and metabolic syndrome status and assess interactions between DNA methylation and disease states to assess effect modification. Specifically, we predict that we will find an interaction between adiposity traits, diabetes and metabolic syndrome and DNA methylation markers with regard to cancer risk. We will attempt to model whether weight gain after visit 2, the time when blood was collected for DNA methylation analysis, modifies risks associated with markers related to both BMI and cancer risk. We will assess the gene ontology of methylated loci associated with cancer to provide mechanistic clues about etiology and pathogenesis.

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).**

We will use ARIC visits 1-5 (all)

**Inclusions/Exclusions:**
- Must have DNA methylation data for blood with 450K Illumina chip of adequate quality
- Must have data for adiposity measures at visit 2
- Must have covariate data that could produce substantial confounding, potentially including use of exogenous hormones, smoking, parity, diabetes;
  - Useful if data also available for other factors, such as age at menarche, age at menopause, infertility, Polycystic Ovary, menstrual abnormalities, family history of cancer, current medication use (e.g. insulin, metformin, other medications for diabetes, weight loss, diabetes, hypertension), physical activity, ethanol use.
- For EWAS replication among 577 White women, also: WBC counts/differential and assay factors (batch, plate, etc.)
- Exclude women with cancer diagnosed prior to visit 2 treated with systemic agents
- Exclude women who are status-post hysterectomy in endometrial cancer analyses.

**Sample Size Estimates:**
2,250 postmenopausal women with EWAS data, including 183 with incident endometrial or breast cancer diagnosed after visit 2.

**Demographics:**
Patient ID, Date of DNA collection, Center, Age, Education, Income

**Obesity Traits:**
Weight at age 25 years; at all study visits for which data are available: weight, height, BMI, waist circumference, hip circumference, waist-hip ratio; CT measures of visceral and subcutaneous adipose tissue, if available

**Person-time:**
Time from collection of blood for analysis of DNA methylation (study visit 2) to: 1) diagnosis of postmenopausal breast or endometrial cancer, 2) loss to follow-up, 3) death or 4) diagnosis of other cancers requiring systemic therapy. For endometrial cancer, censor at benign hysterectomy.

**Methylation Markers:**
We will analyze relationships of methylation markers to cancer identified in the ARIC EWAS and the subset of these markers, which were independently confirmed (see Demerath *Hum Mol Genet* 2017)
- Any of 76 probes related to BMI, 164 related to waist circumference or 8 related to BMI change in ARIC
- Any of 37 independently replicated probes in blood related to BMI or one related to waist circumference and 16 replicated in adipose tissue

**Statistical Analysis:**
We will confirm the previously identified associations of DNA methylation with BMI, waist circumference and BMI change among 577 White women for general consistency of direction of association and comparability of beta values, representing percentage CpG island methylation. We will exclude markers related to SNPs that vary between races and are associated with CpGs. Markers that do not show a confirmatory relationship with methylation among White women and were not independently confirmed will be dropped. We will evaluate the distribution of beta values for individual methylation markers to determine optimal variable formatting (continuous, categorical or binary). We will develop univariate regression models assessing levels of DNA methylation for markers, transformed as needed for normalization, versus cancer risk. We will assess correlations among obesity-related methylation markers and develop appropriate multivariate models including DNA methylation marker levels that are independently predictive of cancer risk, as specified above, adjusted for confounders. It is unclear which DNA methylation markers may be downstream of obesity, weight gain, waist circumference, diabetes, metabolic syndrome (on the pathway) versus a possible cause of these conditions. Further, DNA methylation could modify the associations between established risk associations and obesity-
related cancers. Wahl et al used Mendelian Randomization analyses to argue that BMI is the cause of DNA methylation marks, rather than the converse and data from the ARIC EWAS show that diabetes only weakly modifies associations of BMI with specific DNA methylation loci. Accordingly, we will develop models both including and excluding diabetes and metabolic syndrome and compare the associations. When appropriate, we will test for significant interactions between DNA methylation and disease states or conditions (e.g., diabetes, metabolic syndrome) with regard to cancer risk. We will assess the gene ontology of genes associated with DNA methylation that is related to cancer risk.

**Possible Scenarios to account for relationships of BMI, DNA methylation, chronic non-neoplastic diseases and cancer**

- **Scenario #1**: BMI influences DNA methylation; specific methylation marks are associated with diabetes and metabolic syndrome, whereas others are linked to cancer.
- **Scenario #2**: DNA methylation is the cause of obesity, diabetes and metabolic syndrome and these predisposing conditions drive associations with cancer.
- **Scenario #3**: Obesity causes diabetes and metabolic syndrome, which in turn causes cancer. BMI is linked to DNA methylation, but the molecular changes are not causally related to non-neoplastic diseases or cancer.

These analyses are considered preliminary and exploratory. If associations between DNA methylation and cancer risk are identified, we will propose to extend the analysis to a nested case-control analysis with ARIC by assessing DNA methylation in circulating leukocytes for most predict markers discovered through this work.

In a companion analysis, we will assess DNA methylation markers in EWAS conducted in other populations (Wahl et al *Nature* 2017, Geurts *Int J Obesity* 2017) in relation to cancer risk.

We will perform analyses using Cox regression with adjustment for confounders to assess hazard rates associated with the development of breast or endometrial cancer and breast cancer alone.
We will confirm proportional hazard assumptions. If ER status is available for breast cancer, we will perform a sensitivity analysis restricted to ER-positive cancers.

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? __X__ Yes  ____ 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”? __X__ 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)?

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

11.b. If yes, is the proposal

___ A. primarily the result of an ancillary study (list number* __________)
___ 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/](http://publicaccess.nih.gov/) are posted in [http://www.csc.unc.edu/aric/index.php](http://www.csc.unc.edu/aric/index.php), under Publications, Policies & Forms. [http://publicaccess.nih.gov/submit_process_journals.htm](http://publicaccess.nih.gov/submit_process_journals.htm) shows you which journals automatically upload articles to PubMed central.