1.a. Full Title: The Association of a Multi-Systemic Biological Risk Index with Risk for Colorectal Cancer in Black and White Adults.

b. Abbreviated Title (Length 26 characters): Biological Risk and Colorectal Cancer

2. Writing Group:
Writing group members:
Andrew Odegaard
Anna Prizment
Corinne Joshu
Other interested ARIC Cancer Working Group members

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

First author: Teofilia Acheampong, MPH
Address: 222 Irvine Hall, School of Medicine,
University of California- Irvine, Irvine California, 92617
Office Phone: (949) 824-7401
Cell Phone: (646)-643-0160
E-mail: Tacheamp@uci.edu

Co-author: Andrew Odegaard, MPH, PhD
Address: 222 Irvine Hall, School of Medicine,
University of California- Irvine, Irvine California, 92617
Office Phone: (949) 824-7401
E-mail: aodegaard@uci.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: Anna Prizment
Address: U of MN, Division of Epidemiology and Community Health, 1300 2nd St, Mpls, MN 55454
Phone: 612-62
E-mail: prizm001@umn.edu

Name: Corinne Joshu
Address: JHU, 615 N. Wolfe Street Room E6148 Baltimore, Maryland 21205
Phone: 443-287-3821
E-mail: cjoshu1@jhu.edu
3. **Timeline:** Once the data is received we plan to immediately begin the analysis and proceed to drafting a manuscript within 6-8 months.

4. **Rationale:**
   Allostatic load is a metric of health risk used to express shared physiologic variance in multiple biological systems.\(^1\)\(^2\) It is based on the hypothesis that recurrent exposure to external stressors leads to progressive dysregulation in multiple physiological systems.\(^3\) Conceptually, allostatic load aims to explain the underlying biological relationship between chronic stressors (environmental, psychosocial) and disease.\(^4\) The concept of allostatic load has been widely used in theoretical and empirical studies as a way to measure the stress response in humans.\(^4\)\(^-\)\(^6\)

   In theory, allostatic load as a metric captures the complex biological cascade that occurs in the response to stress by measuring biomarkers in different physiological domains that are the end result of the activation of different autonomic and neuroendocrine pathways triggered by stressors. Each domain (cardiovascular, metabolic, immune) has been strongly implicated as an target outcome regarding biological pathways in stress within both animal and human models\(^7\)-\(^10\) A working model from Duong et al\(^11\) is noted below in Figure 1.

   ![Diagram of allostatic load model](image)

   **Fig. 1** The path from psychosocial stress to disease. HPA is an abbreviation for the hypothalamic-pituitary axis. SAM is the abbreviation for the sympathetic-adrenal medullary system

   Indeed, there is no uniform, agreed upon operationalization standard for defining allostatic load.\(^4\)\(^,\)\(^11\) However, multiple studies have established validity of the construct by demonstrating common variance and statistical coherence between measures from psychometric and stress related instruments, prominent primary mediators of the stress response (e.g. stress related hormones), and secondary mediators reflecting the ensuing biological alterations that accumulate over time – both risk factors and causal measures in the etiology of disease.\(^12\)-\(^17\) Importantly, summary allostatic load indices have demonstrated stronger prediction, or magnitude of association with outcomes, than the individual components/domains that inform the calculation of the index.\(^13\),\(^18\) Furthermore previous research has demonstrated strong, positive associations between higher levels of allostatic load indices (representing greater multi-systemic biological risk) and cognition, physical function, cardiovascular disease, and mortality.\(^4\),\(^19\),\(^20\)

   The theory of allostatic load/multi-systemic biological risk may be of particular relevance in terms of advancing prevention and sub-clinical use of some of the risk factors that have been shown to be important in the development of colorectal cancer (CRC). Thus far, studies have shown that sub-clinical, and clinical cut-points for biomarkers in the domains of metabolic,\(^21\)-\(^27\) immune,\(^28\)-\(^31\) and vascular\(^32\)-\(^35\) health are pertinent predisposing factors to colorectal cancer development. Therefore to further examine improved risk estimation for factors related to colorectal cancer risk, we propose utilizing an index of multi-systemic biological risk.

   In terms of metabolism there are many mechanism at play. It is suggested that insulin resistance affects the proliferation and neoplastic transformation of normal epithelial cells directly, or through its effect on IGF-1 and IGFBP and both normal colorectal epithelium and colon cancer cells have IR and IGF-1 receptors.\(^22\),\(^25\),\(^36\)-\(^38\) In ARIC, Ahmed et al. reported a multivariate-adjusted association of metabolic syndrome with colorectal cancer in men (RR, 1.78; 95%CI, 1.0-3.6), and a dose-response association between CRC incidence and number of metabolic syndrome components present at baseline (P for trend = \(.006\))\(^22\), indicating that an accumulation of components increased risk for CRC.

   Chronic inflammation has also been implicated in several CRC studies.\(^29\),\(^30\) Prizment et al. reported that after multivariate adjustment, for the highest versus lowest quartile, there was a positive
association of CRC risk with fibrinogen: (HR = 1.50 95% CI, 1.05–2.15), P = 0.03; inflammation z-score: (HR = 1.65 95% CI, 1.15–2.35), P = 0.01; and CRP (an acute phase reactant): (HR = 1.97 95% CI, 1.13–3.43), P = 0.02.29 It was also noted that there was no association detected for any of the other individual markers in the study and thus, a certain combination of inflammatory markers may be a better indicator of inflammatory processes associated with CRC risk than any of the markers individually.

Lastly, the relationship between the autonomic nervous system, and colon neoplasms has not been studied outside of one animal study.39 In principle, autonomic nerve impulses innervate intestinal tone and motility.40–42 External stressors, in both human and rodent studies, has been shown to impact the brain-gut axis.43–45 Some psychosocial stressors may trigger the onset of gastrointestinal disorders,43 including those that increase risk for colorectal malignancies.31,46–48 That being said, the autonomic nervous system regulates the cardiovascular system, and dysfunction has been strongly implicated in heart disease.49 Although studies have been mixed, a recent meta-analysis has shown that ischemic heart disease was associated with increased odds for colorectal neoplasms (OR 1.87, 95 % CI 1.38–2.54, p < 0.001).50 Reviews and a study done in ARIC, has shown that there are many obesity related shared risk factors for cancer and cardiovascular disease, which may indicate that this link may be bidirectional.36,51 Furthermore, stress has been implicated in cardiovascular disease10 as well as gastrointestinal disorders,43,52 which puts individuals at an increased risk for CRC.47

Nevertheless, we are not aware of any studies that have examined an index of allostatic load/multi-systemic biological risk, or long-term exposure to chronic stress with colorectal cancer incidence. Thus, a prospective analysis examining the association of an index of multi-systemic biologic risk with incident colorectal cancer would address a major gap in the literature, giving credence to in vivo and in vitro studies of stress and cancer outcomes. Furthermore, the results from this analysis may provide further insight into the etiologic heterogeneity of colorectal cancer, one of the most common cancers in the U.S. This could potentially inform public health approaches to targeted cancer prevention and disparities (black v. white), and it may also have clinical utility for cancer prediction since the index relies upon commonly measured biomarkers.

5. Main Hypothesis/Study Questions:
The main hypothesis of this research aims to test whether allostatic load is associated with colorectal cancer incidence. Since ARIC does not have direct measures of stress at baseline from validated psychometric or stress related assessment instruments, we are unable to truly validate the allostatic load concept in the cohort, thus it is more apt to refer to the hypothesis we are testing as “Multi-systemic Biological Risk” and risk for incidence colorectal cancer as done in the literature.

We have one main study question.
1) Is an index of Multi-systemic Biological Risk (proxy for allostatic load) associated with risk of incident colorectal cancer? We hypothesize that a higher score on the index is positively associated with colorectal cancer 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 methodological limitations or challenges if present).

Study Design: Prospective Analysis of ARIC data

Inclusion and exclusion: Participants with a history of cancer at baseline, participants missing data from any of the main components of the index will be excluded.

Outcome and other variables of interest:

Aim 1: Incident colorectal cancer as ascertained from 1987 to 2012 through linkage with state cancer registries in Minnesota, North Carolina, Maryland and Mississippi, and supplemented by active surveillance of the cohort, which includes recording of hospital discharge codes for all participants.
The central exposure variable of interest will be an index of multi-systemic biological risk comprised of a panel of biomarkers aligned with different biologic systems measured at the baseline (ARIC visit 1) visit. As noted above, there is no standardized definition or operationalization of allostatic load. Previous studies have combined biomarkers into an index that are considered primary mediators of this concept (i.e. indicators of the response to stress with measures of cortisol, catecholamines, DHEA as examples) and secondary mediators reflecting biological dysfunction over time from multiple domains. As the evidence base on the topic has grown, the three primary physiological domains that are consistently represented in validated research examining allostatic load and outcomes, are measures related to dysfunction in the cardiovascular, metabolic, and immune domains.

Thus, to provide the best proxy for allostatic load, and create an index of multi-systemic biologic risk, we will create an index using measures under these biologic domains informed by clinical cut points and treatment guidelines for risk, or based upon evidence in the literature base indicating a threshold of risk for disease. This approach will provide a strong basis for clinical applicability of the results. The scoring rubric is noted below:

- **High risk= 2**
- **Moderate risk=1**
- **Low risk=0**

- **Cardiovascular:**
  - Heart Rate Variability (SDNN)\(^{53-57}\): Greater than 100 ms= 0, Between 51-100 ms=1, Less than or equal to 50 ms= 2.
  - Blood pressure (SBP/DBP)\(^{58,59}\): Less than 120/80 mm/Hg=0, between 120-139/80-89=1, greater than or equal to 140/90 mm/Hg or treatment with hypertensive medications=2.

- **Metabolic:**
  - HOMA-IR ([glucose (nmol/L) x insulin (µU/mL)/22.5])\(^{60-69}\): Less than or equal to 2.6 =0, between 2.6 - 4.65 =1, greater than 4.65, or diabetes=2.
  - Triglycerides\(^{70,71}\): <150 mg/dL =0, 150 to 199 mg/dL =1, > 200 mg/dL = 2.
  - Waist Circumference: Less than 94 cm (M); 80 cm (W)= 0, between 94-102 cm (M); 80-88 cm (W)=1, greater than 102 cm (M); 88 cm (W)=2.

- **Immune:**
  - White blood cell count\(^{72,73}\): 4,500 to 11,000 cells/mcL=0, between 11,001-50,000 cell/mcL=1, greater than 50,001 cells/mcL=2.
  - Fibrinogen\(^{74-77}\): 200 - 400 mg/dL = 0, between 400-530 mg/dL =1, greater than 530 mg/dL=2.

Other Variables to be included in the analysis: age, sex, race, education, income, health insurance status, hormone therapy use (women), physical activity, smoking, diet quality alcohol, BMI, co-morbidities (CVD, Diabetes, Depression), other medication use.

**Summary of Data Analysis:**

- For each biomarker each person will either be assigned a 0 or 1 or 2, with the summary score ranging from (0-14), with a higher score indicating higher multi-systemic biological risk. We will examine the association of the resulting score with incident colorectal cancer. The nature of the distribution of the score will determine the statistical approach that is most apt to examine the association (e.g. whether the distribution allows for a full comparison across the distribution, or it is more appropriate to rank the distribution into quantiles, using the lowest as referent in both). This approach will allow us to test whether any association observed is monotonic in nature or more of a threshold nature.
- Cox proportional hazards models will be used to estimate hazard ratios (HR) and 95% confidence intervals (CI) for cancer incidence in aim 1. We will carry out tests of the proportional hazards
assumption, and if violated we will account for this as appropriate. If power allows, a sub-
analyses, will assess colon and rectal separately.
  o Model 1: Age, sex, race, center, education, income, health insurance status, hormone-
therapy use (women).
  o Model 2: Model 1 + Physical activity, smoking, alcohol, diet quality score (DASH score)
  o Model 3: Model 2 + BMI, co-morbidities.
  o We will test for effect modification by sex, race, smoking, BMI, alcohol use. Stratified
estimates will be reported if there is evidence that the results differ in a material manner.
We will test for a trend by including the index modeled as a continuous variable in the
Cox models.

• Sensitivity analyses:
  o (1) According to previous ARIC studies, approximately less than 25 %, 4%, 5% and 3%
of participants reported use of anti-hypertension, glucose lowering, anti-depressant and
cholesterol lowering medication respectively at baseline. Based on the literature, we
expect that many of these medications (especially beta-blockers) may attenuate the
association. Therefore, we will exclude participants with documented use of these
medications at baseline to see how this affects the association by comparing estimates
with these participants and without.
  o (2) To inform the interpretation of the results from the overall index, we will also
examine the association between each of the individual domains and CRC incidence.
    ▪ Aggregate cardiovascular domain: Heart rate variability, Blood pressure
    ▪ Aggregate Metabolic domain: Insulin Resistance, Triglycerides, Waist
      Circumference
    ▪ Aggregate Immune domain: Fibrinogen, White blood cell count
  o (3) Under the premise of latency period between exposure and its impact on disease risk,
we will exclude cases that occurred within 2 years of baseline to inform interpretation of
results.

Anticipated limitations and challenges:

Ideally, we would be able to carry out the analyses with a repeated measure of the index, but the
components are not universally measured at a follow up visit so we will rely on the single baseline
measure. Although, it is important to note that all of the measures track strongly on the population level
over time. Statistical inference into the association between the index and the outcomes will have less
power in the noted stratified analyses and will thus require more cautious interpretation. However, based
on estimated calculations (N~13,000 with expected exclusions, and ~. 0296 overall incidence utilizing
previous studies) we expect an effect size of about 1.38 (HR), suggesting that these analyses will have
ample power for the main outcomes. Lastly, the main lifestyle confounders (physical activity, diet,
alcohol intake, and to some extent smoking) are each measured with more error on average than the
clinical measures, and even with adjustment there is likely to be some residual confounding from these
measures.

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.csec.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? X Yes  ____ No

11.b. If yes, is the proposal
   
   _X__  A. primarily the result of an ancillary study (list number* 2011.07, 1995.04)
   
   ____  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.cscen.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.

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


