1.a. Full Title: Advanced glycation end product biomarkers in association with diabetes and diabetes-related traits

b. Abbreviated Title (Length 26 characters): AGEs and diabetes

2. Writing Group:

Writing group members: Stephanie Loomis, Yuan Chen, David Sacks, Mark Halushka, Eric S. Christenson, Robert H. Christenson, Elizabeth Selvin; others welcome

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

First author: Stephanie Loomis
Address: Johns Hopkins University
2024 E. Monument Street, Suite 2-600
Baltimore, MD 21287

Phone: 860 912-7169
Fax:
E-mail: sloomis5@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: Elizabeth Selvin
Address: Johns Hopkins University
2024 E. Monument Street, Suite 2-600
Baltimore, MD 21287

Phone: 410-955-0495
Fax: 410-955-0476
E-mail: eselvin@jhu.edu
3. **Timeline:**

Data are available. We anticipate a rapid timeline for this project and aim to have a first draft of the manuscript to co-authors in 6 months.

4. **Rationale:**

Advanced glycation end products (AGEs) are proteins, nucleic acids or lipids that are nonenzymatically glycated by aldose sugars such as glucose. They are synthesized endogenously or ingested through smoking or consumption of foods that have undergone the Maillard reaction, a browning that occurs during the cooking process.\textsuperscript{1,2,3} AGEs are thought to be central in diabetes, a disorder defined by chronically elevated glucose levels.\textsuperscript{4}

AGEs are members of the immunoglobulin super family that can induce a signal transduction cascade leading to an inflammatory response.\textsuperscript{5} AGEs bind to different receptors including the membrane-bound advanced glycation end product receptor (RAGE), which activates the proinflammatory NF-kB cascade. This interaction leads to increased expression of RAGE, which induces more NF-kB, causing chronic inflammation.\textsuperscript{6,5} RAGE is also present in a soluble form, sRAGE, which circulates in the blood and lacks the domain that activates NF-kB, blocking the proinflammatory pathway. Endogenous secreted RAGE (esRAGE) is an alternatively spliced isoform of RAGE that also lacks the NF-kB activating domain and does not induce inflammation.\textsuperscript{1,2,3} A prevailing hypothesis is that unbound AGE ligand levels are reduced if sRAGE and esRAGE levels are elevated enough to outcompete RAGE for AGE binding, leading to a decreased risk of inflammation and inflammatory disease.

The AGE-RAGE involvement in inflammation is of interest for its role as a potential biomarker for inflammatory disease, but the signal transduction pathway that includes AGE and RAGE involves other molecules which add complexity to the AGE-RAGE interaction. AGEs bind to many other receptors such as AGE receptor 1 (AGE-R1) and AGE receptor 2 (AGE-R2),\textsuperscript{7} and RAGE has alternative ligands such as S100/calgranulin and high-mobility group protein B1 (HMGB1).\textsuperscript{8} Each of these molecules may compete with each other for ligands and receptors, affecting the amount of proinflammatory stimuli present in the cell. In addition, AGEs are much more prevalent than sRAGE (approximately 1000-fold greater), limiting the ability of sRAGE to have a major impact in neutralizing the pro-inflammatory effect of AGEs.\textsuperscript{9} There are also multiple forms of AGEs, the most common type being N(6)-Carboxymethyllysine (CML-AGE), making understanding which AGEs are relevant in disease etiology complicated.\textsuperscript{7}

Despite the complexity of the pathway, population-based studies have implicated various AGE measures as potentially useful biomarkers of diabetes and independent predictors of diabetic complications,\textsuperscript{10,11,12,13,14,15,8,16} although the literature has been mixed for particular biomarkers, most notably sRAGE. Some studies have shown a positive association between sRAGE and risk of cardiovascular disease,\textsuperscript{17,13,14} mortality,\textsuperscript{17,18} or renal outcomes,\textsuperscript{18} among diabetic patients, while others have shown an inverse
The variable sRAGE results could also be related to the diversity of assays being employed in research studies. Gas or liquid chromatography mass spectrometry is sensitive, but not without technical limitations. While mass spectrometry has been used, it is not likely feasible either in epidemiologic studies or as a clinical biomarker. Most recent studies have utilized an enzyme-linked immunosorbent (ELISA) assay in serum or plasma, which uses antibodies to selectively bind to epitopes on molecules of interest. These epitopes may occur on multiple molecular structures extracted from the serum, making the test not highly specific to AGEs. In addition, ELISAs lack standardization, leading to variable results. AGEs can also be measured by fluorescence through the skin, as in the Diabetes Control and Complications Trial (DCCT). This method is noninvasive, but it does not detect all AGEs because not all of them fluoresce, it is not specific to individual AGEs, nor is it quantitative.

In addition to problems with measurement, variable sRAGE values may be due to genetics. Maruthur et al showed that a variant in the AGER gene explained up to 26\% of the variation in sRAGE levels in blacks and whites, but this variant was not associated with incident death, coronary heart disease, diabetes, heart failure, or chronic kidney disease in the ARIC study.

The overarching objective of this proposal is to examine the associations of three of the most common serum AGE biomarkers that have been examined in clinical and
epidemiologic studies: sRAGE, esRAGE and AGE-CML measured in the serum using commercially available ELISA assays, with diabetes and diabetes related traits in the community-based ARIC study.

5. Main Hypothesis/Study Questions:

In this study, we will characterize the associations of sRAGE, esRAGE, and AGE-CML with diabetes and diabetes-related traits in participants from the ARIC Carotid MRI (CARMRI) Study, a subset of the ARIC Study.32

**Hypothesis 1:** Diabetes and certain diabetes risk factors will be positively associated with higher levels of AGE-CML, and inversely associated with the two AGE receptors, esRAGE and sRAGE.

**Hypothesis 2:** The associations between AGE-CML, esRAGE and sRAGE with diabetes and diabetes related traits will differ by race and history of cardiovascular disease.

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 population:** Approximately 2,000 individuals in the CARMRI subset of the ARIC study with biomarker data available for serum AGE-CML, sRAGE, and esRAGE measured as part of a grant from the American Heart Association (PI: Selvin). Participants were aged 63-83 in 2004-2005, when they underwent carotid MRI imaging and provided blood and urine samples as well as anthropometric and lifestyle questionnaire data.

**Study design:** Cross sectional analysis using data and biospecimens collected in 2004-2005 in the CARMRI study

**Inclusion/exclusion:** We will limit our analysis to ARIC individuals who participated in the CARMRI study and have valid AGE biomarker data available, and exclude those with missing values for other variables of interest.

**Exposure variables:** Three biomarkers were measured in stored serum samples obtained between 2004-2005 as part of the CARMRI study: AGE-CML, sRAGE and esRAGE. Samples were analyzed using ELISA assays to determine biomarker levels (CML-AGE: AGE-CML ELISA, Microcoat, Penzberg, Germany; esRAGE: MyBioSource Inc., San Diego, CA; sRAGE: Quanatikine Human RAGE Immunoassay, R&D Systems Inc., Minneapolis, MN). Analysis was performed in a CLIA licensed laboratory accredited by the College of American Pathologists.

**Outcome:**
1) Diabetes: defined as
   - no diabetes: fasting glucose < 100 mg/dl
   - prediabetes: fasting glucose >=100 mg/dl, <126 mg/dl
   - diabetes: fasting glucose >=126 mg/dl, self report physician diagnosis, or on diabetes medication

2) the following key diabetes related traits:
   - Hypertension: defined based on systolic blood pressure and diastolic blood pressure
   - Hyperlipidemia: defined based on lipid-lowering medication use, HDL, LDL, cholesterol, triglycerides
   - Inflammation: defined by C-reactive protein (CRP)
   - Overweight or obese: defined by body mass index (BMI), waist-to-hip ratio
   - Medication: defined by any CVD related medication use
   - Family history: defined by family history of diabetes
   - Disease history: defined by history of cardiovascular disease

All variables were measured during one study visit between 2004 and 2005 as part of the CARMRI study.

Covariates: Age (years) and sex. We will test for interactions by race (only African Americans and Caucasians were included in the CARMRI study, thus race is a binary variable) and history of cardiovascular disease.

Potential effect modifiers: race, history of cardiovascular disease

Data analysis:

The CARMRI study contains an oversampling of individuals with thicker carotid arteries as measured from a previous ultrasound exam in ARIC visit 3 or 4 and thus is a nonrandom sample of ARIC participants. To achieve the sampling scheme, site specific cutpoints were created, and approximately 100% of the participants with arterial thickness above the cutpoints were included, while approximately 20% of the participants with arterial thickness below the cutpoints were included. We will adjust for this sampling scheme by including weights in all analyses, using the inverse of the sampling fractions of the eight sampling strata (four study sites, each with a percent included above and below the cutpoint). We will use the svyset command in STATA to incorporate the weights and improve validity and representativeness of the ARIC study for this analysis.

We will compare mean levels of each AGE biomarker stratified by diabetes status and by race using t tests. We will also compare age adjusted means using the same stratifications.

We will then perform univariate regression analyses between each trait and each biomarker, modeling outcomes as continuous for linear regression and dichotomized comparing the lowest quartile to the upper three quartiles for logistic regression. We will
examine the linearity of the biomarkers with diabetes and diabetes related traits to determine if splines are necessary and implement them if applicable. We will run several multiple regression models, controlling for different covariates:

Model 1: age, sex, race
Model 2: age, sex, race, smoking
Model 3: age, sex, race, CRP
Model 4: age, sex, race, BMI
Model 5: age, sex, race, BMI, CRP

outcome = diabetes (fasting glucose>126 mg/dl), diabetes related outcomes: systolic blood pressure, diastolic blood pressure, hypertension, HDL, LDL, cholesterol, triglycerides, C-reactive protein (CRP), body mass index (BMI), waist-to-hip ratio, any CVD related medication use, family history of diabetes, family history of cardiovascular disease

biomarker = CML-AGE, sRAGE, or esRAGE

We will then repeat these models, stratifying by race and history of cardiovascular disease.

We will also add interaction terms for exposure of interest x race to determine if there are statistical interactions between race and diabetes/diabetes traits.

Limitations:

As with any cross-sectional analysis, we will not be able to determine the temporality of any observed associations. Also, we will only be able to evaluate single measurements of AGE biomarkers measured in serum. The relatively small number of African Americans and participants with a history of cardiovascular disease in this study will limit the power to test for interactions and for any stratified analyses.

7.a. Will the data be used for non-CVD analysis in this manuscript? ____ Yes ___x__ 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 ___x__ 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)?

Manuscript Proposal #2436 Galectin 3 and risk heart failure and death in a subsample of the Atherosclerosis Risk in Communities (ARIC) Study

Manuscript Proposal #2170 sRAGE, progression of subclinical cardiac damage, and risk of heart failure

Manuscript Proposal #2330 Carboxymethyl lysine, an advanced glycation end-product, and incident diabetes – the ARIC Study

Manuscript Proposal # 1890 Determinants of sRAGE and its Association with Cardiovascular Disease, Diabetes, and Mortality in a Community-based Population

Manuscript Proposal # 1905 The Association of Lifestyle Factors with circulating levels of the Soluble Receptor for Advanced Glycation End Products (sRAGE)

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*2006.02)

_____ 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.

13. Per Data Use Agreement Addendum for the Use of Linked ARIC CMS Data, approved manuscripts using linked ARIC 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 ____ No.

References:


