1.a. **Full Title**: Egg consumption, overall diet quality, and risk of type 2 diabetes: a pooling project of US prospective cohorts

b. **Abbreviated Title (Length 26 characters)**: Egg consumption and diabetes

2. **Writing Group**:
   Luc Djousse, Brigham and Women’s Hospital; ldjousse@rics.bwh.harvard.edu
   Lyn M. Steffen, University of Minnesota School of Public Health; steffen@umn.edu 612-625-9307
   Xia (Summer) Zhou; University of Minnesota School of Public Health; zhoux062@umn.edu
   Alexis Wood-Frazier; Baylor University, Houston, TX; leikki.wood@gmail.com
   Brian Steffen; University of Minnesota, Dept of Lab Medicine and Pathology; steff293@umn.edu
   Mike Tsai; University of Minnesota, Dept of Lab Medicine and Pathology; tsai001@umn.edu
   JoAnn Manson; Brigham and Women’s Hospital; imanson@rics.bwh.harvard.edu
   Kenneth J Mukamal, Beth Israel Deaconess Medical Center; kmukamal@bidmc.harvard.edu
   Mary Lou Biggs, Department of Biostatistics, University of Washington, Seattle, WA; milbiggs@uw.edu
   I-Min Lee, Harvard University; ilee@rics.bwh.harvard.edu
   Julie E Buring, Harvard University; jquiring@rics.bwh.harvard.edu
   J Michael Gaziano, Brigham and Women's Hospital
   Cheryl Clark, Brigham and Women's Hospital; cclark@partners.org
   Catherine Tucker, University of Massachusetts, Lowell; katherine_tucker@uml.edu
   Edmond Kabagambe, Vanderbilt University; edmond.kabagambe@vanderbilt.edu

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: Luc Djousse**
Address: Brigham and Women’s Hospital, Division of Aging, OBC 3rd Floor, 1620 Tremont St, Boston, MA 02120
Phone: 617-525-7591 Fax: 
E-mail: ldjousse@rics.bwh.harvard.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: Lyn M. Steffen
Address: UMN School of Public Health, Division of Epidemiology and Community Health; 1300 South Second St, Suite 300; Minneapolis, MN 55454
Phone: 612-625-937 Fax: 612-624-0315
E-mail: steffen@umn.edu
3. **Timeline:**
   1 year
   6 months to conduct the data analysis at each center; begin preparing the manuscript
   6 months to conduct the meta-analysis; complete drafting the manuscript

4. **Rationale:**
   Type 2 diabetes (DM) is highly prevalent in the US and is associated with higher health care costs and societal burden. Thus, it remains important to identify modifiable risk factors that could help reduce the risk of DM and its cardiovascular consequences. Several dietary factors including red/processed meats and whole grains have been associated with elevated and lower risk of DM, respectively. In contrast, the association of egg consumption with incidence of DM has been inconsistent. Egg is not only a major source of dietary cholesterol (~200 mg/egg) but also contains other important nutrients such as minerals, vitamins, proteins, carotenoids, saturated (~1.5 g/egg), polyunsaturated (~0.7 g/egg), and monounsaturated (~1.9 g/egg) fatty acids\(^1\),\(^2\). While several of these nutrients have been associated with an increased risk of DM (i.e. saturated fat and cholesterol\(^3\)-\(^5\)), other nutrients may confer a lower risk of DM (i.e. polyunsaturated fat\(^6\)). In animal experiments, a diet rich in fat has been shown to induce hyperglycemia and hyperinsulinemia\(^6\). In addition, a diet enriched with egg yolk was associated with elevated plasma glucose compared with a control diet in male Wistar albino rats\(^7\). However, data from the Zutphen Study\(^8\) have reported a positive association between egg consumption or dietary cholesterol and fasting glucose. In a randomized trial of 28 overweight/obese subjects on a carbohydrate restricted diet, consumption of 3 eggs/d has no effect on fasting glucose compared to 0 egg/d\(^9\).

Our analysis of the Physicians’ Health Study and the Women’s Health Study showed a positive and linear relation between egg consumption and incident DM\(^10\). However, data from the Cardiovascular Health Study\(^11\) and the Jackson Heart Study\(^12\) did not show an association between egg consumption and incident DM. Our recent meta-analysis of 12 prospective studies showed an elevated risk of DM with 3+ eggs/d in US cohorts but not in non-US cohorts\(^13\). Due to the observational design of studies included in the above meta-analysis, it remains unclear whether observed elevated risk of DM with 3+ egg/d is causal or partially explained by unmeasured confounding by diet quality. It is possible that frequent egg consumption might be associated with frequent intake of other food items known to increase the risk of DM (i.e., red/processed meats\(^14\) or fried foods\(^15\)).

People with DM are at 3-6 fold increased risk of cardiovascular disease. It is noteworthy that evidence linking dietary cholesterol to elevated plasma LDL-cholesterol is weak at present\(^16\)-\(^18\). Although, data on association of egg consumption with CHD or stroke have been inconsistent in the general population\(^19\)-\(^22\), few studies suggested that egg consumption may be associated with a higher risk of mortality, CHD or stroke among people with DM\(^20\),\(^23\)-\(^27\). Shin et al\(^20\) in a prior meta-analysis of prospective cohort studies demonstrated no relation between eggs and CHD or stroke in general, but a 69% increased risk of CVD among diabetic subjects. However, data on dose-response relation between egg consumption and incident CHD among people with DM are lacking. It is unclear whether reported heightened risk of CVD with egg consumption among people with DM is causal. Furthermore, those results have not been replicated in other ethnic groups or general population.

To address above gaps, we are proposing to **pool 9 large US cohorts** to examine whether the association of egg consumption and risk of DM is independent of overall diet quality. In a secondary aim, we will examine whether egg consumption is associated with incidence of coronary heart disease independent of overall diet quality among people with DM. Elucidating the role of egg consumption in the development of DM or CHD is important given the affordability of eggs worldwide as a good source of proteins. Regardless of the direction of the results, our findings will have major public health impact by providing critical information necessary support or revised current dietary recommendations on egg consumption.

9 prospective studies are included in this pooling project:
Women's Antioxidant Cardiovascular Study (WACS)
Cardiovascular Health Study (CHS)
Women’s Health Study (WHS)
The Physicians’ Health Study (PHS)
The Jackson Heart Study (JHS)
The Multi-Ethnic Study of Atherosclerosis (MESA)
The Reasons for Geographic and Racial Differences in Stroke (REGARDS) study
The Atherosclerosis Risk in Communities (ARIC) Study
The Coronary Artery Risk Development in Young Adults (CARDIA) Study

5. Main Hypothesis/Study Questions:

Primary hypothesis: There is a positive association between the frequency of egg consumption and incident DM that is independent of diet quality
Secondary hypothesis: There is a positive association between egg consumption and risk of CHD among people with DM independent of diet quality

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).

This is a meta-analysis of 9 prospective studies about the relation between egg consumption and incident diabetes.

Exposures: Egg consumption assessed at baseline in each cohort using either the picture-sort National Cancer Institute food frequency questionnaire at baseline (response categories: never, 5-10/year, 1-3/month, 1-4/week, and 5+/week) or Willett food frequency questionnaire (response categories: Never, 1-3/month, 1/week, 2-4/week, 5-6/week, 1/d, 2-3/d, 4-5/d, and 6/d.), including the ARIC modified FFQ, or CARDIA Diet History. For meta-analysis, we will use the midpoint as average intake of egg per week in each category. If the highest category of egg consumption has an open upper boundary, we multiplied the lower boundary by 1.5 to obtain an estimate of average egg consumption in that category as previously described. For egg consumption measured in grams per week, we will assume that one average egg is equivalent to 50 grams for conversion.

Outcomes: For the primary analysis, incident DM will be defined using fasting glucose of at least 126 mg/dl, A1C of at least 6.5% or use of hypoglycemic medications.

Diet quality: We will use information from diet data to compute DASH (Dietary Approach to Stop Hypertension) score which will be used to assess overall diet quality. The rationale for selecting DASH over other indices for diet quality (aHEI, aMED, and RFS scores) is supported by the paper by de Koning et al. showing a lower risk of DM with DASH after adjustment for aHEI/aMED and not the reverse. Components of DASH score will include fruits, vegetables, whole grains, nuts/seeds/legumes, low-fat dairy, red and processed meats, sugar-sweetened beverage and fruit juice, and sodium. DASH score ranges from 8 to 40 points.

For each study, analysis will be conducted at each study center. Each cohort will conduct sex-specific Cox-regression analysis to estimate hazard ratio (95% CI) for DM using no egg
consumption as reference group. Adjustment will be made for age, center (if applicable), body mass index, smoking, alcohol intake, education, physical activity, and DASH score (quintiles).

Hazard ratios obtained from each cohort will be used to complete pooled meta-analysis centrally using fixed effect (using inverse weighted variance). We will initially pool relative risks and their 95% CI from the highest versus lowest category of egg consumption in each study. For studies that stratified analyses by gender, we will consider each gender as an independent study. We will assess heterogeneity using Q statistic, I-squared, and p value <0.05. We will assess the presence of influential study using removal of one study at a time method. To assess dose-response relation and evaluate the shape of egg-DM/CHD relation, we will use generalized least squares regression described by Greenland and Longnecker and fit cubic splines with knots at 5, 35, 65, and 95 percentile of egg distribution. Two sided p value was used with alpha level of 0.05. All data analyses will be performed using SAS software (version 9.4, SAS Institute, Cary, NC, USA).

7.a. Will the data be used for non-CVD analysis in this manuscript? ____ 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 ___ 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 overlap

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 ___ X___ 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.cscce.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.cscce.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.