1a. Full Title: Circulating and Tissue Biomarkers of Omega-6 Fatty Acids and Incident Diabetes in the Fatty Acids and Outcomes Research Consortium (FORCE)

b. Abbreviated Title (Length 26 characters): N6 fatty acids and diabetes

2. Writing Group:
Writing group members: Jason Wu, Lyn M. Steffen, Jyrki Virtanen, Fumiaki Imamura, Rozenn Lemaitre, Dariush Mozaffarian, Liana Del Gobbo, Graham Giles, Ulf Riserus, Cecilia Samieri, Markku Laakso, Alexis Frazier-Wood, Nathan Tintle, Allison Hodge, Matti Marklund, Kalina Rajaobelina, Andres Ardisson, Frank Hu

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

First author: Jason Wu, PhD
Address: The George Institute for Global Health | AUSTRALIA
        PO Box M201 | Missenden Rd | NSW 2050 Australia
Phone: Fax:
E-mail: jwu1@georgeinstitute.org.au

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: 1300 South Second St, Suite 300; Minneapolis, MN 55454
Phone: 612-625-9307 Fax: 612-624-0315
E-mail: steffen@umn.edu

3. Timeline:
November 2015, begin analysis;
March 2016, begin drafting manuscript;
June 2016, completion of manuscript draft

4. Rationale:
While the American Heart Association and 2010 Dietary Guideline for Americans recommend increasing intake of long-chain polyunsaturated fatty acid (PUFA) from vegetable oil, primarily the omega-6 (n-6) linoleic acid (LA), as healthy replacement for
saturated fatty acid, (1, 2) the influence and dose-response relationship of n-6 PUFA to health outcomes remains contentious (3-7). Experimental and clinical studies support metabolic benefits of LA, including its well established LDL-cholesterol lowering effect (8, 9). Conversely, a key concern that has been raised relates to endogenous conversion of LA to arachidonic acid (AA), which may lead to increased production of downstream pro-inflammatory and pro-thrombotic AA metabolites (10). Another hypothesized harm of high intake of n-6 PUFA is competition and interference with the potential cardiometabolic benefits of omega-3 (n-3) PUFA, due to their shared metabolic pathways (7). Since LA is the major dietary PUFA from vegetable oils, its effects on health is of considerable scientific and public health importance, and additional studies are needed to improve the evidence base for dietary recommendations. The effect of AA on health also remains poorly established and requires further investigation.

The majority of prior studies has investigated potential cardiovascular effects of n-6 PUFA, and much less is known on noncardiovascular outcomes including type 2 diabetes (T2D). Emerging animal and metabolic experiments suggest LA may improve glucose-insulin homeostasis and reduce liver fat accumulation, an important risk factor for T2D (11-13). Nevertheless, only a limited number of prior prospective studies have assessed the association of n-6 PUFA with incident T2D, with mixed findings (14-18). Many of the prior studies examined self-reported n-6 PUFA intake, which are potentially limited by recall bias and measurement error. An alternative approach to assess exposure to n-6 PUFA is measurement of circulating or adipose level of these fatty acids. An important advantage of objective n-6 PUFA biomarkers is the ability to more precisely assess individual fatty acids, which is important as the FA may differ in their biologic functions (19). Circulating and adipose tissue LA are objective biomarkers of LA intake (20). Conversely, circulating levels of AA tend to show weaker associations with their dietary intake, which suggest endogenous metabolism may play more dominant roles in determining exposure to AA (20, 21).

The overall aim of the present study is to prospectively assess the association of n-6 PUFA biomarkers (LA and AA) with incident T2D in cohorts in the Fatty Acids and Outcomes Research Consortium (FORCE), and pooling the findings by meta-analysis. By collaborating with participating cohorts and using a standardized approach to define exposures, covariates, and outcome variables, our analyses will have reduced heterogeneity and substantially increased statistical power to more precisely quantify the association of n-6 PUFA and T2D. Conducting analyses across cohorts from diverse demographic and dietary backgrounds will also enhance the generalizability of our findings. Finally, the large sample size will enable detailed evaluation of potential effect modifiers of the relation between n-6 PUFA and T2D.

The Fatty acids and Outcomes Research Consortium (FORCe) is an extension of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Fatty Acids Working Group, originally developed to investigate effects of genetic variation on biomarker fatty acid levels. Using direct expert contacts and online searches, we identified large studies that had measured circulating or tissue biomarkers of fatty acids in general populations and had ascertained incident CHD. We focused on prospective (cohort, nested case-control) studies; retrospective studies were included if fatty acids were measured in adipose tissue at the time of the first event given stability of adipose
measures. Studies were asked to join the consortium and participate in standardized pooled analyses of biomarker fatty acids and clinical events.

5. **Main Hypothesis/Study Questions:**

**Aims:**
Specific Aim 1: To investigate whether LA and AA, as assessed by objective biomarkers, are associated with risk of incident T2D in all participating cohorts with available exposures of interest.
Specific Aim 2: To investigate potential effect modification (age, BMI, aspirin use, sex, race, n-3 PUFA biomarker concentration, genotype at rs174547) in the association of LA and AA with incident T2D. To address this aim, we will pool stratified estimates from each study for each potential effect modifier using meta-analysis.

**Hypotheses:**
a. Higher LA biomarkers will be associated with lower incidence of T2D.
b. Circulating and tissue AA will not be associated with incidence of T2D.

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).**

**Search strategy & inclusion criteria:**
A search will be conducted to identify all eligible populations (adults ≥18yrs) for analysis of the association of n-6 PUFA biomarkers with T2D, irrespective of whether or not associations have been published in a given population. All prospective (cohort studies and nested case-control studies) are eligible. Only retrospective case control studies that measured adipose tissue fatty acids at the time of the T2D ascertainment will be included, given the stability of adipose tissue fatty acids.20 During pooling, we will evaluate results after excluding retrospective studies to minimize possible reverse causation and selection bias. If results are appreciably different, the primary analyses will present prospective studies only, and secondary analyses will present all studies combined.

**Methods:** a detailed analysis plan and data extraction template will be sent to participating cohorts.

**Population:** Included participants should meet the following criteria:
3. Adults (≥ 18yrs)
4. For prospective studies, no prevalent type 2 diabetes.

**Exposures:** Exposures to be assessed include all available biomarkers [eg. plasma (total plasma, phospholipid, cholesterol esters, or triglyceride), serum, erythrocytes, adipose, or other, all as % total fatty acids] of the following omega-6 fatty acids:

3) Linoleic acid (LA; 18:2 n-6)
4) Arachidonic acid (AA; 20:4 n-6)

Each exposure will be analyzed in two ways:
1.) as a continuous variable (% total fatty acids, per x unit increment);  
2.) in quintiles, using cohort specific cut-points.

**Outcome definition**

Type 2 diabetes will be defined as 1) fasting glucose concentration ≥ 126 mg/dL (7 mmol/L), 2) 2-hour post oral glucose tolerance test glucose concentration ≥ 200 mg/dL (11 mmol/L), 3) new use of an insulin or oral hypoglycemic medication, 4) Fasting or non-fasting HbA1C concentration ≥ 6.5%, or 5) otherwise as defined by the study.

**Covariates:** Variables will be classified across studies in a standardized fashion based on data available in each cohort, please confirm with Jason jwu1@georgeinstitute.org.au before proceeding with the analysis. These will include:

1.) age (continuous)  
2.) sex (binary; male/female)  
3.) race (binary; Caucasian/non-Caucasian, or study-specific)  
4.) clinical centre, if applicable (study-specific categories)  
5.) BMI (continuous)  
6.) education (categorical; < high school, high school graduate, some college or vocational school, college graduate)  
7.) smoking (categorical; current, former, never, if only 2 categories available, use binary: current, not current)  
8.) physical activity (4 categories; first preference is quartiles of METs. If METs are not available, use four categories of physical or leisure activity as defined in your study)  
9.) alcohol intake (4 categories; none, 1-6 drinks/week, 1-2 drink/day, >2 drink/day. If your study’s alcohol unit is grams, please convert to drinks using the conversion 14 grams alcohol=1 standard drink)  
10.) treated hypertension (binary; yes= hypertension drug use, or no. If this information is not available, use in the following order: a.) diagnosed/history of hypertension, or b.) study-specific definition)  
11.) treated hypercholesterolemia (binary; yes=lipid-lowering drug use, or no. If this information is not available, use in the following order: a.) diagnosed/history of hypercholesterolemia, or b.) study-specific definition)  
12.) prevalent coronary heart disease (binary, yes/no)  
13.) n-3 PUFA (sum of EPA, 20:5n-3, DPA, 22:5n-3, and DHA, 22:6n-3) biomarker concentrations (continuous; % total fatty acids)

**Missing data:** To retain study power, missing indicator categories will be used for missing covariates.

**Survival Analysis:** For prospective cohort studies, Cox proportional hazards models, with robust variance, will be used to estimate the hazard ratio for incident T2D. Follow-up time will be calculated from baseline (biomarker measurement) to date of failure, end of follow-up, loss to follow-up, or death, whichever occurred first.

**Heterogeneity:** To examine heterogeneity, stratified analyses will be conducted. For the following variables, the β coefficient for each n-6 PUFA (as continuous variable) and their robust standard error (SE) will be recorded for each specified strata:

**Primary:**
1.) BMI (<25kg/m², ≥25kg/m²)  
2.) age (< 60 years, ≥ 60 years)

**Secondary:**
1.) sex (male, female)
2.) race (Caucasian, race #2, race #3 etc)
3.) n-3 PUFA biomarker concentrations (EPA+DPA+DHA, < or ≥ median value in the specific study)

Effect estimates in each study-specific strata will be pooled, and statistical significance of differences between subgroups of potential sources of heterogeneity will be assessed using meta-regression.

In studies with available genetic data, we will also investigate potential interaction by variants in FADS desaturase genes (rs174547). To assess interactions for this SNPs, linear regression analysis using an additive genetic model, i.e. regression of phenotype on the number of reference alleles, or equivalently the imputed dosage for imputed genotypes, will be conducted. Interaction terms for each SNP will be constructed by creating a cross-product term of omega-6 fatty acid exposure of interest (continuous) by the SNP (ordinal; 0, 1, or 2 T alleles) and added to the fully adjusted model:
S(x) = \exp(\beta_{n6} + \beta_{SNP} + \beta_{n6 \times SNP} + \ldots)

For each SNP, the \(\beta\) coefficient and its robust standard error (SE) will be recorded for the main effect of the omega-6 fatty acid exposure, the interaction term, and the covariance matrix (In STATA, the variance-covariance matrix is stored in e(V) and can be obtained using the lincom command).

Sensitivity analyses: two separate sensitivity analyses will be conducted on the main models only (models without interaction terms):
3.) Cases identified in first 2 years after biomarker sampling will be excluded to minimize effect of reverse causation due to pre-existing health condition
4.) Participants will be censored at the first 6 years of follow-up to minimize exposure misclassification due to within-person variation over time.

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?  ____ 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
11. 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.csc.c.unc.edu/ARIC/search.php

____X__ Yes _______ No

12. 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*1997.04 Aaron Folsom’s ancillary study of fatty acids)

____ 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 __X__ No.
References


5. Allport S. The queen of fats. Why omega-3s were removed from the western diet and what we can do to replace them. University of California Press: 2008.


