

ARIC Manuscript Proposal #2243

PC Reviewed: 10/8/13
SC Reviewed: _____

Status: A
Status: _____

Priority: 2
Priority: _____

1.a. Full Title: Calibration of analytes over twenty-five years in the Atherosclerosis Risk in Communities Study: The impact of calibration on chronic kidney disease prevalence and incidence

b. Abbreviated Title (Length 26 characters): Calibration and CKD

2. Writing Group:

Writing group members: CM Parrinello, ME Grams, D Couper, CM Ballantyne, RC Hoogeveen, Eckfeldt JH, E Selvin, J Coresh, others?

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

First author: Christina Parrinello
Address: 2024 E. Monument St
Baltimore, MD 21205

Phone: (443) 287-4679 Fax:
E-mail: cparrine@jhsph.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: Josef Coresh, MD, PhD
Address: 2024 E. Monument St
Baltimore, MD 21205

Phone: 410-955-0495 Fax: 410-955-0476
E-mail: coresh@jhu.edu

3. Timeline: We plan to submit the manuscript within one year.

4. Rationale:

Comparability of laboratory measures across multiple visits in large cohort studies is of central importance for studies of disease prevalence, progression, and change over time. Calibration is especially crucial when a disease is categorically defined by marker levels above or below a certain cutoff. Even a small amount of systematic bias can lead to substantial misclassification of disease. While small differences may seem negligible on the individual level, on the cohort level, small differences can shift the entire distribution of the marker, resulting in biased estimates of prevalence and incidence. Careful calibration is needed to ensure comparability of analytes across multiple visits to enable accurate comparisons over time.

Our group has previously successfully conducted lab calibration of markers in large epidemiologic studies (Selvin et al, AJKD 2013; Selvin et al, AJKD 2007, OTHERS?) using Deming regression. For these types of data, Deming regression is the preferred alternative to ordinary least squares regression, since it results in a single regression line, regardless of which measurement is chosen as the independent versus dependent variable. Additionally, since both the uncalibrated and the reference values are likely measured with some degree of error, this approach accounts for measurement error in both variables (Cornbleet PJ and Gochman N, Clin Chem 1979). We intend to use this approach to calibrate fifteen laboratory measures across five visits in the Atherosclerosis Risk in Communities (ARIC) Study. Such a large, comprehensive effort to calibrate such a large number of analytes across multiple visits has not previously been undertaken in ARIC, and will be useful for future longitudinal studies that incorporate these measures.

5. Main Hypothesis/Study Questions:

Aim 1: We will assess the comparability of different analytes across ARIC visits (visits 1 through 5) focusing on those where different assays were used, longitudinal measurements are of scientific interest, and laboratory drift is likely; and to determine calibration corrections for those that lack comparability. The 15 analytes to be included are: creatinine, uric acid, C-reactive protein (CRP), total cholesterol, high density lipoprotein cholesterol, low density lipoprotein cholesterol, triglycerides, glucose, N-terminal probrain natriuretic peptide (NT-proBNP), Troponin T (hs-cTnT), alanine aminotransferase (ALT), aspartate aminotransferase (AST), gamma-glutamyl transpeptidase (GGT), β_2 -microglobulin (B2M) and beta-trace protein (BTP).

Hypothesis 1: There will be systematic differences in certain analytes across visits attributable to different lab methodology (different lab location or different assay), specimen type, or calendar time of measurement, which will require correction or calibration to a reference measure.

Aim 2: We will graphically assess trends of each analyte over time, adjusted for key covariates, and compare these trends before and after calibration.

Hypothesis 2: Trends in each analyte over time will be better aligned after calibration.

Aim 3: To assess the impact of calibration on prevalence and incidence estimates using a specific example of CKD (defined by creatinine-based eGFR [eGFR_{cr}]).

Hypothesis 3: Using calibrated creatinine in the estimation of GFR to define CKD will substantially affect the prevalence and incidence of CKD in ARIC participants.

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).

Design and Methods

Aim 1

We will conduct a lab calibration of fifteen biomarkers that have been measured in participants in the Atherosclerosis Risk in Communities (ARIC) Study using blood specimens from the baseline visit (1987-89) through visit 5 (2011-13). Measurements were obtained at different times, at different labs, using different assay methods and different specimen types (plasma versus serum) across visits. These methodologic differences are potentially problematic and may affect consistency of measures over time.

Among participants who had plasma available at all five visits (and had attended visit five by the time of the calibration study), 200 participants were selected using stratified random sampling with 16 strata based on 5-year baseline age group, gender and race/ethnicity. The ARIC Coordinating Center identified additional exclusions for participants who had events between visits.

We will assess whether systematic error exists across visits for each marker, and if so, quantify the amount and direction of bias in order to make necessary corrections. In general, there are three scenarios concerning:

1. Assays conducted in 2011-2013 at the University of Minnesota (UMN) on frozen serum from visits 2 and 5 (ALT, AST, GGT and B2M). Since the assays are done during the span of a single visit with continuous QC pools (tests for drift will be repeated at the end of the visit), they will be assumed to be stable. Thus, visit 2 and 5 results will be assumed to be similar for assays conducted at UMN during 2011-2013.
2. Assays conducted on frozen visit 4 plasma around 2005 at Baylor (CRP, liver function tests, B2M and BTP). These will likely need calibration to correspond with the other assays done on frozen serum from visit 2 (or V5; or ongoing studies for BTP; n=11,000) during 2011-2013. To do this, UMN results on frozen visit 4 serum (n=200) will be compared to the previously obtained values on

- plasma at Baylor (n~11,000). Thus, visit 4 Baylor plasma values measured in ~2005 will be calibrated to V4 serum values at UMN on the same participants accounting for differences between specimen type, lab, and calendar time in a single regression. It will then be assumed that frozen V4 serum behaves comparably to frozen V2 serum assayed contemporaneously during 2011-2013 at UMN.
3. NT-proBNP and hs-cTnT assays at Baylor on V4 frozen plasma (~2005) and visit 5 plasma (2011-2013 for NT-proBNP and in a batch during 2013 for hs-cTnT). These need to be calibrated to visit 2 assays on frozen serum at UMN conducted during 2011-2013. Baylor has provided data indicating that plasma assays were stable in the lab from ~2005 to 2013, obviating the need for a comparison of visit 4 to visit 5 plasma. Therefore, a serum plasma comparison will be made at Baylor using visit 5 specimens, which are more plentiful. In addition, these assays will be conducted on frozen visit 4 serum at UMN as a part of this calibration study. The frozen visit 4 serum values will be calibrated to Baylor frozen visit 4 plasma values in a regression, which will account for any differences between labs and specimen type. Assuming that frozen visit 4 and visit 2 serum behave similarly at UMN during 2013, this regression can be used to calibrate the frozen visit 2 results (n~12,000) to be comparable to the values at Baylor.

We will calculate descriptive statistics for each the “original” and “new” measurement, as well as for the difference and the mean of the two measurements. Scatter plots will be used to visually compare measurements. We will conduct “iterative outlier removal” using Bland-Altman plots. A Bland-Altman plot of the mean of the 2 measurements versus the difference will be plotted. Observations that are greater than 3 standard deviations away from the mean difference will be considered outliers and will be removed. Another Bland-Altman plot will then be created, excluding the outliers previously determined. The new standard deviation will then be calculated and any additional outliers will be excluded, and so on. This procedure will be used until no additional outliers remain. Deming regression of the “new” versus “original” measure will be used, omitting all outliers identified during the “iterative outlier removal” process, to assess the significance of each the intercept and slope. If calibration is warranted, calibration equations will be determined by the Deming regression and/or the mean difference between the two measures. Differences of less than 5-10% will be considered too small to calibrate.

We will additionally include an appendix describing laboratory methods and calibration recommendations for each analyte.

Aim 2:

We will assess the impact of calibration on trends of the analytes over time. We may expect to see an increase or decrease in certain markers across visits, which may reflect

either actual change in the measure or artifactual change (for instance, due to a change in lab methodology or specimen degradation). After calibration, we expect to see an improvement in alignment of measures across visits, and any trend that remains will be assumed to be due to true change over time, rather than due to methodologic differences.

We will graphically assess the impact of calibration of each analyte on trends versus age across all visits. On the X-axis, we will plot age at each visit. On the Y-axis, we will plot the mean of the analyte at each visit plus the predicted residual from the regression of the analyte on several covariates (gender, race-center, BMI, diabetes [self-reported or medication use], current smoking status and hypertension [DBP \geq 90, SBP \geq 140 or medication use]). Visual assessment of these graphs will aid in determining if changes over time remain after calibration.

We will use a complete case analysis, and include only those participants who have complete data available at each visit that the marker was measured. (For example, for markers measured at visit 5, we will restrict the analysis to only participants who were seen at visit 5 and have data available for both the marker and the adjustment variables.)

Aim 3:

Lastly, we will assess the impact of calibration using the example of chronic kidney disease (CKD), as defined by eGFR (estimated glomerular filtration rate) calculated using creatinine. We consider CKD a salient example since accurately characterizing trajectories of eGFR and disease frequency are particularly important in identifying cases over time.

Creatinine (which was measured at visits 1, 2, 4 and 5) will be used to calculate eGFR_{cr} using the CKD-Epi equation (Levey et al., 2009). We will calculate the prevalence and incidence of CKD (stage 3+) using both uncalibrated and calibrated creatinine to estimate GFR. Prevalent CKD at visit 1 will be defined as eGFR_{cr}<60 mL/min/1.73 m². Incident CKD at visits 2, 4 and 5 will be defined as eGFR_{cr}<60 mL/min/1.73 m² and eGFR_{cr} decline of \geq 25% since baseline.

We will calculate the prevalence of CKD at each visit, using eGFR_{cr} as measured by either uncalibrated or calibrated creatinine. Using both the Kaplan-Meier and a method which accounts for the competing risk of death, we will calculate the cumulative incidence of CKD by visit 5 among persons with no CKD at baseline, again comparing results before and after calibration.

For this aim, we will include all ARIC participants whose CKD status is known at visit 1. For analyses of incident CKD, we will only include participants who are free of CKD at baseline.

Limitations:

We have assumed that the calibration equations derived from the subset of 200 participants included in the calibration substudy apply to the entire ARIC study population. While this is a reasonable assumption, these 200 samples may not have covered the entire range of values for each analyte, and there may be instances in which the calibration is an under- or overestimate.

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.csc.c.unc.edu/ARIC/search.php>
 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)?

The following manuscript proposals intend to analyze measurements across visits, and will therefore apply results from this calibration study to their analyses:

Proposal #2162: Trends in uric acid levels over 25 years in Atherosclerosis Risk in the Communities Study (ARIC)

Proposal #1784: Is the effect of the genetic urate score on the risk of gout fully mediated by serum uric acid levels in the Atherosclerosis Risk in the Communities Study

Proposal #2140: 6-year changes in n-terminal pro-Brain Natriuretic Peptide (NT-proBNP) and metabolic changes: The Atherosclerosis Risk in Community Study

Proposal #2111: The association of cardiac troponin T measured by a highly sensitive assay and arterial stiffness

Proposal #2207: Associations of C-reactive protein over six years with incident diabetes, cardiovascular events and mortality

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* [2009.16](#), [2009.09](#)

_____)

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.csc.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.csc.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.