ARIC Manuscript Proposal #4237

PC Reviewed: 5/09/23	**Status: _____**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Updating cardiovascular risk prediction: A collaborative meta-analysis

b. Abbreviated Title (Length 26 characters): CVD risk prediction update

2. Writing Group:

Writing group members: Sadiya Khan, Kunihiro Matsushita, Yingying Sang, Shoshana Ballew, Josef Coresh, and others for the CKD Prognosis Consortium

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

First author: **Sadiya Khan**
Address: Northwestern University Feinberg School of Medicine, Chicago, IL, USA
Email: s-khan-1@northwestern.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, MHS**
Address: 2024 E. Monument, Suite 2-630
Baltimore, MD 21287
Phone 410-955-0495, Fax 410-955-0476
Email: coresh@jhu.edu

3. Timeline:

Data analysis will start immediately. A manuscript is expected to be prepared within 6 months.

4. Rationale:

Cardiovascular (CVD) risk prediction may benefit from a broader range of outcomes (including heart failure and cardiovascular mortality in addition to CHD, stroke and CHD mortality) and predictors, particularly social determinants of health and measures of kidney function.

The CKD Prognosis Consortium (CKD-PC) is an international consortium established in 2009 after a controversies conference sponsored by the Kidney Disease: Improving Global Outcomes (KDIGO). Since then CKD-PC has been aiming to conduct sophisticated meta-analyses to inform CKD clinical guidelines and improve CKD clinical practice and research. Indeed, several articles from CKD-PC have been cited in the KDIGO 2012 clinical guidelines for CKD and

create a basis for new CKD staging system based on both glomerular filtration rate (GFR) and albuminuria. CKD-PC has published add-on modules to improve the risk prediction of existing recommended equations (pooled cohort equation, PCE and SCORE[1] and SCORE2[2]) with measures of CKD (estimated GFR and albuminuria). These papers included 7 to 9 Million individuals across a wide range of cohorts. The current proposal aims to conduct a more focused analysis to test the ability to improve CVD risk prediction. The primary focus be on US cohorts starting with the core cardiovascular predictors in the PCE but testing the hypothesis that adding SDOH metrics will allow for good discrimination and calibration without explicitly using race.

In addition, the incremental value of adding kidney function estimates and other risk factors which have been widely explored for CVD risk prediction (e.g. biomarkers including NT-pro-BNP and hs-troponin) will be explored. These variables which are less widely available now, are less useful as core predictors for wide clinical use. However, the add-on methodology[3] we used previously may allow for separate addition of factors which are clinically measured and useful at the individual level but may not be warranted for population-wide screening.

GFR estimation equations using serum creatinine, e.g., the MDRD Study or the CKD-EPI equations, have been commonly used in clinical practice and epidemiologic studies.[4] Newer estimation equations using cystatin C with and without serum creatinine have been created and evaluated in previous CKD-PC work.[5,6] The CKD Prognosis Consortium provides a great opportunity to evaluate how the new recommended 2021 eGFR equations predict CVD risk. **ARIC contributes excellent data on serum creatinine, cystatin C, SDOH and risk of subsequent CVD events.**

5. Main Hypothesis/Study Questions:
Risk prediction equations for CVD (overall and its components of CHD, stroke, HF, and cardiovascular mortality) can yield good discrimination and calibration.

Calibration will be adequate across meaningful subgroups (age, sex, race, SDOH quartiles) despite not including race explicitly.

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:
All ARIC participants aged 30-79 at baseline, with data on traditional cardiovascular risk factors will be included. Comorbidity data will be used to exclude prevalent cardiovascular disease.

Exposures:
- Traditional risk factors
    - Age, sex, smoking (current, former, never), diabetes
    - Cholesterol levels (total, HDL, LDL, triglycerides), systolic blood pressure (also diastolic blood pressure and hypertension status)
    - Medications for hypertension, dyslipidemia, and diabetes; fasting status
- Social determinants of health (SDOH) metrics and indices

    - Race, ethnicity, income, education
    - Geographic based indices of SDOH – Social deprivation index (SDI) is preferred since it allows for different geographic levels but ADI, SVI and other indices will be evaluated if available[7].
- Kidney measures
    - Serum creatinine and cystatin C for use in 2021 CKD-EPI equations (eGFRcr, eGFRcys, eGFRcr-cys).[4,5]
    - Albuminuria (urinary albumin-to-creatinine ratio). Albuminuria will be expressed as urinary albumin-to-creatinine ratio (ACR)
- Additional risk factors (for secondary analyses)
    - BMI (height, weight) – usually has independent risk for heart failure but not CHD
    - In diabetes – HbA1c (also fasting glucose)
    - ECG metrics, including heart rate and QT interval
    - Biomarkers, including NT-pro-BNP and hs-troponin T
    - Cardiac imaging (CAC) is only available in selected cohorts and populations

Outcome variables:

- Incident coronary heart disease (CHD) including a hospitalized myocardial infarction (MI), fatal CHD, or cardiac procedure (composite and individually)
- Incident fatal and non-fatal stroke (composite and individually)
- Incident heart failure (HF)
- Cardiovascular mortality
- All-cause mortality

Brief analysis plan and methods:

Various cohorts from the United States will be meta-analyzed. The primary approach will be a multivariate random effects meta-analysis of the results within each cohort from an individual participant level analysis. Participating cohorts are required to send data to the CKD-PC data coordination center, that is, Johns Hopkins University, Baltimore, MD. The data to be sent should not contain a participant IDs. We have IRB approval for our analyses. Since all predictors are conventional and usually measured in the entire study population in most cohorts, we will mainly conduct complete case analysis but as appropriate we will apply the multiple imputation technique. We will conduct the following analysis among the entire study population and subsequently validate risk equations both internally and externally.

We will first evaluate the associations of the major cardiovascular risk factors with each CVD subtype (MI, stroke, HF, CVM and a composite first of any of these). Subsequently, we will assess hypothesized potential interactions (e.g. age*major risk factors and treatment by corresponding risk factor for hypertension and cholesterol). Next, we will examine models which add other risk factors (listed above). Improvement in model performance through addition of new candidate variables in multivariate Cox proportional hazards regression models will be tested using metrics described below. The time horizons for risk prediction will be 2, 5, and 10 years.

In this process, we will also explore whether a new method we have recently developed, "predictor patch"[8], will improve prediction. This new "predictor patch" approach calibrates predicted risk based on two elements: 1) a difference between observed value and expected value

of an additional predictor (in this specific study kidney measures) and 2) hazard ratio of an outcome of interest related to the difference in an additional predictor. Expected values and hazard ratios are obtained from outside datasets.
The overall modeling strategy will include development of both a CKD risk patch and new questions and validation in additional cohorts.

Statistical metrics to evaluate model performance

i. **Discrimination:** Discrimination refers to the ability of a model to correctly distinguish between those with and without outcomes. Concordance statistics (C statistics) and integrated discrimination improvement will be computed as measures of discrimination.
ii. **Calibration:** Calibration describes how closely the predicted probabilities agree numerically with the observed outcomes. We will compare the observed vs. predicted risk of outcomes of interest for each quintile of predicted risk and determined the magnitude of the deviation using slope (observed vs. predicted) and the Gronnesby and Borgan test[9].
    A. **Overall**
    B. **Subgroups:** Within meaningful subgroups by age, sex, race, SDOH quartiles. This will allow us to address the second hypothesis. We recognize that the second hypothesis may or may not be true within any given cohort or overall which is part of the importance of testing it.
iii. **Goodness of Fit:** Overall model fit for sequential models will be compared using the Akaike Information Criterion (AIC).
iv. **Reclassification**: Reclassification improvement will be quantified using the net reclassification improvement (NRI) statistic. To evaluate the effect of definition of risk categories on reclassification, we will calculate NRI using an alternative method that does not require categories (continuous NRI).

Summary/conclusion:
By meta-analyzing various cohorts using individual participant level data and the same models; we will be able to rigorously assess the risk prediction models. These results will serve as key work for future guidelines and patient care. REGARDS would be valuable for appropriate inference in its population compared to all the other research cohorts and large electronic medical records datasets.

7.a. Will the data be used for non-ARIC analysis or by a for-profit organization in this manuscript? ____ Yes ___X_ No

b. If Yes, is the author aware that the current derived consent file ICTDER05 must be used to exclude persons with a value RES_OTH and/or RES_DNA = "ARIC only" and/or "Not for Profit" ? ____ Yes ____ No
(The 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 current derived consent file ICTDER05 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/mantrack/maintain/search/dtSearch.html

__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)?

The most related manuscript proposal is MP 1915 but this further expands risk factors and a composite CVD and subtypes.

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 https://sites.cscc.unc.edu/aric/approved-ancillary-studies

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.

References

1. Matsushita K, Jassal SK, Sang Y, et al. Incorporating kidney disease measures into cardiovascular risk prediction: Development and validation in 9 million adults from 72 datasets. EClinicalMedicine 2020;27:100552.

2. Matsushita K, Kaptoge S, Hageman SHJ, et al. Including measures of chronic kidney disease to improve cardiovascular risk prediction by SCORE2 and SCORE2-OP. Eur J Prev Cardiol 2023;30:8-16.
3. Matsushita K, Sang Y, Chen J, et al. Novel "Predictor Patch" Method for Adding Predictors Using Estimates From Outside Datasets - A Proof-of-Concept Study Adding Kidney Measures to Cardiovascular Mortality Prediction. Circ J 2019;83:1876-82.
4. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med 2009;150:604-12.
5. Inker LA, Schmid CH, Tighiouart H, et al. Estimating glomerular filtration rate from serum creatinine and cystatin C. N Engl J Med 2012;367:20-9.
6. Shlipak MG, Matsushita K, Arnlov J, et al. Cystatin C versus creatinine in determining risk based on kidney function. N Engl J Med 2013;369:932-43.
7. Trinidad S, Brokamp C, Mor Huertas A, et al. Use Of Area-Based Socioeconomic Deprivation Indices: A Scoping Review And Qualitative Analysis. Health Aff (Millwood) 2022;41:1804-11.
8. Matsushita K, Sang Y, Chen J, et al. Novel "Predictor Patch" Method for Adding Predictors Using Estimates From Outside Datasets- A Proof-of-Concept Study Adding Kidney Measures to Cardiovascular Mortality Prediction. Circulation journal : official journal of the Japanese Circulation Society 2019.
9. Demler OV, Paynter NP, Cook NR. Tests of calibration and goodness-of-fit in the survival setting. Stat Med 2015;34:1659-80.