1.a. Full Title: End-stage renal disease (ESRD) prediction in CKD subpopulation

b. Abbreviated Title (Length 26 characters): ESRD prediction in CKD

2. Writing Group:
   Writing group members: Josef Coresh, Mark Woodward, Morgan Grams, Brad Astor, and others for the CKD Prognosis Consortium

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

First author: 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

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 to be used in this proposal are already available. Analyses and manuscript preparation will be performed over the next 6 months.

4. Rationale:
A few risk prediction models for ESRD have been recently published. However, it is unknown to what extent these models perform in broad range of settings. We will adopt a similar approach to the Framingham Study Investigators in the validation of the recently published prediction models for ESRD and conduct meta-analysis with the Chronic Kidney Disease Prognosis Consortium (CKD-PC) as a platform. This analysis will
inform physicians about which patients to treat, possibly delaying treatment in those who ultimately progress to kidney failure, or unnecessarily treating those who do not progress.¹

5. **Main Hypothesis/Study Questions:**
Do recently published prediction models for ESRD¹ using routinely collected data perform well among individuals with CKD stage 3 to 5 from various cohorts from all over the world?

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).**

**Data:**

**Inclusion:**
As done in the original JAMA paper,¹ we will primarily study individuals with eGFR <60 ml/min/1.73 m² (CKD stage 3-5) using the CKD-EPI equation.

**Exposure Variables from ARIC visit 4:**
- Predictors of interest shown in table below (models 1-7 will be tested)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline GFR, per 5 mL/min/1.73m²</td>
<td>1  2   3  4  5  6  7</td>
</tr>
<tr>
<td>Age, per 10 y</td>
<td>X X X X X X X</td>
</tr>
<tr>
<td>Sex</td>
<td>X X X X X X X</td>
</tr>
<tr>
<td>Log urine ACR</td>
<td>X X X X X X X</td>
</tr>
<tr>
<td>Diabetes</td>
<td>X X X X X X X</td>
</tr>
<tr>
<td>Hypertension</td>
<td>X X X X X X X</td>
</tr>
<tr>
<td>Systolic BP, per 10 mm Hg</td>
<td>X X X X X X X</td>
</tr>
<tr>
<td>Diastolic BP, per 10 mm Hg</td>
<td>X X X X X X X</td>
</tr>
<tr>
<td>Body weight, per 10 kg</td>
<td>X X X X X X X</td>
</tr>
<tr>
<td>*Serum albumin, per 0.5 g/dL</td>
<td>X X X X X X X</td>
</tr>
<tr>
<td>*Serum phosphate, per 1.0 mg/dL</td>
<td>X X X X X X X</td>
</tr>
<tr>
<td>†Serum bicarbonate, per 1.0 mEq/L</td>
<td>X X X X X X X</td>
</tr>
<tr>
<td>*Serum calcium, per mg/dL</td>
<td>X X X X X X X</td>
</tr>
</tbody>
</table>

*We are aware these 3 variables are only available at visit 1; †We are aware this variable is not available. The two models which include these 4 variables would not be formally tested but each of three variables available will be assessed as a sensitivity analysis.

- eGFR (serum creatinine). eGFR will be assessed by CKD-EPI epi equation.²
- Albuminuria (urinary albumin-to-creatinine ratio). Albuminuria will be expressed as urinary albumin-to-creatinine ratio (ACR).
- Also included are: race, smoking, and hemoglobin (or hematocrit)

**Outcome Variables:**
- End-stage renal disease (initiation of dialysis, kidney transplantation, death coded due to kidney disease) + Follow-up time

**Analysis Plan:**
The analysis plan will involve comparisons of the original models (fixed beta coefficients and fixed baseline hazards) to recalibrated models (fixed beta, updated baseline hazard), and best fit models (updated beta and updated baseline hazard). The metrics of comparison will include discrimination, calibration and reclassification.

The models shown in the table above will be evaluated for validation. We will also assess a few variables that did not contribute to better prediction in the original JAMA paper but are available in most of the collaborating cohorts (i.e., hemoglobin, smoking, and race).

We will examine the performance of the original prediction models overall, and furthermore by subgroups of interest (e.g., blacks, diabetics, females, etc.). We anticipate that we may see poorer performance in a general population cohort like ARIC due to differences in characteristics compared to the original development and validation cohorts consisting of CKD patients followed by nephrologists. We will analyze the model performance statistics and evaluate them against pre-set thresholds for C statistic, Chi square statistic and net reclassification improvement (NRI), as described below. In the event of limited performance of the original models, we will explore several strategies to improve the accuracy of prediction. These would include recalibration by updating the baseline hazard and/or alternative model coefficients (best cox) overall as well for specific subgroups (e.g., sex or race).

In further model development we will examine the risk prediction at eGFR>60 and additional variables including previous progression of CKD. The estimates in ARIC then will be used for meta-analysis along with results from other CKD-PC collaborating cohorts.

**Risk prediction metrics to be used:**

**Discrimination:** The C statistic will be used as the primary metric of discrimination. For each equation, we will calculate a c statistic at 1, 3 and 5 years, and compare the values from the original model to the best fit model. A C statistic of < 0.8 in external validation, or a drop in C (>0.03) when comparing the best fit model to the original model, will be considered a significant drop in discrimination.

**Calibration:** We will compute chi square statistics comparing quintiles of observed and predicted risk, at 1, 3 and 5 years using the original model, the recalibrated model and the best fit model. A chi square statistic <20 will be considered as evidence of adequate calibration.

**Reclassification:** We will compare nested models (Models 2-7, and 2-5 when applicable) to each other using continuous and categorical measures of reclassification. In addition, we will compare the predictions from the original cox models, to the best cox models, for equations 3 and 6. A continuous NRI >30 % will be considered as significant, whereas a categorical NRI (0, 10, 20 at one year and 0, 5, 15 at five years) of > 10 % will be considered as clinically significant.
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  ____ 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)?

   1362 - Chronic kidney disease and risk of end-stage renal disease: The Atherosclerosis Risk in Communities Study; Bash, LD.

   1574 - Comparison of novel markers of kidney function and prediction of cardiovascular events, mortality, and kidney failure: the Atherosclerosis Risk in Communities (ARIC) Study; Astor, BC.

   1581 - Novel markers of kidney function and prediction of incident chronic kidney disease and end-stage renal disease: the Atherosclerosis Risk in Communities (ARIC) Study; Astor, BC.

   1873 - Combined association of Cystatin C-based and Creatinine-based estimated Glomerular filtration rate (eGFR) with Mortality, Cardiovascular and Renal Outcomes; Waheed, S.
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.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.

References
