

**A Novel “Predictor Patch” Method for Adding Predictors using Estimates from Outside  
Datasets: A Proof-of-Concept Study Adding Kidney Measures to Cardiovascular  
Mortality Prediction**

Kunihiro Matsushita, MD, PhD<sup>1</sup>, Yingying Sang, MS<sup>1</sup>, Jingsha Chen, MS<sup>1</sup>, Shoshana H. Ballew,  
PhD<sup>1</sup>, Michael Shlipak, MD, MPH<sup>2</sup>, Josef Coresh, MD, PhD<sup>1</sup>, Carmen A. Peralta, MD, MAS<sup>2</sup>,  
Mark Woodward, PhD<sup>1,3,4</sup>

<sup>1</sup>Johns Hopkins University, Baltimore, MD

<sup>2</sup>University of California San Francisco, San Francisco, CA

<sup>3</sup> The George Institute for Global Health, University of New South Wales, Australia

<sup>4</sup> The George Institute for Global Health, University of Oxford, UK

Running Title: Matsushita Patch method for risk prediction models

Disclosures: Kunihiro Matsushita received consultancy and research funding from Kyowa  
Hakko Kirin outside of the work and consultancy from Akebia outside of the work.

Correspondence to Dr. Kunihiro Matsushita, Department of Epidemiology, Johns Hopkins  
Bloomberg School of Public Health, Division of Cardiology, Johns Hopkins School of Medicine,  
Welch Center for Prevention, Epidemiology, and Clinical Research. 2024 E. Monument St.,  
Suite 2-600, Baltimore, MD 21287 (Tel: (443) 287-8766, Fax: (410) 367-2384, e-mail:  
[kuni.matsushita@jhu.edu](mailto:kuni.matsushita@jhu.edu)).

Word count: Abstract 219 words and main text 3,263 words

3 tables, 3 figures, and 1 supplementary tables and figures file

# Abstract

**Background:** Cardiovascular guidelines include risk prediction models for decision making which lack the capacity to include novel predictors.

**Methods and Results:** We explored a new “predictor patch” approach to calibrate the predicted risk from a base model according to two components from outside datasets, 1) the difference in observed vs. expected values of novel predictors and 2) hazard ratios for novel predictors, in a scenario of adding kidney measures for cardiovascular mortality. Using four US cohorts (n=54,425) we alternately chose one as the base dataset and constructed a base prediction model with traditional predictors for cross-validation. In the three other “outside” datasets, we developed a linear regression model with traditional predictors for estimating expected values of glomerular filtration rate and albuminuria and obtained their adjusted hazard ratios of cardiovascular mortality, together constituting a “patch” for adding kidney measures to the base model. The base model predicted cardiovascular mortality well in each cohort (c-statistics 0.78-0.91). The addition of kidney measures using a patch significantly improved discrimination (cross-validated  $\Delta$ c-statistic 0.006 [0.004-0.008]) to a similar degree as refitting these kidney measures in each base dataset.

**Conclusions:** The addition of kidney measures using our new “predictor patch” approach based on estimates from outside datasets improved cardiovascular prediction based on

traditional predictors, providing an option to incorporate novel predictors to an existing prediction model.

Key words: risk prediction; novel biomarkers; cardiovascular disease, chronic kidney disease

# Introduction

Risk prediction is a central element for disease prevention and clinical management.<sup>1,2</sup> This is mainly because therapies generally reduce the relative risk of outcomes of interest in a consistent way across different subgroups. Thus, treating high-risk populations is more efficient than treating low-risk populations for the same level of an individual risk factor.<sup>3</sup> For example, blood pressure reduction has been shown to reduce the risk of cardiovascular events by 15-20% regardless of the patient's baseline risk and thus the number needed to treat for preventing an event is much lower in the case of treating higher vs. lower risk patients (e.g., 26 in 5 years if predicted risk >21% vs. 71 if <11%).<sup>3</sup> Indeed, several major clinical guidelines incorporate risk prediction models based on traditional risk factors (e.g., blood pressure and diabetes for cardiovascular disease) for clinical decision making.<sup>1,2,4</sup>

Unfortunately, none of the existing cardiovascular risk prediction tools in major clinical guidelines have the practical ability to incorporate novel biomarkers, despite that several promising biomarkers have been identified (e.g., coronary artery calcium for coronary disease risk prediction<sup>5</sup>). Some of those promising predictors may be expensive to measure and cannot be widely incorporated. However, some other predictors are routinely assessed in clinical practice, but the base datasets, from which risk prediction models in clinical guidelines were derived, often do not have data on those promising predictors, precluding the inclusion of

novel predictors in established risk prediction models.<sup>5</sup>

In this situation, some guidelines categorize specific groups of patients as high-risk outside of a risk prediction scheme (e.g., diabetic patients with albuminuria are considered to be at very high cardiovascular risk in a European guideline<sup>6</sup>) (S1 Fig). However, this approach does not take into account risk variation due to other risk factors or provide any guidance to patients not included in these specific groups (e.g., non-diabetic patients with very high albuminuria). As an alternative approach, some investigators have developed their own risk prediction models with novel predictors to overcome this situation. However, given the difference in the baseline risk of datasets that were used to derive the guideline prediction models and the investigators' models, those two models may provide quite different predicted risk for the same person, which may well lead to lack of uptake amongst clinicians.

To overcome these issues, here we propose a new approach, using a “predictor patch” based on data from outside datasets, to add non-traditional predictors to a base prediction model without remodeling non-traditional predictors in its base dataset. Our approach calibrates predicted risk based on two elements (green text in Fig 1): 1) a difference between observed value and expected value of an additional predictor (e.g., if a predictor has a positive association with an outcome, individuals with higher observed values of an additional predictor than expected would have a higher risk than predicted risk solely based on traditional

predictors) and 2) hazard ratio of an outcome of interest related to the difference in an additional predictor (e.g., calibration for a difference between observed vs. expected would be larger if the association of an additional predictor with an outcome is stronger). Expected values and hazard ratios are obtained from outside datasets.

To explain our methodology we chose a scenario of wanting to add two chronic kidney disease (CKD) measures (glomerular filtration rate [GFR] and urine albumin-to-creatinine ratio [ACR] <sup>7</sup>) to a base risk prediction model with traditional predictors in the context of cardiovascular mortality risk. This is a practical example, since we have previously shown that these kidney markers do, indeed, add to cardiovascular risk prediction.<sup>8</sup> Also, major cardiovascular disease guidelines acknowledge these kidney measures as important predictors but none of the established risk models in ~~cardiovascular disease~~those guidelines ~~do not~~ include these kidney markers. ~~Also~~In addition, both GFR and albuminuria are often measured in clinical practice.<sup>8,9</sup>

## Materials and methods

### Study populations

To test our novel approach and expect reasonable generalizability, we used data from four US community-based cohort studies, each of which had data on traditional cardiovascular

predictors, kidney measures, and cardiovascular mortality. These were the Atherosclerosis Risk in Communities (ARIC) Study, the Multi-Ethnic Study of Atherosclerosis (MESA), the National Health and Nutrition Examination Survey (NHANES) III, and NHANES 1999-2010. Details of each study were reported previously.<sup>10-12</sup> Briefly, ARIC enrolled 15,792 mostly white and black men and women from four US communities (Washington County, Maryland, suburban Minneapolis, Minnesota, Jackson, Mississippi, and Forsyth County, North Carolina) in 1987-1989. For this study, we used data from 9,351 participants who attended visit 4 (1996-1998, age range 52-75 years) and were free of a history of cardiovascular disease (coronary heart disease and stroke). MESA enrolled 6,814 participants aged 45-84 years without a history of cardiovascular disease in 2000–2002 from six US communities: Los Angeles County, California; Chicago, Illinois; Baltimore, Maryland; St Paul, Minnesota; northern New York City, New York; and Forsyth County, North Carolina). MESA was designed to include whites, blacks, Hispanic, and Chinese free of cardiovascular disease at baseline. For this study, we included 6,704 participants with data on kidney measures and traditional risk factors detailed below. For NHANES, we included white, black, and Hispanic men and women without a history of cardiovascular disease with data on GFR and albuminuria, comprising 14,103 participants from NHANES III and 24,267 participants from NHANES 1999-2010.



## **Kidney measures**

Estimated GFR (eGFR) was calculated using the CKD-EPI creatinine equation in all four studies.<sup>7,13</sup> Serum creatinine was measured using a modified kinetic Jaffé method in ARIC,<sup>14</sup> rate reflectance spectrophotometry using thin film adaptation of the creatine amidinohydrolase method in MESA,<sup>10</sup> and a kinetic Jaffé method in NHANES.<sup>15</sup> Urine albumin was measured by nephelometry in ARIC,<sup>14</sup> the Array 360 CE Protein Analyzer in MESA,<sup>10</sup> and solid-phase fluorescent immunoassay in NHANES.<sup>15</sup> Urine creatinine was measured using the Jaffé method in ARIC and NHANES and the Vitros 950IRC instrument in MESA.

## **Traditional cardiovascular predictors**

We considered all the predictors used in the American Heart Association (AHA) and the American College of Cardiology (ACC) Pooled Cohort Equation as traditional risk factors: age, sex, race/ethnicity, systolic blood pressure, use of antihypertensive medications, diabetes, total and high-density lipoprotein cholesterol, and smoking.<sup>5</sup> Data collection of these traditional predictors was conducted according to a standard protocol in each study, as reported previously.<sup>10-12</sup> Race/ethnicity was categorized into whites, blacks, Hispanics, and Asians. Diabetes mellitus was defined as a fasting glucose  $\geq 7.0$  mmol/L, self-reported history of diabetes, or use of glucose lowering medications. Smoking status was dichotomized as

current vs. former/never.

## Outcome

Given its consistent availability in all four cohorts, the outcome of interest was cardiovascular mortality. Cardiovascular mortality was defined as death from myocardial infarction, stroke, heart failure, or sudden cardiac death.<sup>16</sup>

## Statistical analysis

Analyses were conducted using Stata/MP 14 ([www.stata.com](http://www.stata.com)). A P-value below 0.05 was considered significant. Baseline characteristics were summarized as mean (SD) or median [interquartile interval, IQI] if continuous variables and number (%) if categorical, across the four cohorts.

A scheme for developing a “patch” in outside datasets and applying it to a base dataset is shown in S2 Fig. We first selected three outside datasets and developed a linear regression model for estimating expected values of eGFR and log-ACR based on traditional predictors (ARIC, NHANES III, and NHANES 1999-2010 in S2 Fig as an example). Then, we obtained the log hazard ratio ( $\beta$ ) for cardiovascular mortality for eGFR ( $\beta_{\text{eGFR}}$ ) and log-ACR ( $\beta_{\text{logACR}}$ ), adjusted for traditional predictors.  $\beta_{\text{eGFR}}$  and  $\beta_{\text{logACR}}$  were estimated in each of the three outside

datasets first and then meta-analyzed using fixed-effect models. In estimating the log hazard ratios, we fixed the log hazard ratios for traditional predictors at their values from the base dataset. The linear model for estimating expected eGFR and ACR according to traditional predictors and the log hazard ratio for eGFR and ACR together from outside datasets constituted the “CKD patch”.

Then, we applied the “CKD patch” to a base dataset (MESA in S2 Fig) and calibrated the predicted risk according to the difference between observed kidney measures in a base dataset vs. expected kidney measures (based on traditional predictors) and the log hazard ratio for kidney measures in each participant in the base dataset. The formula for this calibration is as follows:

$$h_i(t)_{\text{new}} = h_i(t)_{\text{original}} * \exp(\beta_{\text{eGFR}} * [\text{observed eGFR}_i - \text{expected eGFR}_i]) + \beta_{\text{logACR}} * [\text{observed logACR}_i - \text{expected logACR}_i],$$

where  $i$  indicates person  $i$  in the base dataset;  $h(t)_{\text{new}}$ , calibrated hazard incorporating observed values of CKD measures;  $h(t)_{\text{original}}$ , original predicted hazard with traditional predictors in the base dataset;  $\beta_{\text{eGFR}}$  ( $\beta_{\text{logACR}}$ ), log hazard ratio of eGFR (logACR) from the three outside datasets, observed  $\text{eGFR}_i$  ( $\text{logACR}_i$ ), observed values of eGFR (logACR) in person  $i$  in the base dataset; expected  $\text{eGFR}_i$  ( $\text{logACR}_i$ ), expected values of eGFR (logACR) based on observed traditional predictors in person  $i$  in the base dataset according to a linear

model from the three outside datasets.

It is important that the expected value of the kidney measures is derived based on the traditional predictors since the coefficients of traditional predictors in the base model may include residual confounding by the kidney measures. For example, the coefficient for hypertension will include the risk associated with the higher average level of albuminuria in this group but not individual deviations in albuminuria from this average confounding.

To evaluate the performance of the “CKD patch”, we plotted predicted and observed risk from the base model with only traditional predictors as well as when we used “CKD patch”. Also, we compared the CKD patch to a fully refit model incorporating both traditional predictors (i.e., age, sex, race/ethnicity, systolic blood pressure, use of antihypertensive medications, diabetes, total and high-density lipoprotein cholesterol, and smoking) and kidney measures (i.e., eGFR and ACR) in a base dataset. This simulates the scenario in which the base dataset has data of the non-traditional predictors. We also quantified the difference in Harrell’s c-statistics<sup>17</sup> among the base model, the CKD patch, and the fully refit model. In addition, we assessed categorical net reclassification improvement (NRI),<sup>18</sup> with risk categories of 5% and 10% in 10 years.<sup>4</sup> We repeated the entire process for each cohort as the base dataset (and the remaining three cohorts as outside datasets) for cross-validation.

# Results

## Study characteristics

Characteristics of each study are summarized in Table 1. The average age was similar between ARIC and MESA (~62 years) and between the two NHANES cohorts (~45 years). Approximately 20-30% were blacks across the four cohorts. Reflecting age differences, risk factor profiles were generally better (lower prevalence of diabetes, lower blood pressure, higher kidney function) in the NHANES cohorts than in the other two cohorts. However, ACR was higher on average in the NHANES cohorts than in ARIC and MESA. The prevalence of current smokers was higher in the NHANES cohorts than the other two cohorts, but the combined prevalence of current and former smokers was similar in all four cohorts. We did not see an evident difference in total or HDL cholesterol across the cohorts.

## Estimation of expected kidney measures

In three outside datasets, using traditional predictors, we developed a linear regression model to estimate expected levels of eGFR and log-ACR conditional on the traditional predictors (estimation model) (S1 and S2 Tables). Then, we applied each estimation model to a base dataset. The root mean square errors ranged from 13.91-15.67 for eGFR (S1 Table) and 0.50-0.73 for log-ACR (S2 Table).

Results for expected values of CKD measures (based on an estimation model from MESA, NHANES III, and NHANES 1999-2010) and their observed values in ARIC are shown in Fig 2. There were a number of participants with considerably lower eGFR and higher ACR than expected, indicating that, in those individuals, predicted risk based on only traditional predictors was likely to underestimate their actual risk. The opposite is true for participants with considerably higher eGFR or lower ACR than expected. In the scenario of Fig 2, there were 29% of ARIC participants who had 15 ml/min/1.73m<sup>2</sup> lower or higher eGFR than expected and 17% who had 8-fold higher or lower ACR than expected. Results for the other three scenarios with each of MESA, NHANES III, and NHANES 1999-2010 as the base dataset are shown in S1-S3 Figs.

## **Adjusted hazard ratios of kidney measures with cardiovascular mortality**

Based on meta-analysis of three outside datasets, both kidney measures were generally associated with cardiovascular mortality independently of each other and traditional cardiovascular risk factors (Table 2). However, results for lower eGFR were not necessarily consistent across the four combinations of three outside datasets. For example, 15 ml/min/1.73m<sup>2</sup> lower eGFR in the range below 60 ml/min/1.73m<sup>2</sup> was significantly associated

with cardiovascular mortality except when MESA, NHANES III, and NHANES 1999-2010 were treated as the three outside datasets, whereas lower eGFR in the range 60-90 ml/min/1.73m<sup>2</sup> showed a significant association only with MESA, NHANES III, and NHANES 1999-2010 or ARIC, MESA, and NHANES 1999-2010 as the outside datasets. In contrast, ACR was consistently related to higher risk of cardiovascular mortality regardless of the combinations of outside datasets, with adjusted hazard ratios between 1.4-1.6 per 8-fold higher values. S3 Table displays hazard ratios for kidney measures from each dataset used for the meta-analyzed hazard ratios in Table 2. S4 Table summarizes adjusted hazard ratios of kidney measures from a fully refit model including traditional predictors in each dataset.

## **Comparison of risk prediction**

Using estimation models of eGFR and ACR as well as adjusted hazard ratios of cardiovascular mortality for these two kidney measures from three outside datasets, we calibrated predicted risk according to observed values of eGFR and ACR in every participant in a base dataset using the formula in the Methods. Fig 3 demonstrates the predicted vs. observed risk in each scenario when ARIC, MESA, NHANES III, and NHANES 1999-2010 served as a base dataset, respectively, and the remaining three were treated as outside datasets. Black dots represent predicted vs. observed risk for a base model with only traditional risk factors in a base dataset.

Red dots reflect the calibrated risk prediction based on the “CKD patch” (“base model + CKD patch” in Fig 3), whereas blue dots represent predicted risk with a model including both traditional predictors and the two kidney measures in a base dataset (“fully refit model” in Fig 3).

Overall, all three lines were around the diagonal line of identity, indicating overall good calibration. In all four scenarios, the highest risk decile based on the “CKD patch” (red dot) was shifted towards the right upper corner from the highest risk decile for a base model (black dot), indicating that the addition of kidney measures with the “CKD patch” contributed to identifying individuals at higher risk than originally predicted by the base model with only traditional predictors and indeed they had higher risk. In general, the shift for the remaining risk deciles was less evident than that for the highest risk decile. The patterns for fully refit models (blue dots) were largely similar when the “CKD patch” was used (red dots).

Subsequently, we contrasted c-statistics across three approaches (base model, base model + “CKD patch”, and fully refit model) in the four scenarios (ARIC, MESA, NHANES III, and NHANES 1999-2010) of a base dataset (Table 3). C-statistics ranged from 0.779 to 0.909 for the model with traditional predictors in a base model in each study. The fully refit models showed modest but significant improvement of c-statistics in all the scenarios except MESA. The “CKD patch” significantly improved the risk discrimination of cardiovascular mortality from



a base model in all studies except MESA as well. The degree of improvement by the “CKD patch” was largely similar to fully refit models in all studies, with the pooled cross-validated difference in the c-statistic of 0.006 (95% CI 0.004-0.008) in both approaches. [We also confirmed a significantly positive NRI with CKD patch approach \(S5 Table\).](#)

## Discussion

Here, we demonstrate a new approach, using a “predictor patch” based on estimates from outside datasets, to incorporate non-traditional predictors into a base model without fully refitting both non-traditional and traditional predictors in a base dataset from which a base prediction model was derived. In the scenario of adding kidney measures, our new approach with a “CKD patch” demonstrated risk prediction improvement compared to a base model. The degree of improvement was similar to fully refitting models, although they were not identical in each study. Importantly, we confirmed cross-validation of our new approach in four scenarios by treating one of four US community-based cohorts as a base dataset and the remaining three as outside datasets in turn. This approach will be an option to take into account non-traditional predictors in clinical practice when it is not practical to develop a new model in datasets used to derive relevant existing prediction models. For example, kidney measures are not uniformly measured in cohorts that derived the AHA/ACC Pooled Cohort Equations.<sup>5</sup>

We are not aware of any attempts to use the approach we explored in this study in order to implement new predictors in existing prediction models even when relevant derived datasets do not have data on relevant new predictors. Theoretically, our approach can be applied to any predictors and any outcomes. Our approach would be particularly helpful when predictors are already collected in clinical practice, but healthcare providers could not effectively incorporate information of those predictors. This is exactly the scenario for kidney measures tested here. Indeed, serum creatinine for eGFR is measured approximately 300 million times in the US every year,<sup>19</sup> and the assessment of albuminuria is recommended in patients with diabetes, hypertension, and CKD,<sup>7,20,21</sup> but none of the major clinical guidelines adopt cardiovascular risk prediction tools incorporating these kidney measures.

There are some conditions for our approach to be effective in settings beyond what we tested here for cardiovascular mortality by adding CKD measures. First, the strong associations between non-traditional predictors and outcomes of interest would be required as is true for any prediction model.<sup>22</sup> Second, the estimation of expected values of non-traditional predictors using traditional predictors should fit reasonably well but should have enough residual variance to model a difference between observed and expected values. If this estimation is perfect without any residual variation, then the information in the additional predictors is fully incorporated into the existing risk factors and there is no need for a patch.

However, it seems extremely rare to see such a perfect estimation for most predictors.<sup>23</sup>

There are several limitations in our study. All cohorts were from the US and community-based, and thus, whether and to what extent our new approach works in other regions, ~~or~~ countries, or specific clinical populations should be tested in future studies. In this context, cardiovascular mortality risk prediction tested in this study is particularly relevant to Europe since the European Society of Cardiology Guideline for Cardiovascular Prevention is based on predicted risk of cardiovascular mortality using SCORE <sup>6</sup>. Given our experience and findings of strong associations, we explored kidney measures and cardiovascular mortality for this proof-of-concept study.<sup>9,24</sup> Thus, confirmatory studies for other non-traditional predictors and other outcomes would be needed. Also, it is not necessarily certain as to how many predictors can be appropriately added using this patch approach, although we explored a scenario of adding two predictors and can technically add more predictors. It is likely to depend on correlations among additional and traditional predictors as well as the degree of confounding by additional predictors. Nonetheless, this should be explored in future studies.—

In conclusion, we have demonstrated a new approach, “predictor patch” based on estimates from outside datasets, to incorporate additional predictors to an existing prediction model without remodeling those non-traditional predictors in its derivation base dataset. Although confirmatory investigations are needed, theoretically, the “predictor patch” approach

described here can be applied to a wide range of settings and will allow researchers and healthcare providers to efficiently adopt additional predictors for improving risk prediction in the context of existing prediction models.

## Acknowledgements

The ARIC study has been funded in whole or in part with Federal funds from the National Heart, Lung, and Blood Institute, National Institutes of Health, Department of Health and Human Services, under Contract nos. HHSN268201700001I, HHSN268201700002I, HHSN268201700003I, HHSN268201700005I, HHSN268201700004I. MESA was supported by National Heart, Lung, and Blood Institute Contracts N01-HC-95159–N01-HC-95169 and National Center for Research Resource Grants UL1-RR-024156 and UL1-RR-025005. The authors thank the staff and participants of the ARIC study for their important contributions. The authors also thank the other investigators, the staff, and the participants of MESA for their valuable contributions. A full list of participating MESA investigators and institutions can be found at <http://www.mesa-nhlbi.org>.

## References

1. Stone NJ, Robinson JG, Lichtenstein AH, et al. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation* 2014;**129**:S1-45.
2. Whelton PK, Carey RM, Aronow WS, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *J Am Coll Cardiol* 2017.
3. Blood Pressure Lowering Treatment Trialists C, Sundstrom J, Arima H, et al. Blood pressure-lowering treatment based on cardiovascular risk: a meta-analysis of individual patient data. *Lancet* 2014;**384**:591-598.
4. Piepoli MF, Hoes AW, Agewall S, et al. 2016 European Guidelines on cardiovascular disease prevention in clinical practice: The Sixth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of 10 societies and by invited experts)Developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). *Eur Heart J* 2016;**37**:2315-2381.
5. Goff DC, Jr., Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation* 2014;**129**:S49-73.
6. European Guidelines on cardiovascular disease prevention in clinical practice (version 2012): The Fifth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of nine societies and by invited experts)Developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). *Eur Heart J* 2012;**33**:1635-1701.
7. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. *Kidney International Supplement* 2013;**3**:1-150.
8. Matsushita K, Coresh J, Sang Y, et al. Estimated glomerular filtration rate and albuminuria for prediction of cardiovascular outcomes: a collaborative meta-analysis of individual participant data. *The lancet Diabetes & endocrinology* 2015;**3**:514-525.
9. Matsushita K, van der Velde M, Astor BC, et al. Association of estimated glomerular

filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: a collaborative meta-analysis. *Lancet* 2010;**375**:2073-2081.

10. Matsushita K, Sang Y, Ballew SH, et al. Subclinical atherosclerosis measures for cardiovascular prediction in CKD. *J Am Soc Nephrol* 2015;**26**:439-447.
11. Hui X, Matsushita K, Sang Y, Ballew SH, Fulop T, Coresh J. CKD and cardiovascular disease in the Atherosclerosis Risk in Communities (ARIC) study: interactions with age, sex, and race. *Am J Kidney Dis* 2013;**62**:691-702.
12. Shafi T, Matsushita K, Selvin E, et al. Comparing the association of GFR estimated by the CKD-EPI and MDRD study equations and mortality: the third national health and nutrition examination survey (NHANES III). *BMC Nephrol* 2012;**13**:42.
13. Levey AS, Stevens LA. Estimating GFR using the CKD Epidemiology Collaboration (CKD-EPI) creatinine equation: more accurate GFR estimates, lower CKD prevalence estimates, and better risk predictions. *Am J Kidney Dis* 2010;**55**:622-627.
14. Matsushita K, Sang Y, Ballew SH, et al. Cardiac and kidney markers for cardiovascular prediction in individuals with chronic kidney disease: the atherosclerosis risk in communities study. *Arterioscler Thromb Vasc Biol* 2014;**34**:1770-1777.
15. Shafi T, Matsushita K, Selvin E, et al. Comparing the association of GFR estimated by the CKD-EPI and MDRD study equations and mortality: the third national health and nutrition examination survey (NHANES III) examination survey (NHANES III). *BMC Nephrol* 2012;**13**:42.
16. Matsushita K, Coresh J, Sang Y, et al. Estimated glomerular filtration rate and albuminuria for prediction of cardiovascular outcomes: a collaborative meta-analysis of individual participant data. *The lancet Diabetes & endocrinology* 2015;**3**:514-525.
17. Pencina MJ, D'Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Stat Med* 2004;**23**:2109-2123.
18. Pencina MJ, D'Agostino RB, Sr., Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med* 2011;**30**:11-21.
19. Inker LA, Levey AS. Pro: Estimating GFR using the chronic kidney disease epidemiology collaboration (CKD-EPI) 2009 creatinine equation: the time for change is now. *Nephrol Dial Transplant* 2013;**28**:1390-1396.
20. American Diabetes Association. Executive summary: Standards of medical care in diabetes--2012. *Diabetes Care* 2012;**35 Suppl 1**:S4-S10.
21. Mancia G, De Backer G, Dominiczak A, et al. 2007 ESH-ESC Practice Guidelines for the Management of Arterial Hypertension: ESH-ESC Task Force on the Management of Arterial Hypertension. *J Hypertens* 2007;**25**:1751-1762.

22. Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am J Epidemiol* 2004;**159**:882-890.
23. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009;**150**:604-612.
24. Matsushita K, Coresh J, Sang Y, et al. Estimated glomerular filtration rate and albuminuria for prediction of cardiovascular outcomes: a collaborative meta-analysis. *Lancet Diabetes-Endocrinol* 2015;**3**:514-525.



## Table and Figure Legends

**Table 1. Characteristics of four US community-based cohorts**

**Table 2. Pooled hazard ratios (95% CI) for cardiovascular mortality according to kidney measures in three "outside" datasets**

**Table 3. C-statistics for cardiovascular mortality across three models.**

**Figure 1. Conceptual scheme for calibrating the risk based on 1. the difference between observed and expected values of CKD measures and 2. the relative risk according to CKD measures**

**Figure 2. Scatter plot of residual (observed minus expected) vs. expected eGFR and  $\log_8$ -ACR in ARIC.** Expected values are based on a linear regression model from the other three cohorts (MESA, NHANES III, and NHANES 1999-2010), regarded as outside datasets.

**Figure 3. Calibration plot (observed vs. predicted risk) for cardiovascular mortality risk for three models.** “base model” indicates a model with traditional predictors in a base dataset, “base model + CKD patch” represents a calibrated predicted risk using CKD patch, and “fully refit model” reflects a model with traditional predictors and CKD measures in a base dataset.

## Supporting Information

**S1 Fig. Summary of current approaches for dealing with non-traditional predictors and their limitations.**

**S2 Fig. A scheme of our approach, “CKD patch”.** This scheme summarizes key steps of developing, applying, and evaluating “CKD patch”. This scheme shows a scenario where MESA was treated as base dataset but we repeated the process with each of the four studies to be the base dataset.

**S3 Fig. Scatter plot of residual vs. expected eGFR and  $\log_8$ -ACR in MESA.** Expected values are based on a linear regression model from the other three cohorts (ARIC, NHANES III, and NHANES 1999-2010), regarded as outside datasets.

**S4 Fig. Scatter plot of residual vs. expected eGFR and log-ACR in NHANES III.** Expected values are based on a linear regression model from the other three cohorts (ARIC, MESA, and NHANES 1999-2010), regarded as outside datasets.

**S5 Fig. Scatter plot of residual vs. expected eGFR and log-ACR in NHANES 1999-2010.** Expected values are based on a linear regression model from the other three cohorts (ARIC, MESA, and NHANES III), regarded as outside datasets.

**S1 Table. Coefficients to estimate eGFR for each traditional risk factor in combinations of three outside datasets.**

**S2 Table. Coefficients to estimate  $\log_8$ -ACR for each traditional risk factor in combinations of three outside datasets.**

**S3 Table. Hazard ratios (95% CI) for kidney measures in each of three outside datasets by plugging  $\beta$  (log hazard ratio) for traditional predictors from a base dataset.**

**S4 Table. Hazard ratios (95% CI) for kidney measures from the fully refit model including traditional predictors in each cohort.**

**S5 Table. Categorical net reclassification improvement (NRI) for cardiovascular mortality after adding CKD measures using CKD patch and fully refit models to the base model**