The kidney failure risk equation

Grams, Morgan E.; Brunskill, Nigel J.; Ballew, Shoshana H.; Sang, Yingying; Coresh, Josef; Matsushita, Kunihiro

Published in:
Journal of the American Society of Nephrology

DOI:
10.1681/ASN.0000000000000050

Publication date:
2023

Document Version
Peer reviewed version

The Kidney Failure Risk Equation: Evaluation Of Novel Input Variables Including eGFR Estimated Using the CKD-EPI 2021 Equation in 59 cohorts

Morgan E Grams, MD, PhD*, Nigel Brunskill, MD, PhD*, Shoshana H Ballew, PhD, Yingying Sang, MSc, Josef Coresh, MD, PhD, Kunihiro Matsushita, MD, PhD, Aditya Surapaneni, PhD, Samira Bell, MD, Juan J Carrero, PharmD, PhD, Gabriel Chodick, PhD, Marie Evans, MD, PhD, Hiddo JL Heerspink, PhD, Lesley A Inker, MD, MS, Kunitoshi Iseki, MD, PhD, Philip A Kalra, MD, H Lester Kirchner, PhD, Brian J Lee, MD, Adeera Levin, MD, Rupert W Major, MD, PhD, James Medcalf, MD, PhD, Girish N Nadkarni, MD, MPH, David MJ Naimark, MD, MSc, Ana C Ricardo, MD, Simon Sawhney, MD, PhD, Manish M Sood, MD, MSc, Natalie Staplin, PhD, Nikita Stempniewicz, MSc, Benedicte Stengel, MD, PhD, Keiichi Sumida, MD, MPH, Jamie P Traynor, MD, Jan van den Brand, PhD, Chi-Pang Wen, MD, DrPH, Mark Woodward, PhD, Jae Won Yang, MD, Angela Yee-Moon Wang, MD, PhD†, Navdeep Tangri, MD, PhD†

*Indicates co-first authors; †Indicates co-last authors

Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD (Grams, Ballew, Sang, Coresh, Matsushita, Surapaneni, Woodward)

Department of Medicine, New York University Grossman School of Medicine, New York, NY (Grams)

Department of Cardiovascular Sciences, University of Leicester, Leicester, United Kingdom (Brunskill, Major, Medcalf)

John Walls Renal Unit, Leicester General Hospital, University Hospitals of Leicester NHS Trust, Leicester, United Kingdom (Brunskill, Major, Medcalf)

Renal Unit, Ninewells Hospital, Dundee, United Kingdom and Division of Population Health and Genomics, School of Medicine, University of Dundee, Dundee, United Kingdom (Bell)

Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Huddinge, Sweden (Carrero)

Medical Division, Maccabi Healthcare Services, and Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel (Chodick)

Department of Clinical Intervention, and Technology (CLINTEC), Karolinska University Hospital and Karolinska Institutet, Stockholm, Sweden (Evans)

Department of Clinical Pharmacy and Pharmacology, University of Groningen, University Medical Center, Groningen, Netherlands (Heerspink)

Tufts Medical Center, Boston, MA (Inker)

Okinawa Heart and Renal Association, Okinawa, Japan (Iseki)

Department of Renal Medicine, Salford Royal Hospital, Northern Care Alliance NHS Foundation Trust, Salford, United Kingdom (Kalra)

Department of Population Health Sciences, Geisinger, Danville, PA (Kirchner)

Kaiser Permanente, Hawaii Region, and Moanalua Medical Center, Honolulu, Hawaii (Lee)

Division of Nephrology, University of British Columbia, Vancouver, Canada (Levin)

UK Renal Registry, The Renal Association, Bristol, UK and Department of Cardiovascular Sciences, University of Leicester, Leicester, UK (Medcalf)

Department of Medicine, Division of Nephrology, Icahn School of Medicine at Mount Sinai, New York, New York (Nadkarni)

Sunnybrook Hospital, University of Toronto, Toronto, ON, Canada (Naimark)

Department of Medicine, University of Illinois, Chicago, Illinois (Ricardo)

University of Aberdeen, Aberdeen, Scotland, UK (Sawhney)

Division of Nephrology, Department of Medicine, University of Ottawa, Ottawa, Canada, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada and Division of Nephrology, Department of Medicine, The Ottawa Hospital, Ottawa, Ontario, Canada (Sood)

MRC Population Health Research Unit, Clinical Trial Service Unit and Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, Oxford, UK (Staplin)

AMGA (American Medical Group Association), Alexandria, Virginia and OptumLabs Visiting Fellow (Stempniewicz)

University Paris-Saclay, Inserm, Clinical Epidemiology team, CESP, Villejuif, France (Stengel)

Division of Nephrology, Department of Medicine, University of Tennessee Health Science Center, Memphis, TN (Sumida)

Glasgow Renal and Transplant Unit, Queen Elizabeth University Hospital Glasgow Scotland, UK (Traynor)

Department of Nephrology, Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, The Netherlands (van den Brand)

National Health Research Institutes, Miaoli, Taiwan and China Medical University Hospital, Taichung, Taiwan (Wen)

George Institute for Global Health, University of New South Wales, Sydney, Australia (Woodward)

George Institute for Global Health, Imperial College London, London, United Kingdom (Woodward)

Department of Internal Medicine, Yonsei University Wonju College of Medicine, Wonju, Republic of Korea (Yang)

Department of Medicine, Queen Mary Hospital, The University of Hong Kong, Hong Kong (Wang)

Division of Nephrology, Department of Medicine, University of Manitoba, Winnipeg, Canada (Tangri)

Running title: KFRE and novel input variables

Word count: abstract: 250; **text (excluding methods/references/tables/figures):** 2231

Address for corresponding author: Chronic Kidney Disease Prognosis Consortium (Co-PIs: Drs. Josef Coresh and Morgan Grams), 2024 E. Monument Street, Baltimore, MD, 21287; email: ckdpc@jhmi.edu

Significance Statement

The widely-used kidney failure risk equation (KFRE) predicts the 2- and 5-year the risk of kidney failure in populations with eGFR <60 ml/min/1.73 m^2. We assessed whether predictive performance and calibration was acceptable in 59 cohorts when using the new CKD-EPI 2021 eGFR equation, and whether the KFRE could be improved with additional kidney/cardiovascular variables and accounting for the competing risk of death. The KFRE generally performed well using CKD-EPI 2021 eGFR. Adding previous eGFR slope and cardiovascular variables resulted in no improvement in performance. However, the KFRE showed systematic underprediction in eGFR 45-59 ml/min/1.73 m^2, overprediction over the 5-year time frame for age ≥65 years, and underprediction for eGFR ≥60 ml/min/1.73 m^2 (in which the KFRE was not intended for use). A new equation that incorporates a spline term for eGFR and the competing risk of death was slightly better calibrated in CKD Stage G3a, although inter-cohort variation persisted. In summary, the original KFRE is well calibrated using the CKD-EPI 2021 equation in most populations with eGFR <45 ml/min/1.73 m^2. For populations with eGFR 45-59 ml/min/1.73 m^2, a new equation can be used. Alternative equations should be used in eGFR ≥60 ml/min/1.73 m^2.

Abstract:

Background: The kidney failure risk equation (KFRE) uses age, sex, glomerular filtration rate (GFR), and urine albumin-to-creatinine (ACR) to predict the risk of kidney failure in people with GFR <60 ml/min/1.73 m^2.

Methods: Using 59 cohorts from the CKD Prognosis Consortium, we tested several modifications to the KFRE: 1) using the CKD-EPI 2021 creatinine equation for eGFR; 2) substituting 1-year average ACR for single-measure ACR and 1-year average eGFR in participants with high eGFR variability; and 3) adding 2-year prior eGFR slope and cardiovascular comorbidities. We also assessed calibration of the KFRE in subgroups of eGFR and age before and after accounting for the competing risk of death.

Results: The KFRE remained accurate overall using CKD-EPI 2021 eGFR (median 2-year c-statistic 0.921 (interquartile range [IQR], 0.903, 0.939); median 2-year calibration slope of 1.11 (IQR, 0.87-1.27)). There was no improvement in performance when using 1-year average ACR, using 1-year average eGFR among participants with high variability, adding previous 2-year eGFR slope, or adding cardiovascular variables (heart failure, coronary heart disease, atrial fibrillation, and stroke). The calibration of the KFRE was less good in the low-risk group of eGFR ≥45 ml/min/1.73 m^2, and in older adults using the 5-year time horizon. We developed and tested a new model with a spline term in eGFR and incorporating the competing risk of mortality, which resulted in a median calibration slope of 0.834 (0.453-1.237) from 1.944 (1.129-2.954) in the validation cohorts for eGFR 45-60 ml/min/1.73 m^2 and 0.906 (0.436-1.102) from 0.735 (0.418-0.835) over 5 years in older adults.

Conclusions: The original KFRE is accurate in eGFR <45 ml/min/1.73 m^2 when using the CKD-EPI 2021 equation, but calibration may be improved in low-risk settings with longer time horizons by incorporating competing risk methodology and splines for eGFR. Inclusion of historical averages, eGFR

slopes, or a competing risk design did not meaningfully alter the KFRE performance in most circumstances.

Introduction

Accurate prediction of the risk of adverse kidney outcomes can facilitate shared decision-making, enable risk-based care, and is an important priority for patients and physicians. Tangri et al. previously developed a kidney failure risk equation (KFRE) that uses demographic and laboratory data to predict the progression of chronic kidney disease (CKD) to kidney failure.[1] The risk equation was validated in more than 30 countries and 700,000 participants and has been incorporated into electronic health records, reimbursement criteria, and clinical practice guidelines for CKD management.[2-6]

The 4-variable KFRE incorporates age, sex, glomerular filtration rate (GFR) and albuminuria; however, a new GFR estimating equation (2021 CKD-EPI creatinine equation)[7] is now recommended for use and has not been fully tested in the context of the KFRE. In addition, other identified risk factors may improve the accuracy of the KFRE or allow for implementation in broader populations. For example, the KFRE does not incorporate historical information on kidney function. The use of average recent measurements of kidney function or prior trajectory (slope) in estimated GFR (eGFR) may improve risk prediction. The presence of cardiovascular diseases such as coronary heart disease, atrial fibrillation, heart failure, and stroke – all relatively common in CKD – may affect risk estimates, as may accounting for the competing risk of all-cause mortality.[8] Finally the performance of the KFRE has not been rigorously investigated within subgroups of eGFR and age.

In order to address these questions, we conducted a multinational observational study of patients with CKD enrolled in the CKD Prognosis Consortium. For eGFR, we used the 2021 CKD-EPI creatinine equation. Our overall goal was to evaluate the original KFRE in diverse study populations, including by subgroup of age and eGFR, and improve upon its accuracy wherever possible.

Methods

Study population

Included cohorts were drawn from the CKD Prognosis Consortium, a global consortium of cohorts with data on kidney function and outcomes and at least 1000 participants (www.ckdpc.org).[9] For the present study, cohorts were required to have at least two years of observation with at least two outpatient measures of eGFR prior to the index value, and at least two years of observation thereafter. In total, 59 cohorts had adequate data and agreed to participate. Differences between the study population and that used in the original KFRE validation paper are highlighted in **Appendix 1**. All cohorts were used in analyses of existing equations. For the purpose of equation development, we divided cohorts into development and validation subsets, with development occurring in cohorts able to send data to the Data Coordinating Center as well as a random selection of 50% of the cohorts from OptumLabs® Data Warehouse (OLDW), and validation occurring in the remaining cohorts. Comparisons between newly developed equations and existing equations were performed in validation cohorts. The OLDW is a longitudinal, real-world data asset with de-identified administrative claims and electronic health record (EHR) data.[10]

Procedures

In all cohorts, eGFR was estimated using the CKD-EPI 2021 creatinine equation[7] and serum or plasma creatinine. To mirror inputs available in clinical practice, prior eGFR slope was estimated using all outpatient creatinine measures and linear regression over the previous 1-, 2-, 3- and 5-years at the individual level rather than using mixed models in the population overall. Slopes were categorized as <-3

ml/min/1.73 m^2 per year, between -3 and -1 ml/min/1.73 m^2, -1 to 1 ml/min/1.73 m^2 per year, and >1 ml/min/1.73 m^2 per year. The category -1 to 1 ml/min/1.73 m^2 per year was used as a reference. Other key variables included demographics and urine albumin divided by urine creatinine (ACR). For participants with measured urine protein-to-creatinine ratio (PCR) but not ACR, values were converted to ACR using the unadjusted conversion equation.[11] In analyses including participants with only dipstick proteinuria measurement, urine protein was categorized as negative, trace, +, ++, and ≥+++, and values were converted to ACR.[11] Other variables tested for inclusion were the following cardiovascular conditions: history of heart failure, history of coronary heart disease, atrial fibrillation, and history of stroke (**Appendix 2**)**.**

Outcomes

The primary outcome was kidney failure, defined as the receipt of kidney replacement therapy (dialysis or kidney transplantation). Acute or temporary dialysis was not considered kidney failure. In analyses of competing events, all-cause mortality was simultaneously assessed.

Predicting kidney failure in eGFR <60 ml/min/1.73 m^2: Are new input variables needed?

First, we estimated the time-dependent C-statistic and calibration of the KFRE in all cohorts using GFR estimated using the CKD-EPI 2021 creatinine equation in the population with quantified albuminuria as well as that where only urine dipstick protein was available. As a comparison, we performed the same analyses using the CKD-EPI 2009 equation.[12] Time dependent C-statistics were estimated using the inverse probability of censoring weight; variance and covariance of the C-statistics were estimated by the jackknife method.[13] We quantified calibration by plotting deciles of predicted vs. observed risk

(accounting for the competing risk of death) and estimating the slope ("calibration slope") within each cohort. Perfect calibration has a slope of 1. We summarized model discrimination and calibration using the median and 25^{th}-75^{th} percentile across cohorts. Differences in C-statistics were estimated within each cohort and then summarized using random-effects meta-analysis. Cohort deviations from a calibration slope of 1 by >30% (i.e., calibration slope <0.7 or >1.43) were investigated through meta-regression, considering average age, sex, eGFR, history of coronary heart disease, stroke, heart failure, atrial fibrillation, diabetes, hypertension, and median albuminuria as potential risk factors.

Second, we evaluated whether using a potentially more precise estimate of ACR as an input would improve the performance of the original KFRE. For this analysis, we included all cohorts with participants with an index eGFR <60 ml/min/1.73 m^2, a concomitant value of ACR, and at least one additional measurement of ACR in the year prior. In each cohort, we evaluated the C-statistic and calibration slope of the KFRE using ACR at the index date and compared it to that achieved with the KFRE estimated using the average ACR over the previous year.

Third, we tested whether using average eGFR over the previous year as an input in the original KFRE would improve performance over that estimated solely with eGFR at the index date. For this analysis, we included all cohorts and participants with eGFR <60 ml/min/1.73 m^2 and high variability in eGFR measurements over the previous year (defined as a standard deviation ≥4 ml/min/1.73 m^2, approximately the top quartile for most cohorts), comparing C-statistic and calibration slope of the KFRE estimated using average eGFR to that using index eGFR as above.

Fourth, we evaluated whether adding previous eGFR slope to the KFRE variables would improve performance. To do this, we developed a model in the development cohorts, adding categories of eGFR slopes over 1, 2, 3, and 5-years prior to the index date, fitting a Cox model for kidney failure on age, sex, eGFR, and ACR along with the 4 eGFR slope categories (<-3 ml/min/1.73 m^2 per year, between -3 and -1 ml/min/1.73 m^2, -1 to 1 ml/min/1.73 m^2 per year, and >1 ml/min/1.73 m^2 per year). We then meta-analyzed coefficients to develop a new model of kidney failure. We tested the new model's discrimination and calibration and compared it to that of the original KFRE in the validation cohorts.

Fifth, we evaluated whether adding indicators of heart failure, coronary heart disease, atrial fibrillation, and stroke to the KFRE variables would improve risk estimation. Coefficients were estimated in each development cohort by adding variables to the model with KFRE variables and previous 2-year slope categories. We then meta-analyzed coefficients to develop a new model of kidney failure. We compared the C-statistic and calibration of the new model to that of the original KFRE within each validation cohort, summarizing across cohorts as above.

Evaluating the performance of the KFRE by subgroups of age and eGFR: Is a competing risk model needed?

We evaluated the performance of the original 4-variable KFRE for predicting kidney failure in participants stratified by eGFR (<30 ml/min/1.73 m^2, 30 to <45 ml/min/1.73 m^2, 45 to <60 ml/min/1.73 m^2, and, although the KFRE is not intended for this population, eGFR ≥60 ml/min/1.73 m^2) as well as by subgroups of age (<65 years; ≥65 years).

Because we found that the original KFRE was poorly calibrated in the subgroup with eGFR 45 to <60 ml/min/1.73 m^2, as well as among older adults for the 5-year time horizon, we developed a new model in the development cohorts that incorporated age, sex, log-ACR, and a linear spline term for eGFR with a knot at 45 ml/min/1.73 m^2 and accounted for the competing risk of mortality using the method of Fine and Gray.[14] We then evaluated discrimination and calibration of the model in the validation cohorts, comparing discrimination and calibration with the original KFRE.

All analyses were done in Stata version 14 (StataCorp) using complete case analysis. Statistical significance was determined using a 2-sided test with a threshold P value of <0.05.

Data availability

The CKD Prognosis Consortium has agreed with collaborating cohorts not to share data outside the consortium. Each participating cohort has its own policy for data sharing.

Results

Predicting kidney failure in eGFR <60 ml/min/1.73 m^2: Novel inputs in the original KFRE

Testing the 2021 CKD-EPI equation for eGFR in the original KFRE

In the study population with eGFR <60 ml/min/1.73 m^2 and available ACR, there were 59 cohorts included, with 312,424 participants and 20,728 kidney failure events (**Table 1**). Overall, mean age was 73 years, and mean eGFR was 44 ml/min/1.73 m^2 (**eTable 1**). The KFRE performed well when using the CKD-EPI 2021 equation, with a median cohort 2-year c-statistic of 0.921 (25th percentile to 75th percentile, 0.903-0.939) and calibration slope of 1.111 (0.872-1.272) (**Figure 1A**). The 5-year KFRE also had good

discrimination (C-statistic, 0.898, IQI, 0.883-0.919) and calibration (calibration slope 0.828, 0.736-1.031) (**Figure 1B**). There were eight (out of 58) cohorts with substantial overprediction (>30%) of the 2-year risk and six cohorts with substantial underprediction of the 2-year risk (**eTable 2**), but no cohort-level factor was associated with large miscalibration after accounting for multiple comparisons (**eTable 3**). For the 5-year KFRE, there were three out of 20 cohorts with substantial overprediction of risk and no cohorts with substantial underprediction; no cohort-level factor evaluated was associated with these deviations. Compared to the new CKD-EPI 2021 eGFR equation for the eGFR variable, using the previous CKD-EPI 2009 eGFR resulted in slightly worse discrimination of the 2-year KFRE overall (difference in C-statistic, -0.001, -0.001 to -0.001) and similar calibration (14 out of 58 cohorts with large deviations from a calibration slope closer to 1 using both equations) (**Table 2, columns 1-3**). Results were similar for the 5-year KFRE.

Testing historical eGFR or ACR in the original KFRE

Among participants with at least one additional ACR measure in the year prior to the index date, using average ACR rather than index ACR did not improve the discrimination nor the calibration of the 2- and 5-year KFRE (**Table 2, column 4**). Among participants with highly variable eGFR in the year prior to index date, using average eGFR rather than index eGFR negatively impacted the discrimination and calibration, particularly for the 2-year KFRE (**Table 2, column 5**).

Testing the use of dipstick protein in the original KFRE

In the study population with eGFR <60 ml/min/1.73 m^2 and dipstick protein measures, mean age was 74, mean eGFR was 45 ml/min/1.73 m^2, and 11% had dipstick values of ++ or greater (**eTable 4**). Discrimination and calibration of the KFRE using imputed ACR from these studies showed a median C-statistic of 0.914 (0.902, 0.932) and slight underprediction of 2-year but not 5-year risk (**Table 2, column 6; eFigure 1**).

Predicting kidney failure in eGFR <60 ml/min/1.73 m^2: Introducing new variables into a modified KFRE

Testing whether incorporating previous eGFR slope improves prediction of kidney failure

Model development was conducted among 125,566 participants with available slope measures in the development cohorts. The median 1-, 2-, 3-, and 5-year slope was -3.7 ml/min/1.73 m^2 per year, -2.8 ml/min/1.73 m^2 per year, -2.4 ml/min/1.73 m^2 per year, and -2.1 ml/min/1.73 m^2 per year (**eTable 5, top half**). When categories of slope were incorporated with the variables in the KFRE (age, sex, eGFR, ACR), slope <-3 ml/min/1.73 m^2 per year had small but significant associations with kidney failure, with stronger associations when slopes were observed for at least 2 years (**eTable 6**). On an individual level, previous eGFR slope did not predict future eGFR slope (**eTable 7**).

In the 187,234 participants in the validation cohorts, the new model that incorporated previous 2-year slope had a c-statistic of 0.923 (0.906, 0.943) and 0.896 (0.886, 0.938) for the 2-year and 5-year timeframe, respectively (**Table 3**). Calibration slope was 0.868 (0.647-0.955) for the 2-year model, with no suggestion of improvement over the original KFRE, and 0.725 (0.564-0.822) for the 5-year model.

Testing whether incorporating history of cardiovascular diseases improves prediction of kidney failure

When the additional indicators of heart failure, coronary heart disease, atrial fibrillation, and stroke were included in the model that incorporated 2-year eGFR slope, only heart failure had a statistically significant association with kidney failure (meta-analyzed hazard ratio, 1.21, 95% CI: 1.11-1.33) (**Table 3, last column**). The C-statistic of this model was 0.925 (0.907, 0.945) and 0.899 (0.887, 0.939) for the 2- and 5-year timeframe in the validation cohorts. Calibration slope was 0.883 (0.668-0.960) for the 2-year

model and 0.741 (0.558-0.834) for the 5-year model, suggesting overprediction in the majority of cohorts in both settings.

Predicting kidney failure in subgroups of eGFR and age: the original KFRE

The KFRE, which was developed for use in eGFR <60 ml/min/1.73 m^2, performed poorly in eGFR ≥60 ml/min/1.73 m^2 (**eTable 8-9**). The median c-statistic of 0.750 (25^{th} -75^{th} percentile of cohorts, 0.725-0.776) and there was systematic underprediction of 2- and 5-year risk (**eFigure 2**). The risk of kidney failure was extremely low in these cohorts, with the top quintile of participants having an observed risk of kidney failure of 0.1%-0.4% at 2 years and 0.2%-1.2% at 5 years.

In subgroups of eGFR in participants with eGFR <60 ml/min/1.73 m^2, there was also evidence of underprediction at higher levels of eGFR. Although the KFRE had high discrimination in all subgroups, the median calibration slope was 1.944 (1.129-2.954) for 2-year KFRE and 1.431 (1.256,1.988) for 5-year KFRE in the eGFR 45-59 ml/min/1.73 m^2, 1.301 (0.810, 1.522) and 1.089 (0.864, 1.591) in the eGFR 30-44 ml/min/1.73 m^2, and 1.082 (0.858, 1.232) and 0.821 (0.711, 0.976) in the eGFR <30 ml/min/1.73 m^2 (**Figure 2**).

By subgroups of age in participants with eGFR <60 ml/min/1.73 m^2, the median (25^{th} -75^{th} percentile) cohort calibration slope was 1.154 (0.930, 1.279) for the 2-year KFRE and 0.953 (0.824, 1.063) for the 5-year KFRE in ages younger than 65 years, and 1.072 (0.814, 1.270) and 0.766 (0.634, 0.965) for the 2- and 5-year KFRE in ages 65 and older (**eTable 10**).

Testing whether incorporating the competing risk of mortality and a nonlinear eGFR term improves prediction of kidney failure

To test whether a competing risk model with a linear spline term for eGFR might improve calibration of the KFRE, we constructed a new model in the 31 development cohorts. Coefficients for age, sex, eGFR, and ACR were relatively similar to those of the original KFRE; however, the coefficient for eGFR above 45 ml/min/1.73 m^2 was much weaker than that for eGFR below 45 ml/min/1.73 m^2 (**Table 4**). Overall, the median (25^{th}, 75^{th} percentile) cohort C-statistic for the 2-year and 5-year risk in the validation cohorts was 0.923 (0.913, 0.942) and 0.907 (0.896, 0.942) for the competing risk model, and the calibration slope was 1.002 (0.718, 1.107) and 0.837 (0.610, 1.058).

Within subgroups of eGFR, the competing risk model improved the median cohort calibration in validation cohorts with eGFR 45-59 ml/min/1.73 m^2, where the original KFRE had a median cohort (25^{th}-75^{th} percentile) calibration slope of 1.944 (1.129, 2.954) and 1.431 (1.256, 1.988) for 2-year and 5-year timeframes and the corresponding values for the competing risk model were 0.834 (0.453, 1.237) and 0.957 (0.722, 1.447) (**Table 5**). It also shifted calibration for the 5-year timeframe for older adults, with the original KFRE demonstrating a median calibration slope of 0.735 (0.418, 0.865) and the competing risk model showing a calibration slope of 0.906 (0.436, 1.102).

Discussion

In this multinational collaborative meta-analysis including >3 million individuals across 59 cohorts and more than 30 countries, we comprehensively evaluated the KFRE in the estimation of risk of kidney

failure. The KFRE remained accurate overall in the prediction of kidney failure in populations with eGFR <60 ml/min/1.73 m^2 using the new CKD-EPI 2021 creatinine equation for estimating GFR. Performance of the KFRE was not improved when historical ACR or eGFR averages were substituted for index ACR and eGFR values, respectively, or with the inclusion of new variables, such as previous slopes of eGFR or additional cardiovascular comorbidities. However, when evaluated within subgroups, the KFRE underpredicted kidney failure risk at higher eGFR and, among participants over the age of 65 years, overestimated kidney failure risk over the 5-year timeframes. To address this issue, we developed a competing risk model with a non-linear term for eGFR. The new model demonstrated improved calibration in these subgroups for some cohorts, but there remained heterogeneity in absolute risk across cohorts. Taken together, these findings suggest that the previously developed KFRE is fairly generalizable to a variety of global settings; however, there is variation in local calibration, particularly in subgroups with higher eGFR. We suggest that the KFRE continue to be recommended, aligned with the CKD-EPI 2021 equation, and implemented in health systems and clinical encounters for patients with CKD Stages G3-G5; for more accurate prediction of risk in lower-risk settings, such as eGFR >45 ml/min/1.73 m^2, alternative equations or more common endpoints might be preferable.[15, 16]

The current study investigates several previously suggested enhancements to the KFRE in a large number of global cohorts. A priori, we selected average ACR and average eGFR as alternatives to index ACR and index eGFR, and eGFR slope and cardiovascular as novel inputs. None resulted in meaningful improvements in discrimination or calibration in the validation cohorts. We demonstrate that local variation in absolute risk is substantial, highlighting the limitations of applying results from single cohort studies to an entire patient population. Indeed, previous efforts to develop alternative risk equations have not consistently outperformed the KFRE in external validation.[17] Other studies have identified concerns with the use of the KFRE in patients with CKD Stage G4 at high risk of death, and we previously

developed an alternative equation that models several outcomes simultaneously in this population.[8, 18] Our new competing risk model also helps calibration among older adults over the 5-year timeframe, as well as in eGFR 45-60 ml/min/1.73 m^2. Overall, however, the new model does not improve upon the original KFRE or address the variability in local calibration. These findings emphasize the need for implementation studies of the KFRE in health systems with local calibration where possible, given what appears to be the limited potential for improvement in accuracy in risk prediction for kidney failure in the CKD G3-G5 population with additional variables.

The KFRE was inaccurate in patients with eGFR ≥60 ml/min/1.73 m^2, a population in which it was not intended for use. In these individuals, the absolute risk of kidney failure is very low, even when the risk of other clinically meaningful outcomes such as a loss of 30-50% of kidney function may be high. As such, we recommend alternative equations, such as those to predict a ≥40% decline in eGFR. Given the recent success of SGLT2i in preventing/delaying the loss of eGFR in these populations,[19-21] efforts to identify high-risk individuals early may lead to prevention or forestalling of kidney failure over a patient's lifetime rather than simply a delay in disease progression.

There are some limitations to our findings. First, we focused on validating and developing the equations in patients that had available measurements for eGFR and albuminuria. Given that patients with diabetes and those at higher risk of progression are more likely to have albuminuria measured in routine clinical settings, some cohorts may be biased due to an informative measurement process. However, it is important to note that our study also included cohorts where measurements were part of scheduled study visits, and we did not see any differences in accuracy. Second, while we tested the inclusion of several comorbidity-related variables in the KFRE and did not find meaningful improvement, we were

unable to test biomarkers such as cystatin C, neutrophil gelatinase-associated lipocalin (NGAL), or kidney injury molecule-1 (KIM1). These tests are not available in the vast majority of patients with CKD worldwide, and models incorporating these tests would therefore be difficult to implement. As cystatin C use increases in clinical settings, inclusion or substitution of cystatin C-based eGFR for creatinine-based eGFR in the KFREs should be tested. Third, our sample size and follow-up was reduced by the requirement of a two-year lead-in period during which we could estimate eGFR slope, a novel input that did not improve the KFRE performance. These findings differ from results of surrogate endpoint analyses, which suggest that eGFR slope can serve as a valid surrogate endpoint for end-stage kidney disease. Surrogate endpoint analyses evaluate differences in eGFR at a group, not individual level, and adjust for initial eGFR value, not most recent eGFR value, and they include events that occur during the period of slope observation. Finally, our study provides a new competing risk-based KFRE which may improve calibration in certain cases, but it did not decrease the inter-cohort heterogeneity. Our definition of kidney failure is driven in part by provider-level decisions as to if and when to initiate kidney replacement therapy, which is almost certainly variable across providers, cohorts, and countries. The impact of including kidney failure receiving conservative care in the outcome was not addressable in this study.

Strengths of this study include its large sample size and diversity in geography, ethnicity, and health system design. The evidence for generalizability provided in this study suggests that efforts to incorporate slopes and averages, or alternative statistical approaches to modeling kidney failure risk are less likely to improve model performance, and that efforts should instead be focused on knowledge translation, implementation, and local calibration as needed. We believe that this study provides sufficient and robust evidence that guidelines for CKD staging, diagnosis and management should endorse the implementation of the KFREs in clinical care.

In conclusion, the original KFRE using the 2021 CKD-EPI creatinine equation is accurate in most populations with eGFR <60 ml/min/1.73 m^2, and is not improved by the inclusion of eGFR slopes or cardiovascular variables . The performance of the KFRE is less optimal in low-risk settings, such as CKD G3a. Health systems and researchers particularly interested in the low-risk population may wish to use an alternative equation or outcome for risk stratification. Randomized controlled trials and prospective studies evaluating the effect of implementing the KFREs in clinical practice are needed and ongoing.

Contributors: MEG and YS had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. MEG, NB, SHB, JC, KM, AYMW, and NT were responsible for the study concept and design. MEG, SHB, YS, JC, KM, and AS with the CKD-PC investigators/collaborators listed below were involved in the acquisition of data. MEG, NB, SHB, YS, JC, KM, AYMW, and NT drafted the manuscript. All the authors contributed to the analysis and interpretation of data and to the critical revision of the manuscript for important intellectual content as well as the final decision to submit for publication. MEG, and JC guarantee the integrity of the work. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Competing interests: All authors will complete the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author)

Acknowledgements

CKD-PC investigators/collaborators (study acronyms/abbreviations are listed in **Appendix 2** in the Supplement

ADVANCE: John Chalmers, Mark Woodward; **CanPREDDICT:** Adeera Levin, Ognjenka Djurdjev, Mark Canney, Mila Tang; **CRIC:** Chi-yuan Hsu, Ana Ricardo, Amanda Anderson, Panduranga Rao, Harold Feldman; **Geisinger:** Alex R. Chang, Kevin Ho, Jamie Green, H. Lester Kirchner; **GLOMMS-2:** Simon Sawhney, Corri Black, Angharad Marks, Lynn Robertson; **Go-DARTS:** Samira Bell, Moneeza Siddiqui, Colin Palmer; **Hong Kong CKD:** Angela Yee-Moon Wang, Charmaine Fong, Kitty Mei-Ting Kwong, Henry Wu;

ICES-KDT: Amit X. Garg, Eric McArthur, Manish Sood; **KP Hawaii:** Brian J. Lee; **LCC:** Nigel Brunskill, Rupert Major, David Shepherd, James Medcalf; **Maccabi:** Varda Shalev, Gabriel Chodick; **MASTERPLAN:** Jack Wetzels, Peter Blankestijn, Arjan van Zuilen, Jan van de Brand; **MDRD:** Mark Sarnak, Lesley Inker, Andrew S. Levey; **Mt. Sinai BioMe:** Girish N Nadkarni, Erwin P Bottinger, Ruth JF Loos, Stephen B Ellis; **NephroTest:** Benedicte Stengel, Marie Metzger, Martin Flamant, Pascal Houillier, Jean-Philippe Haymann; **Okinawa:** Kunitoshi Iseki; **OLDW:** Nikita Stempniewicz, John Cuddeback, Elizabeth Ciemins; **PSP-CKD:** Nigel Brunskill, Rupert Major, David Shepherd, James Medcalf; **RCAV:** Csaba P. Kovesdy, Keiichi Sumida; **RENAAL:** Hiddo J.L. Heerspink, Michelle Pena, Dick de Zeeuw; **SCREAM:** Juan J Carrero, Marco Trevisan, Carl Gustaf Elinder, Björn Wettermark; **SHARP:** Colin Baigent, Martin Landray, William G Herrington, Natalie Staplin; **SKS:** Philip Kalra, Rajkumar Chinnadurai, James Tollitt, Darren Green; **SRR-CKD:** Marie Evans, Helena Rydell, Maria Stendahl, Mårten Segelmark; **Sunnybrook:** David Naimark, Navdeep Tangri; **Taiwan MJ:** Chi-Pang Wen, Min-Kuang Tsai, Ta-Wei David Chu; **West of Scotland:** Patrick B. Mark, Jamie P. Traynor, Peter C. Thomson, Colin C. Geddes; **YWSCC:** Jae Il Shin, Jae Won Yang, Jae Seok Kim, Jun Young Lee, Miryung Kim

CKD-PC Steering Committee: Josef Coresh (Chair), Shoshana H Ballew, Alex R. Chang, Ron T Gansevoort, Morgan E. Grams, Orlando Gutierrez, Tsuneo Konta, Anna Köttgen, Andrew S Levey, Kunihiro Matsushita, Kevan Polkinghorne, Elke Schäffner, Mark Woodward, Luxia Zhang

CKD-PC Data Coordinating Center: Shoshana H Ballew (Assistant Project Director), Jingsha Chen (Programmer), Josef Coresh (Co-Principal Investigator), Morgan E Grams (Co-Principal Investigator; Director of Nephrology Initiatives), Kunihiro Matsushita (Director), Yingying Sang (Lead Programmer), Aditya Surapaneni (Programmer), Mark Woodward (Senior Statistician)

Funding

The CKD Prognosis Consortium (CKD-PC) Data Coordinating Center is funded in part by a program grant from the US National Kidney Foundation and the National Institute of Diabetes and Digestive and Kidney Diseases (R01DK100446). A variety of sources have supported enrollment and data collection including laboratory measurements, and follow-up in the collaborating cohorts of the CKD-PC. These funding sources include government agencies such as national institutes of health and medical research councils as well as foundations and industry sponsors listed in **Appendix 3**.

Some of the data reported here have been supplied by the United States Renal Data System. The interpretation and reporting of these data are the responsibility of the authors and in no way should be seen as an official policy or interpretation of the US Government.

References

1. Tangri, N, Stevens, LA, Griffith, J, Tighiouart, H, Djurdjev, O, Naimark, D, Levin, A, Levey, AS: A predictive model for progression of chronic kidney disease to kidney failure. *JAMA,* 305**:** 1553-1559, 2011.
2. Tangri, N, Grams, ME, Levey, AS, Coresh, J, Appel, LJ, Astor, BC, Chodick, G, Collins, AJ, Djurdjev, O, Elley, CR, Evans, M, Garg, AX, Hallan, SI, Inker, LA, Ito, S, Jee, SH, Kovesdy, CP, Kronenberg, F, Heerspink, HJ, Marks, A, Nadkarni, GN, Navaneethan, SD, Nelson, RG, Titze, S, Sarnak, MJ, Stengel, B, Woodward, M, Iseki, K: Multinational Assessment of Accuracy of Equations for Predicting Risk of Kidney Failure: A Meta-analysis. *JAMA,* 315**:** 164-174, 2016.
3. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group: KDIGO 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. *Kidney inter, Suppl,* 3**:** 1-150, 2013.
4. Smekal, MD, Tam-Tham, H, Finlay, J, Donald, M, Thomas, C, Weaver, RG, Quinn, RR, Tam, K, Manns, BJ, Tonelli, M, Bello, A, Tangri, N, Hemmelgarn, BR: Patient and provider experience and perspectives of a risk-based approach to multidisciplinary chronic kidney disease care: a mixed methods study. *BMC Nephrol,* 20**:** 110, 2019.
5. Grill, AK, Brimble, S: Approach to the detection and management of chronic kidney disease: What primary care providers need to know. *Can Fam Physician,* 64**:** 728-735, 2018.
6. Farrington, K, Covic, A, Nistor, I, Aucella, F, Clyne, N, De Vos, L, Findlay, A, Fouque, D, Grodzicki, T, Iyasere, O, Jager, KJ, Joosten, H, Macias, JF, Mooney, A, Nagler, E, Nitsch, D, Taal, M, Tattersall, J, Stryckers, M, van Asselt, D, Van den Noortgate, N, van der Veer, S, van Biesen, W: Clinical Practice Guideline on management of older patients with chronic kidney disease stage 3b or higher (eGFR<45 mL/min/1.73 m2): a summary document from the European Renal Best Practice Group. *Nephrol Dial Transplant,* 32**:** 9-16, 2017.
7. Inker, LA, Eneanya, ND, Coresh, J, Tighiouart, H, Wang, D, Sang, Y, Crews, DC, Doria, A, Estrella, MM, Froissart, M, Grams, ME, Greene, T, Grubb, A, Gudnason, V, Gutierrez, OM, Kalil, R, Karger, AB, Mauer, M, Navis, G, Nelson, RG, Poggio, ED, Rodby, R, Rossing, P, Rule, AD, Selvin, E, Seegmiller, JC, Shlipak, MG, Torres, VE, Yang, W, Ballew, SH, Couture, SJ, Powe, NR, Levey, AS, Chronic Kidney Disease Epidemiology, C: New Creatinine- and Cystatin C-Based Equations to Estimate GFR without Race. *N Engl J Med,* 385**:** 1737-1749, 2021.
8. Ravani, P, Fiocco, M, Liu, P, Quinn, RR, Hemmelgarn, B, James, M, Lam, N, Manns, B, Oliver, MJ, Strippoli, GFM, Tonelli, M: Influence of Mortality on Estimating the Risk of Kidney Failure in People with Stage 4 CKD. *J Am Soc Nephrol,* 30**:** 2219-2227, 2019.
9. Matsushita, K, Ballew, SH, Astor, BC, Jong, PE, Gansevoort, RT, Hemmelgarn, BR, Levey, AS, Levin, A, Wen, CP, Woodward, M, Coresh, J: Cohort Profile: The Chronic Kidney Disease Prognosis Consortium. *Int J Epidemiol,* 42**:** 1660-1668, 2013.
10. OptumLabs: OptumLabs and OptumLabs Data Warehouse (OLDW) Descriptions and Citation. Eden Prairie, MN, n.p., June 2020.
11. Sumida, K, Nadkarni, GN, Grams, ME, Sang, Y, Ballew, SH, Coresh, J, Matsushita, K, Surapaneni, A, Brunskill, N, Chadban, SJ, Chang, AR, Cirillo, M, Daratha, KB, Gansevoort, RT, Garg, AX, Iacoviello, L, Kayama, T, Konta, T, Kovesdy, CP, Lash, J, Lee, BJ, Major, RW, Metzger, M, Miura, K, Naimark, DMJ, Nelson, RG, Sawhney, S, Stempniewicz, N, Tang, M, Townsend, RR, Traynor, JP, Valdivielso, JM, Wetzels, J, Polkinghorne, KR, Heerspink, HJL: Conversion of Urine Protein-Creatinine Ratio or Urine Dipstick Protein to Urine Albumin-Creatinine Ratio for Use in Chronic Kidney Disease Screening and Prognosis : An Individual Participant-Based Meta-analysis. *Ann Intern Med,* 173**:** 426-435, 2020.

12. Levey, AS, Stevens, LA, Schmid, CH, Zhang, YL, Castro, AF, 3rd, Feldman, HI, Kusek, JW, Eggers, P, Van Lente, F, Greene, T, Coresh, J: A new equation to estimate glomerular filtration rate. *Ann Intern Med,* 150**:** 604-612, 2009.
13. Uno, H, Cai, T, Pencina, MJ, D'Agostino, RB, Wei, LJ: On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Stat Med,* 30**:** 1105-1117, 2011.
14. Fine, JP, Gray, RJ: A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association***:** 496-509, 1999.
15. Grams, ME, Brunskill, NJ, Ballew, SH, Sang, Y, Coresh, J, Matsushita, K, Surapaneni, A, Bell, S, Carrero, JJ, Chodick, G, Evans, M, Heerspink, HJL, Inker, LA, Iseki, K, Kalra, PA, Kirchner, HL, Lee, BJ, Levin, A, Major, RW, Medcalf, J, Nadkarni, GN, Naimark, DMJ, Ricardo, AC, Sawhney, S, Sood, MM, Staplin, N, Stempniewicz, N, Stengel, B, Sumida, K, Traynor, JP, van den Brand, J, Wen, CP, Woodward, M, Yang, JW, Wang, AY, Tangri, N, Consortium, CKDP, Chalmers, J, Woodward, M, Hsu, CY, Ricardo, AC, Anderson, A, Rao, P, Feldman, H, Chang, AR, Ho, K, Green, J, Kirchner, HL, Bell, S, Siddiqui, M, Palmer, C, Shalev, V, Chodick, G, Stengel, B, Metzger, M, Flamant, M, Houillier, P, Haymann, JP, Stempniewicz, N, Cuddeback, J, Ciemins, E, Kovesdy, CP, Sumida, K, Carrero, JJ, Trevisan, M, Elinder, CG, Wettermark, B, Kalra, P, Chinnadurai, R, Tollitt, J, Green, D, Coresh, J, Ballew, SH, Chang, AR, Gansevoort, RT, Grams, ME, Gutierrez, O, Konta, T, Kottgen, A, Levey, AS, Matsushita, K, Polkinghorne, K, Schaffner, E, Woodward, M, Zhang, L, Ballew, SH, Chen, J, Coresh, J, Grams, ME, Matsushita, K, Sang, Y, Surapaneni, A, Woodward, M: Development and Validation of Prediction Models of Adverse Kidney Outcomes in the Population With and Without Diabetes. *Diabetes Care,* 45**:** 2055-2063, 2022.
16. Matsushita, K, Kaptoge, S, Hageman, SH, Sang, Y, Ballew, SH, Grams, ME, Surapaneni, A, Sun, L, Arnlov, J, Bozic, M, Brenner, H, Brunskill, NJ, Chang, AR, Chinnadurai, R, Cirillo, M, Correa, A, Ebert, N, Eckardt, KU, Gansevoort, RT, Gutierrez, O, Hadaegh, F, He, J, Hwang, SJ, Jafar, TH, Jassal, SK, Kayama, T, Kovesdy, CP, Landman, GW, Levey, AS, Lloyd-Jones, DM, Major, RW, Miura, K, Muntner, P, Nadkarni, GN, Nowak, C, Ohkubo, T, Pena, MJ, Polkinghorne, KR, Sairenchi, T, Schaeffner, E, Schneider, MP, Shalev, V, Shlipak, MG, Solbu, MD, Stempniewicz, N, Tollitt, J, Valdivielso, JM, van der Leeuw, J, Wang, AYM, Wen, CP, Woodward, M, Yamagishi, K, Yatsuya, H, Zhang, L, Dorresteijn, JA, Di Angelantonio, E, Visseren, FL, Pennells, L, Coresh, J, Chronic Kidney Disease Prognosis, C: Including Measures of Chronic Kidney Disease to Improve Cardiovascular Risk Prediction by SCORE2 and SCORE2-OP. *Eur J Prev Cardiol*, 2022.
17. Collins, GS, Omar, O, Shanyinde, M, Yu, LM: A systematic review finds prediction models for chronic kidney disease were poorly reported and often developed using inappropriate methods. *J Clin Epidemiol,* 66**:** 268-277, 2013.
18. Grams, ME, Sang, Y, Ballew, SH, Carrero, JJ, Djurdjev, O, Heerspink, HJL, Ho, K, Ito, S, Marks, A, Naimark, D, Nash, DM, Navaneethan, SD, Sarnak, M, Stengel, B, Visseren, FLJ, Wang, AY, Kottgen, A, Levey, AS, Woodward, M, Eckardt, KU, Hemmelgarn, B, Coresh, J: Predicting timing of clinical outcomes in patients with chronic kidney disease and severely decreased glomerular filtration rate. *Kidney Int,* 93**:** 1442-1451, 2018.
19. Neal, B, Perkovic, V, Mahaffey, KW, de Zeeuw, D, Fulcher, G, Erondu, N, Shaw, W, Law, G, Desai, M, Matthews, DR, Group, CPC: Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes. *N Engl J Med,* 377**:** 644-657, 2017.
20. Wiviott, SD, Raz, I, Bonaca, MP, Mosenzon, O, Kato, ET, Cahn, A, Silverman, MG, Zelniker, TA, Kuder, JF, Murphy, SA, Bhatt, DL, Leiter, LA, McGuire, DK, Wilding, JPH, Ruff, CT, Gause-Nilsson, IAM, Fredriksson, M, Johansson, PA, Langkilde, AM, Sabatine, MS, Investigators, D-T: Dapagliflozin and Cardiovascular Outcomes in Type 2 Diabetes. *N Engl J Med,* 380**:** 347-357, 2019.

21. Zinman, B, Wanner, C, Lachin, JM, Fitchett, D, Bluhmki, E, Hantel, S, Mattheus, M, Devins, T, Johansen, OE, Woerle, HJ, Broedl, UC, Inzucchi, SE, Investigators, E-RO: Empagliflozin, Cardiovascular Outcomes, and Mortality in Type 2 Diabetes. *N Engl J Med,* 373**:** 2117-2128, 2015.

Table 1. Characteristics of cohorts and participants with any quantitative measure of albuminuria and eGFR <60 ml/min/1.73m^2

	All Cohorts	Cohorts Used in Development	Cohorts Used in Validation
Analytic purpose	Evaluation of new inputs to existing equations	Development of new equations	Validation of new equations & comparison to KFRE
# of cohorts	59	31	24
# of participants	312,424	91,578	142,591
Age	72 (11)	73 (11)	74 (11)
% female	53%	54%	46%
eGFR†	43 (13)	43 (13)	45 (12)
ACR**	27 (10-107)	24 (9-93)	40 (18-147)
Kidney failure events, N	20,197	6,230	7,138
Kidney failure follow-up, y	4 (2)	4 (2)	4 (2)
Slope <-3 ml/min/1.73 m^2/year	N/A	47%	53%
-3 ml/min/1.73 m^2/year ≤ slope < -1 ml/min/1.73 m^2/year	N/A	18%	17%
-1 ml/min/1.73 m^2/year ≤ Slope < 1 ml/min/1.73 m^2/year	N/A	16%	14%
Slope ≥ 1 ml/min/1.73 m^2/year	N/A	19%	16%
HF	N/A	19%	22%
CHD	N/A	36%	32%
Afib£	N/A	16%	26%
Stroke€	N/A	21%	11%

ACR: urine albumin-to-creatinine ratio; Afib: atrial fibrillation; CHD: coronary heart disease; eGFR: estimated glomerular filtration rate; HF: heart failure

* Cohorts were required to have available 2-year eGFR slope, history of CHD, and history of heart failure to be included in the development and validation of the 2-year slope model. When atrial fibrillation or stroke information was not available at a cohort level, the variable was omitted in that cohort's model development or validation.

**Including urine protein-to-creatinine ratio that was converted to ACR.

†Using the 2021 CKD-EPI creatinine equation

£ Afib is not available in 13 development cohorts, 2 validation cohorts

€ Stroke is not available in 3 development cohorts

Table 2. Testing the performance of the original kidney failure risk equation when substituting inputs for eGFR and albuminuria

	New CKD-Epi 2021 Equation for **eGFR**	Old CKD-Epi 2009 Equation for **eGFR**	Average 1-year ACR for **ACR**	Average 1-Year eGFR in patients with high variability for **eGFR**	Dipstick protein for **ACR**
2-year Kidney Failure Risk Equation, All Available Cohorts					
N cohorts***	58	58	45	53	44
N participants	310,094	310,094	111,602	143,822	358,491
N events	11,751	11,751	5,427	3,661	7,537
Median Cohort C-statistic, 25^{th} – 75^{th} percentile*	0.921 (0.903, 0.939)	0.919 (0.898, 0.939)	0.924 (0.877, 0.944)	0.890 (0.863, 0.927)	0.914 (0.902, 0.932)
Median Cohort Calibration Slope, 25^{th} – 75^{th} percentile	1.111 (0.872, 1.272)	1.025 (0.803, 1.177)	1.209 (1.018, 1.388)	1.615 (1.319, 1.953)	1.224 (1.012, 1.474)
Cohorts with Large Deviations in Calibration Slope**	8 overpredict 6 underpredict	10 overpredict 4 underpredict	3 overpredict 9 underpredict	1 overpredict 33 underpredict	3 overpredict 13 underpredict
5-Year Kidney Failure Risk Equation, All Available Cohorts					
N cohorts	20	20	17	19	12
N participants	148,339	148,339	70,511	66,124	100,876
N events	9,155	9,155	5,997	2,880	3,842
Median Cohort C-statistic, 25^{th} – 75^{th} percentile*	0.898 (0.883, 0.919)	0.897 (0.881, 0.918)	0.891 (0.882, 0.909)	0.888 (0.854, 0.920)	0.899 (0.864, 0.917)
Median Cohort Calibration Slope, 25^{th} – 75^{th} percentile	0.828 (0.736, 1.031)	0.779 (0.669, 0.988)	0.822 (0.756, 0.964)	1.005 (0.913, 1.241)	1.005 (0.800, 1.203)
Cohorts with Large Deviations in Calibration Slope**	3 overpredict 0 underpredict	5 overpredict 0 underpredict	3 overpredict 0 underpredict	2 overpredict 3 underpredict	2 overpredict 1 underpredict

*There was no statistically significant improvement in C-statistics compared to the KFRE using CKD-EPI creatinine 2021 for any of the input variables when tested in the same cohorts.

**Large deviations defined as >30% difference, which is a ratio of <0.7 (overprediction) or >1.43 (underprediction).

***The ADVANCE cohort did not have sufficient number of events in the 2 year follow up to be included in analyses.

Table 3. Testing the performance of new input variables in addition to the four variables present in the kidney failure risk equation

Variable	Original KFRE	+2-year slope KFRE	2-year slope +CVD KFRE
Age, 10y	**0.80 (0.75, 0.86)**	**0.75 (0.72, 0.78)**	**0.74 (0.70, 0.77)**
Male	**1.28 (1.04, 1.58)**	**1.30 (1.23, 1.37)**	**1.27 (1.20, 1.36)**
eGFR, 5ml	**0.57 (0.54, 0.61)**	**0.58 (0.55, 0.61)**	**0.58 (0.56, 0.62)**
lnACR, mg/g	**1.57 (1.44, 1.71)**	**1.51 (1.44, 1.58)**	**1.52 (1.46, 1.59)**
Slope <-3 ml/min/1.73 m^2/year		**1.32 (1.15, 1.52)**	**1.27 (1.09, 1.49)**
-3 ml/min/1.73 m^2/year ≤ slope < -1 ml/min/1.73 m^2/year		1.09 (0.95, 1.23)	1.06 (0.92, 1.23)
Slope ≥ 1 ml/min/1.73 m^2/year		1.04 (0.92, 1.19)	1.03 (0.87, 1.21)
HF			**1.21 (1.11, 1.33)**
CHD			1.06 (0.95, 1.18)
Afib			0.97 (0.79, 1.18)
Stroke			1.04 (0.95, 1.14)
2-Year Risk Equations, Validation Cohorts			
N cohorts	24	24	24
N participants	142,586	142,586	142,586
N events	3,693	3,693	3,693
Median Cohort C-statistic, 25^{th} – 75^{th} percentile*	0.925 (0.907, 0.943)	0.923 (0.906, 0.943)	0.925 (0.907, 0.945)
Median Cohort Calibration Slope, 25^{th} – 75^{th} percentile	1.075 (0.817, 1.227)	0.868 (0.647, 0.955)	0.883 (0.668, 0.960)
Cohorts with Large Deviations in Calibration Slope**	4 overpredict 1 underpredict	8 overpredict 0 underpredict	8 overpredict 0 underpredict
5-Year Risk Equations, Validation Cohorts			
N cohorts	8	8	8
N participants	92,087	92,087	92,087
N events	4,609	4,609	4,609
Median Cohort C-statistic, 25^{th} – 75^{th} percentile*	0.903 (0.883, 0.937)	0.896 (0.886, 0.938)	0.899 (0.887, 0.939)
Median Cohort Calibration Slope, 25^{th} – 75^{th} percentile	0.761 (0.578, 0.883)	0.725 (0.564, 0.822)	0.741 (0.558, 0.834)
Cohorts with Large Deviations in Calibration Slope**	2 overpredict 0 underpredict	3 overpredict 0 underpredict	3 overpredict 0 underpredict

*There was no statistically significant improvement in C-statistics compared to the KFRE using CKD-EPI creatinine 2021.

**Large deviations defined as >30% difference, which is a ratio of <0.7 (overprediction) or >1.43 (underprediction).

Table 4. Testing the addition of competing risk methodology and a linear spline term for eGFR with the four variables present in the kidney failure risk equation

Variable	Original KFRE	Competing risk with spline term
Age, 10y	**0.80 (0.75, 0.86)**	**0.70 (0.66, 0.73)**
Male	**1.28 (1.04, 1.58)**	**1.26 (1.19, 1.33)**
eGFR, below 45, per 5ml higher	**0.57 (0.54, 0.61)**	**0.61 (0.58, 0.64)**
eGFR, above 45, per 5 ml higher	**0.57 (0.54, 0.61)**	**0.83 (0.79, 0.86)**
lnACR, mg/g	**1.57 (1.44, 1.71)**	**1.49 (1.42, 1.55)**
2-Year Risk Equations, Validation Cohorts		
N cohorts	24	24
N participants	186,847	186,847
N events	4,735	4,735
Median Cohort C-statistic, 25^{th} – 75^{th} percentile*	0.925 (0.907, 0.943)	0.923 (0.913, 0.942)
Median Cohort Calibration Slope, 25^{th} – 75^{th} percentile	1.075 (0.817, 1.227)	1.002 (0.718, 1.107)
Cohorts with Large Deviations in Calibration Slope**	4 overpredict 1 underpredict	5 overpredict 1 underpredict
5-Year Risk Equations, Validation Cohorts		
N cohorts	8	8
N participants	106,510	106,510
N events	5,096	5,096
Median Cohort C-statistic, 25^{th} – 75^{th} percentile*	0.903 (0.883, 0.937)	0.907 (0.896, 0.942)
Median Cohort Calibration Slope, 25^{th} – 75^{th} percentile	0.761 (0.578, 0.883)	0.837 (0.610, 1.058)
Cohorts with Large Deviations in Calibration Slope**	2 overpredict 0 underpredict	2 overpredict 0 underpredict

*There was no statistically significant improvement in C-statistics compared to the KFRE using CKD-EPI creatinine 2021.

**Large deviations defined as >30% difference, which is a ratio of <0.7 (overprediction) or >1.43 (underprediction).

Table 5. Testing the addition of competing risk methodology and a linear spline term for eGFR with the four variables present in the kidney failure risk equation, within subgroups of eGFR and age

Variable	Original KFRE	Competing risk with spline term
2-Year Risk Equations, Validation Cohorts		
eGFR 45-59		
N cohorts	8	8
N participants	87,924	87,924
N events	312	312
Median Calibration Slope, 25th – 75th percentile Overall	1.944 (1.129, 2.954)	0.834 (0.453, 1.237)
Cohorts with Large Deviations in Calibration Slope*	1 overpredict 5 underpredict	3 overpredict 2 underpredict
eGFR 30-44		
N cohorts	12	12
N participants	46,789	46,789
N events	680	680
Median Calibration Slope, 25th – 75th percentile Overall	1.300 (0.843, 1.471)	0.919 (0.674, 1.061)
Cohorts with Large Deviations in Calibration Slope*	3 overpredict 3 underpredict	3 overpredict 1 underpredict
eGFR <30		
N cohorts	24	24
N participants	22,151	22,151
N events	3,612	3,612
Median Calibration Slope, 25th – 75th percentile Overall	1.054 (0.808, 1.187)	0.975 (0.798, 1.178)
Cohorts with Large Deviations in Calibration Slope*	5 overpredict 1 underpredict	5 overpredict 1 underpredict
Age <65		
N cohorts	22	22
N participants	37,633	37,633
N events	2,194	2,194
Median Calibration Slope, 25th – 75th percentile Overall	1.113 (0.935, 1.234)	0.913 (0.693, 1.029)
Cohorts with Large Deviations in Calibration Slope*	2 overpredict 0 underpredict	6 overpredict 0 underpredict

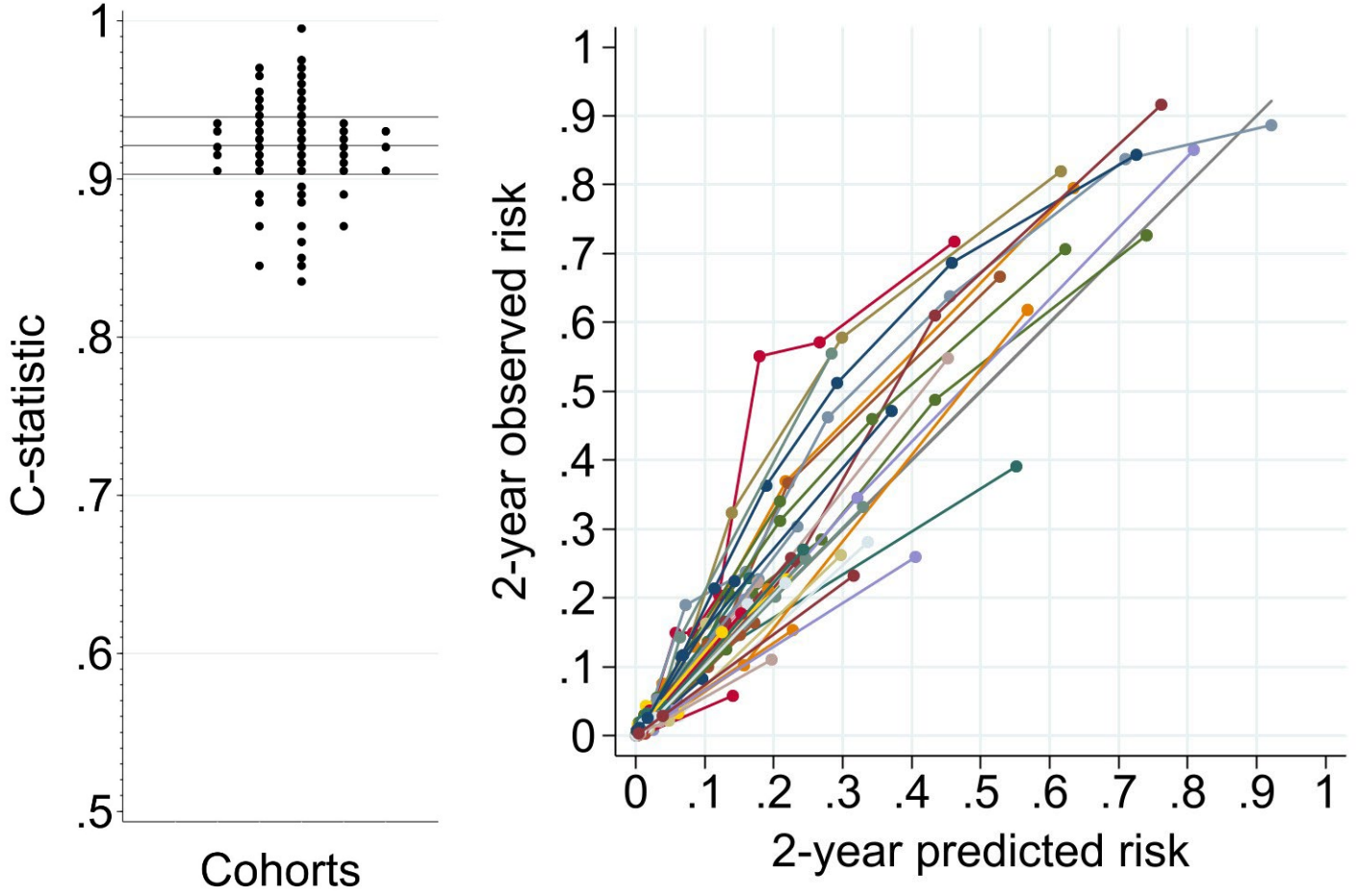
Age 65+		
N cohorts	20	20
N participants	145,704	145,704
N events	2,495	2,495
Median Calibration Slope, 25th – 75th percentile Overall	1.032 (0.814, 1.289)	1.093 (0.723, 1.314)
Cohorts with Large Deviations in Calibration Slope*	4 overpredict 1 underpredict	4 overpredict 2 underpredict
5-Year Risk Equations, Validation Cohorts		
eGFR 45-59		
N cohorts	6	6
N participants	57,828	57,828
N events	499	499
Median Calibration Slope, 25th – 75th percentile Overall	1.431 (1.256, 1.988)	0.957 (0.722, 1.447)
Cohorts with Large Deviations in Calibration Slope*	0 overpredict 3 underpredict	1 overpredict 2 underpredict
eGFR 30-44		
N cohorts	7	7
N participants	31,527	31,527
N events	1,093	1,093
Median Calibration Slope, 25th – 75th percentile Overall	0.864 (0.835, 1.116)	0.995 (0.802, 1.105)
Cohorts with Large Deviations in Calibration Slope*	1 overpredict 1 underpredict	1 overpredict 1 underpredict
eGFR <30		
N cohorts	8	8
N participants	14,976	14,976
N events	3,550	3,550
Median Calibration Slope, 25th – 75th percentile Overall	0.745 (0.586, 0.825)	0.820 (0.622, 0.912)
Cohorts with Large Deviations in Calibration Slope*	2 overpredict 0 underpredict	2 overpredict 0 underpredict
Age <65		
N cohorts	8	8
N participants	14,790	14,790
N events	2,136	2,136

Median Calibration Slope, 25th – 75th percentile Overall	0.912 (0.745, 0.932)	0.863 (0.806, 0.938)
Cohorts with Large Deviations in Calibration Slope*	0 overpredict 0 underpredict	1 overpredict 0 underpredict
Age 65+		
N cohorts	8	8
N participants	91,720	91,720
N events	3,020	3,020
Median Calibration Slope, 25th – 75th percentile Overall	0.735 (0.418, 0.865)	0.906 (0.436, 1.102)
Cohorts with Large Deviations in Calibration Slope*	4 overpredict 0 underpredict	3 overpredict 0 underpredict

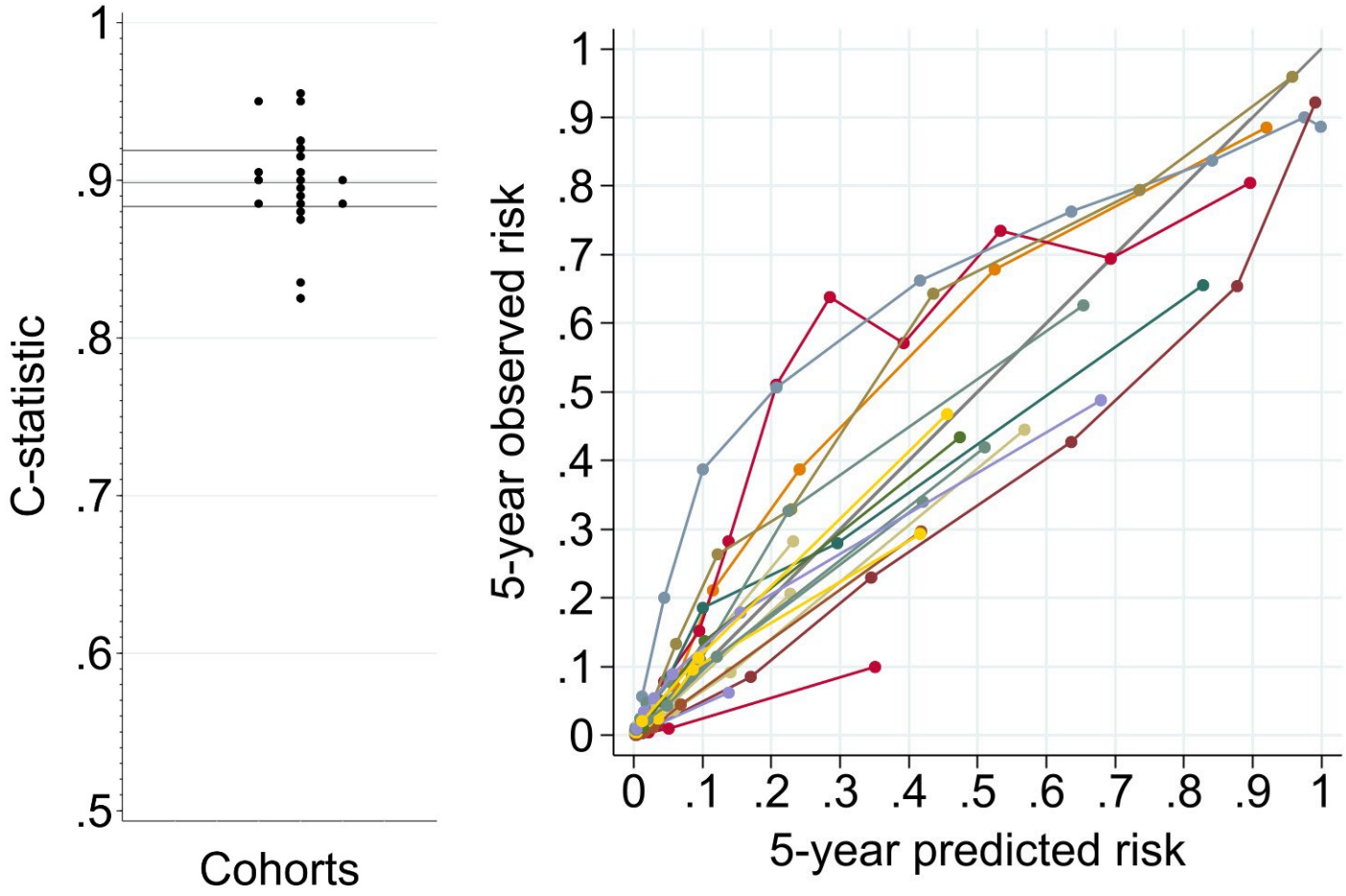
*Large deviations defined as >30% difference, which is a ratio of <0.7 (overprediction) or >1.43 (underprediction).

Figure 1. Display of discrimination (dot plot) and calibration (spaghetti plot) for the 2-year (A) and 5-year (B) 4-variable KFRE in cohorts with eGFR <60 ml/min/1.73 m^2 using the CKD-EPI 2021 creatinine equation for estimating GFR, all cohorts

A)

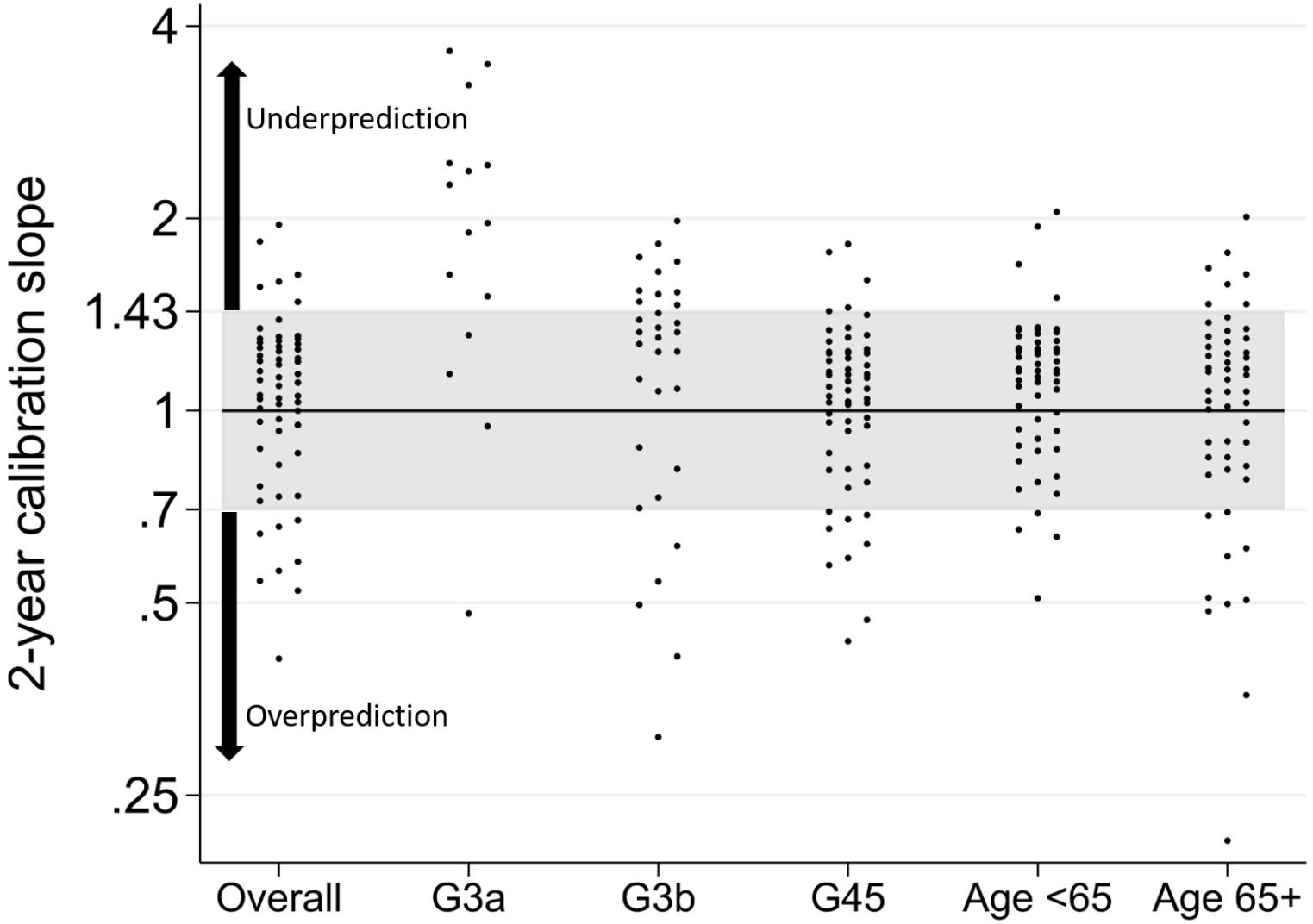


B)

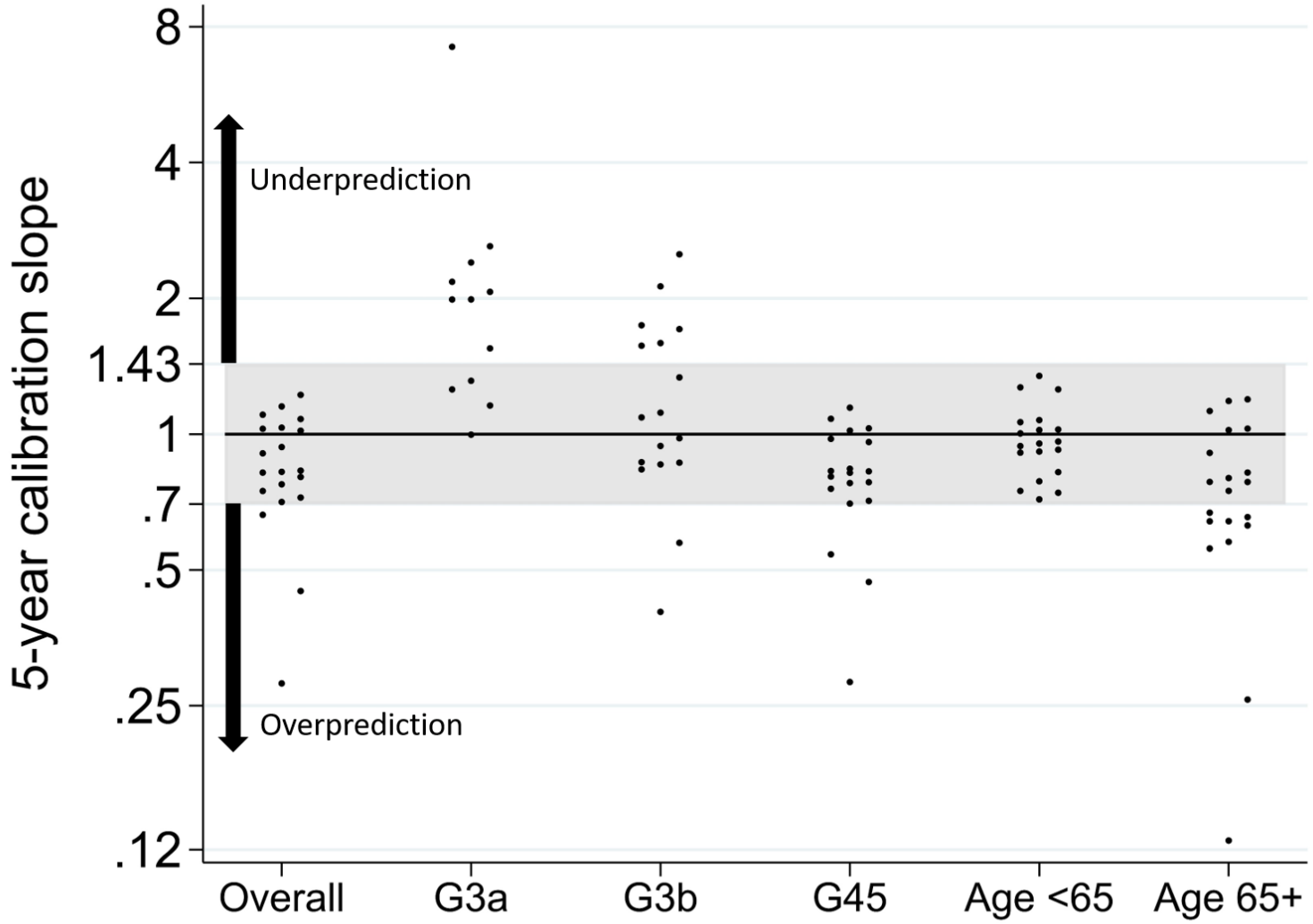


*Each dot on the left-hand graph represents a cohort. Each line on the right represents a cohort, with observed vs. predicted risks plotted by decile of predicted risk.

Figure 2. Calibration slopes of the (A) 2-year and (B) 5-year kidney failure risk equation within subgroups of eGFR and age, all cohorts



(A)


5-year calibration slope
8
4
2
1.43
1
.7
.5
.25
.12
Underprediction
Overprediction
Overall
G3a
G3b
G45
Age <65
Age 65+

(B)

*Each dot represents a cohort. A calibration slope of 1 represents perfect calibration. The gray shaded area represents a calibration slope within 30% of 1 (<0.7 or >1.43).