Importance

Identifying patients at risk of chronic kidney disease (CKD) progression may facilitate more optimal nephrology care. Kidney failure risk equations were previously developed and validated in 2 Canadian cohorts. Validation in other regions and in CKD populations not under the care of a nephrologist is needed.

Objective

To evaluate the accuracy of the risk equations across different geographic regions and patient populations through individual participant data meta-analysis.

Data Sources

Thirty-one cohorts, including 721,357 participants with CKD stages 3 to 5 in more than 30 countries spanning 4 continents, were studied. These cohorts collected data from 1982 through 2014.

Study Selection

Cohorts participating in the CKD Prognosis Consortium with data on end-stage renal disease.

Data Extraction and Synthesis

Data were obtained and statistical analyses were performed between July 2012 and June 2015. Using the risk factors from the original risk equations, cohort-specific hazard ratios were estimated and combined in meta-analysis to form new pooled kidney failure risk equations. Original and pooled kidney failure risk equation performance was compared, and the need for regional calibration factors was assessed.

Main Outcomes and Measures

Kidney failure (treatment by dialysis or kidney transplant).

Results

During a median follow-up of 4 years, 23,829 cases of kidney failure were observed. The original risk equations achieved excellent discrimination (ability to differentiate those who developed kidney failure from those who did not) across all cohorts (overall C statistic, 0.90; 95% CI, 0.89-0.92 at 2 years; C statistic, 0.88; 95% CI, 0.86-0.90 at 5 years); discrimination in subgroups by age, race, and diabetes status was similar. There was no improvement with the pooled equations. Calibration (the difference between observed and predicted risk) was adequate in North American cohorts, but the original risk equations overestimated risk in some non-North American cohorts. Addition of a calibration factor that lowered the baseline risk by 32.9% at 2 years and 16.5% at 5 years improved the calibration in 12 of 15 and 10 of 13 non-North American cohorts at 2 and 5 years, respectively (P = .04 and P = .02).

Conclusions and Relevance

Kidney failure risk equations developed in a Canadian population showed high discrimination and adequate calibration when validated in 31 multinational cohorts. However, in some regions the addition of a calibration factor may be necessary.
Chronic kidney disease (CKD) is increasing in incidence and prevalence worldwide. Rates of progression to kidney failure vary among individuals with CKD and depend on the severity of kidney disease, comorbid conditions, and risk of dying before kidney failure onset. Interventions to slow CKD progression, planning for initiation of dialysis and transplant, and early creation of arteriovenous fistula have been advocated, but these strategies may be expensive and are associated with risks. Treatment would ideally be recommended only for patients at high risk of progression and for whom the benefit exceeds the harm.

Tangri et al previously developed kidney failure risk equations that use demographic and laboratory data to predict progression of CKD to kidney failure. The risk equations were developed in 3449 patients with stages 3 to 5 CKD who were referred for nephrology care in Ontario, Canada, and were validated in referred patients with CKD in British Columbia, Canada. The preferred risk equations (the 4-variable and 8-variable equations) are age-, sex-, and laboratory value-based, thereby enabling automated risk reporting whenever laboratory tests are performed. The 4-variable equation requires age, sex, estimated glomerular filtration rate (eGFR), and urinary albumin to creatinine ratio (ACR), facilitating integration into clinical practice. The kidney failure risk equations are widely used through electronic applications (eg, http://www.qxmd.com/calculate-online/nephrology/kidney-failure-risk-equation), with some initial validation in other countries and health care systems. However, widespread adoption of the risk equations requires validation in additional populations including nonwhite ethnicities, patients not under nephrology care, and cohorts outside North America. Their accuracy in different geographic regions and patient populations is evaluated herein.

Methods

Participating Cohorts

Thirty-one cohorts participating in the Chronic Kidney Disease Prognosis Consortium (CKD-PC) were selected for validation based on data availability. The CKD-PC is a collaborative research group integrating data from more than 50 cohorts spanning 40 countries and involving 2 million individuals. The diverse cohorts include populations across a wide range of baseline risk of kidney failure. For the purpose of this analysis, cohorts were selected to include patients with stages 3 to 5 CKD with an estimated glomerular filtration rate (eGFR) less than 60 mL/min/1.73 m² and an absence of kidney failure at baseline who had follow-up information on kidney failure, defined as treatment by dialysis or a kidney transplant. Data transfer and analysis took place between July 2012 and June 2015. Data in included cohorts were collected from September 1982 through October 2014. This study was approved for use of deidentified data by the institutional review board at the Johns Hopkins Bloomberg School of Public Health, and the need for informed consent was waived.

Measurement of Variables in Cohorts

As in the original kidney failure risk equations, GFR was estimated using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) 2009 creatinine equation. Serum creatinine concentrations were standardized to isotope dilution mass spectrometry-traceable methods where possible. Albuminuria was represented as a log-transformed urine ACR. Alternative measures of urine protein excretion (protein to creatinine ratio, 24-hour urine collection, urinary dipstick) were transformed to the ACR using previously developed equations. When available, baseline values for serum albumin, phosphorous, calcium, and bicarbonate, as well as physical examination measures of weight, systolic and diastolic blood pressure, were derived from each cohort. Age, sex, and ethnicity (black or nonblack), as well as the presence of diabetes and hypertension, were also derived from the individual cohorts, with information on race collected as part of routine clinical care for the health systems and as demographic data for the study cohorts. Diabetes was defined as fasting glucose of at least 126.1 mg/dL (to convert glucose to mmol/L, multiply by 0.0555), nonfasting glucose of at least 200 mg/dL or glycated hemoglobin (Hba1c) of at least 6.5%, use of glucose-lowering drugs, or self-reported diabetes.

Hypertension was defined as a systolic blood pressure of at least 140 mmHg, diastolic blood pressure of at least 90 mmHg, or use of antihypertensive drugs for treatment of hypertension. Potential participants missing any baseline data were excluded from analysis. Information on individual cohorts is provided in eAppendix 1 in the Supplement.

Statistical Analysis

There were 4 kidney failure risk equations developed in the original cohorts: the 3-variable (age, sex, and eGFR), the 4-variable (3-variable + ACR), the 6-variable (4-variable + diabetes and hypertension), and the 8-variable equations (4-variable + calcium, phosphorous, bicarbonate, and albumin). The 4-variable and 8-variable equations demonstrated the best performance in the original cohorts; thus, the focus of this validation effort centered on the 4-variable and 8-variable equations.

Participant-level data were analyzed for each individual cohort. Meta-analysis was performed across studies using a random-effects model. Risk relationships observed in the original cohorts were compared with those seen in the validation cohorts. Cox proportional hazards models were fit using the variables included in each of the original equations within each study, allowing both the regression coefficients and the baseline hazard to vary. All variables were centered (age, 70 years; 56% men, eGFR, 36 mL/min/1.73 m²; ACR, 170 mg/g; phosphorus, 3.9 mg/dL; albumin, 4.0 g/dL; bicarbonate, 25.6 mEq/L; and calcium, 9.4 mg/dL; to convert calcium to mmol/L, multiply by 0.25), as per the original study. The refit coefficients were then pooled across studies using random-effects meta-analysis. Pooled and original coefficients were compared using the z test.
## Table 1. Baseline Characteristics of the Participating Cohorts*

<table>
<thead>
<tr>
<th>Cohort</th>
<th>No. of Participantsb</th>
<th>Follow-up Time, Median (IQI), y</th>
<th>Age, mean (SD), y</th>
<th>No. (%)</th>
<th>eGFR, mean (SD), mL/min/1.73 m²</th>
<th>No. (%) of Participants With Albuminuria, a</th>
<th>No. of Kidney Failure Events a</th>
<th>Kidney Failure Incidence, per 1000 Patient-Years</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>North America</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AASK</td>
<td>898</td>
<td>8 (4-10)</td>
<td>55 (11)</td>
<td>537 (60)</td>
<td>898 (100)</td>
<td>40 (12)</td>
<td>592 (66)</td>
<td>303</td>
</tr>
<tr>
<td>ARIC</td>
<td>722</td>
<td>12 (7-14)</td>
<td>67 (5)</td>
<td>332 (46)</td>
<td>171 (24)</td>
<td>50 (10)</td>
<td>192 (27)</td>
<td>112</td>
</tr>
<tr>
<td>BC CKD</td>
<td>11 331</td>
<td>3 (2-5)</td>
<td>70 (13)</td>
<td>6042 (54)</td>
<td>44 (0.4)</td>
<td>31 (11)</td>
<td>7928 (71)</td>
<td>2091</td>
</tr>
<tr>
<td>CCF ACR</td>
<td>4102</td>
<td>2 (1-4)</td>
<td>71 (11)</td>
<td>1950 (48)</td>
<td>747 (18)</td>
<td>48 (10)</td>
<td>1643 (40)</td>
<td>101</td>
</tr>
<tr>
<td>CCF DIP</td>
<td>12 275</td>
<td>3 (1-4)</td>
<td>72 (13)</td>
<td>5457 (44)</td>
<td>1579 (13)</td>
<td>46 (11)</td>
<td>2835 (23)</td>
<td>300</td>
</tr>
<tr>
<td>CRIC</td>
<td>3099</td>
<td>6 (4-7)</td>
<td>59 (11)</td>
<td>1720 (56)</td>
<td>1315 (42)</td>
<td>40 (11)</td>
<td>1866 (63)</td>
<td>796</td>
</tr>
<tr>
<td>Gesingera</td>
<td>20 720</td>
<td>4 (2-6)</td>
<td>70 (10)</td>
<td>8605 (42)</td>
<td>211 (1)</td>
<td>51 (8)</td>
<td>1961 (44)</td>
<td>453</td>
</tr>
<tr>
<td>ICES-KDTe</td>
<td>100 569</td>
<td>4 (2-6)</td>
<td>73 (11)</td>
<td>46 883 (47)</td>
<td>0</td>
<td>46 (12)</td>
<td>39 611 (39)</td>
<td>3093</td>
</tr>
<tr>
<td>REGARDS</td>
<td>3158</td>
<td>7 (5-8)</td>
<td>72 (9)</td>
<td>1402 (44)</td>
<td>1308 (41)</td>
<td>47 (11)</td>
<td>1079 (36)</td>
<td>240</td>
</tr>
<tr>
<td>Sunnybrook</td>
<td>3173</td>
<td>10 (9-12)</td>
<td>60 (10)</td>
<td>46 578 (45)</td>
<td>166 (11)</td>
<td>45 (11)</td>
<td>478 (32)</td>
<td>100</td>
</tr>
<tr>
<td>RENAAL</td>
<td>1434</td>
<td>3 (2-4)</td>
<td>72 (13)</td>
<td>5457 (44)</td>
<td>1579 (13)</td>
<td>46 (11)</td>
<td>2835 (23)</td>
<td>300</td>
</tr>
<tr>
<td>Okinawa</td>
<td>93</td>
<td>15 (17-17)</td>
<td>69 (10)</td>
<td>419 (25)</td>
<td>0</td>
<td>51 (8)</td>
<td>599 (35)</td>
<td>55</td>
</tr>
<tr>
<td>REGARDS</td>
<td>3158</td>
<td>7 (5-8)</td>
<td>72 (9)</td>
<td>1402 (44)</td>
<td>1308 (41)</td>
<td>47 (11)</td>
<td>1079 (36)</td>
<td>240</td>
</tr>
<tr>
<td>Sunnybrook</td>
<td>3173</td>
<td>10 (9-12)</td>
<td>60 (10)</td>
<td>46 578 (45)</td>
<td>166 (11)</td>
<td>45 (11)</td>
<td>478 (32)</td>
<td>100</td>
</tr>
<tr>
<td>VARK</td>
<td>434 810</td>
<td>4 (3-4)</td>
<td>75 (9)</td>
<td>423 521 (97)</td>
<td>38 893 (9)</td>
<td>47 (11)</td>
<td>14 084 (41)</td>
<td>8836</td>
</tr>
<tr>
<td><strong>Non-North America</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRIB</td>
<td>382</td>
<td>2 (1-5)</td>
<td>65 (13)</td>
<td>1620 (45)</td>
<td>0</td>
<td>42 (14)</td>
<td>970 (63)</td>
<td>525</td>
</tr>
<tr>
<td>Sunnybrook</td>
<td>3098</td>
<td>3 (2-5)</td>
<td>71 (14)</td>
<td>391 (37)</td>
<td>0</td>
<td>49 (39)</td>
<td>313 (30)</td>
<td>55</td>
</tr>
<tr>
<td>REGARDS</td>
<td>3158</td>
<td>7 (5-8)</td>
<td>72 (9)</td>
<td>1402 (44)</td>
<td>1308 (41)</td>
<td>47 (11)</td>
<td>1079 (36)</td>
<td>240</td>
</tr>
<tr>
<td>Sunnybrook</td>
<td>3173</td>
<td>10 (9-12)</td>
<td>60 (10)</td>
<td>46 578 (45)</td>
<td>166 (11)</td>
<td>45 (11)</td>
<td>478 (32)</td>
<td>100</td>
</tr>
<tr>
<td>VARK</td>
<td>434 810</td>
<td>4 (3-4)</td>
<td>75 (9)</td>
<td>423 521 (97)</td>
<td>38 893 (9)</td>
<td>47 (11)</td>
<td>14 084 (41)</td>
<td>8836</td>
</tr>
<tr>
<td>Subtotal</td>
<td>617 604</td>
<td>4 (3-6)</td>
<td>74 (10)</td>
<td>506 751 (82)</td>
<td>50 251 (8)</td>
<td>46 (11)</td>
<td>79 573 (41)</td>
<td>18 926</td>
</tr>
</tbody>
</table>
| **Abbreviations:** AASK, African American Study of Kidney Disease and Hypertension; ACR, urine albumin to creatinine ratio; ARIC, Atherosclerosis Risk in Communities; BC CKD, British Columbia Chronic Kidney Disease; CCF, Cleveland Clinic Foundation; CRIC, Chronic Renal Insufficiency Cohort; DIP, dipstick protein; eGFR, estimated glomerular filtration rate; GCKD, German CKD; GLOMMS, Grampian Laboratory Outcomes, Morbidity and Mortality Studies; HUNT, Nord Trondelag Health Study; ICES-KDT, Institute for Clinical Evaluative Sciences, Provincial Kidney, Dialysis, and Transplantation; IQI, interquartile interval; KEEP, Kidney Early Evaluation Program; KPNW, Kaiser Permanente Northwest; MASTERPLAN, Multifactorial Approach and Superior Treatment Efficacy in Renal Patients With the Aid of a Nurse Practitioner; MDRD, Modification of Diet in Renal Disease; MMKD, Mild to Moderate Kidney Disease; NZDCS, New Zealand Diabetes Cohort Study; REGARDS, Reasons for Geographic and Racial Differences in Stroke Study; RENAAL, Reduction of Endpoints in Non-insulin Dependent Diabetes Mellitus With the Angiotensin II Antagonist Losartan; SRR-CKD, Swedish Renal Registry CKD; VA CKD, Veterans Administration CKD. 

* Representative references for each cohort are provided in eAppendix 3 in the Supplement.
* The number of participants represents the total number with data for the 3-variable equation.
* The proportion of participants with a urine albumin to creatinine ratio of 30 mg/g or higher, urine protein to creatinine ratio of 50 mg/g or higher, or a dipstick protein of 1+ or more. The proportion out of the total number of participants with data for the 4-variable equation is listed in eAppendix 1 in the Supplement.
* Kidney failure is defined as treatment by dialysis or a kidney transplant.
* Denotes cohorts that participated in the validation of the 8-variable equation.

Copyright 2016 American Medical Association. All rights reserved.
A set of pooled risk equations were developed to compare with the original risk equations. Pooled coefficients from the random-effects meta-analysis were combined with a pooled baseline hazard, defined as the average refit baseline hazard weighted by the number of kidney failure events.

Discrimination of the original and pooled kidney failure risk equations was assessed using the Harrell C statistic within each study, which was then meta-analyzed using random-effects models. Performance was also evaluated in predetermined subgroups of black or non-black race, presence or absence of diabetes mellitus, and age older or younger than 65 years. The discrimination of the original and pooled risk equations was compared by assessing the meta-analyzed difference in the C statistic within individual studies. Finally, within each set of original and pooled risk equations, the discrimination of the 4- vs 6- and 4- vs 8-variable risk equations was compared by meta-analyzing the difference in individual study C statistics (6-variable performance is reported in the supplementary materials).

Calibration (the difference between observed and predicted risk) was examined by plotting the observed 2-year and 5-year probability of kidney failure in individual cohorts and comparing it to the predicted risk using the original and pooled risk equations. This was done in 5 risk categories (for 2 years, 0% to <2%, 2% to <6%, 6% to <10%, 10% to <20%, and ≥20%; for 5 years, 0% to <5%, 5% to <15%, 15% to <25%, 25% to <50%, and ≥50%). In the absence of clinical practice guidelines that recommend risk cut-offs or strata for CKD progression, the risk categories were used to adopt the methodology of the original development study and subsequent CKD-PC publications. Calibration varied across cohorts; thus, factors that might explain heterogeneity in baseline risk were investigated by regressing cohort-specific baseline risk on cohort characteristics (eg, region of cohort, mean eGFR, proportion of the cohort with black race, diabetes mellitus, and hypertension). Baseline risk was estimated for each cohort using Cox proportional hazards models, holding the variable coefficients constant and equal to the original risk equations regression coefficients but allowing the intercept to vary. The only cohort characteristic associated with cohort-specific baseline hazard was region of cohort, with higher baseline risk in North American cohorts compared with non-North American cohorts.

Regional variation in baseline risk was addressed through the development of 2 regional calibration factors (North America and non-North America). The regional calibration factors were developed as the ratio of the event-weighted regional mean to the original baseline hazard. A Brier score, the squared difference between the observed vs predicted binary outcomes (observed minus predicted risk), was used to evaluate whether calibration improved with the “regional-calibrated original” risk equations, in each study. The Wilcoxon sign-rank test was used to evaluate the differences in Brier score between original and regional-calibrated original risk equations. An overall Brier score was calculated using event-weighted means. The square root of this overall score was reported as the root-mean-squared error between observed and predicted risk. A P value <.05 was considered statistically significant. All tests were 2-sided. All analyses were performed using Stata MP 13 (StataCorp).

Results

There were 721 357 CKD patients and 23 829 kidney failure events in 31 cohorts with an average follow-up time of 4.2 years (Table 1). A total of 16 cohorts (617 604 patients) were based in North America, and 15 cohorts (103 753 patients) were from Asia, Europe, and Australasia. Missing data varied by cohort (median of 0% for the 4-, 1% for the 6-, and 41% for the 8-variable equations; eAppendix 1 in the Supplement). The amount of missing data was higher in North American cohorts (median missing of 2% for the 4-, 3% for the 6-, and 79% for the 8-variable equations) than in non-North American cohorts (median missing of 0% for the 4-, 1% for the 6-, and 9% for the 8-variable equations). All 31 cohorts had the variables necessary to validate the 4-variable, 29 cohorts had the variables necessary to validate the 6-variable, and 16 cohorts had the variables necessary to validate the 8-variable equations.

The mean age of the study population was 74 years, and the mean baseline eGFR was 46 mL/min/1.73 m². Cohorts ranged from being predominantly men (Veterans Administration CKD, 97%) to majority Asian (Okinawa, 75%). Forty percent of the patients had diabetes, and 84% had hypertension (eTable 1 in the Supplement). Forty percent of the study participants had a baseline urinary ACR of 30 mg/g or greater. The observed incidence of kidney failure ranged from 1.2 events per 1000 person-years in Okinawa to 168.3 events per 1000 person-years.
person-years in the Pima Indian cohort. According to the original 4-variable equation, the proportion of each cohort that had a more than 20% 2-year predicted probability of kidney failure ranged from 0.23% (Okinawa 93 cohort) to 50% (Chronic Renal Impairment in Birmingham cohort).

### Variable Coefficients in the Original and Pooled Kidney Failure Risk Equation

In general, coefficients for the association between different characteristics (eg, age, sex, eGFR, ACR) and the risk of kidney failure were similar in the original and pooled equations (Table 2). Exceptions were eGFR in the 4-variable equations (original vs pooled: HR, 0.57 vs 0.63 per 5 mL/min/1.73 m² higher eGFR) and serum bicarbonate in the 8-variable equations (0.93 vs 0.99 per 1 mEq/L higher serum bicarbonate), both of which were stronger in the original kidney failure risk equation.

### Discrimination

Measures of discrimination for the original 4-variable risk equation were excellent for the 2-year and 5-year predicted probability of kidney failure (Figure 1 and Figure 2). Overall, the 4-variable equation had a pooled C statistic of 0.90 (95% CI, 0.89-0.92) at 2 years, and 0.88 (95% CI, 0.86-0.90) at 5 years. Within individual cohorts, discrimination was also excellent, with a C statistic of at least 0.80 in all but 2 cohorts; the MMKD (Mild to Moderate Kidney Disease) study 2-year C statistic was 0.79 (95% CI, 0.72-0.87) and the MASTERPLAN (Multifactorial Approach and Superior Treatment Efficacy in Renal Patients with the Aid of a Nurse Practitioner) study 5-year C statistic was 0.77 (95% CI, 0.73-0.81). Discrimination for the original 8-variable risk equation was 0.89 (95% CI, 0.88-0.91) at 2 years and 0.86 (95% CI, 0.84-0.87) at 5 years (eFigure 1 in the Supplement). In prespecified subgroups of age, sex, race, region, and diabetes status, discrimination was qualitatively unchanged, with C statistics for the 4-variable equation ranging from 0.90 to 0.92 for 2 years and 0.87 to 0.89 for 5 years (Figure 3). Similar statistics for the 6-variable equation are shown in eFigure 2 in the Supplement.

In general, the pooled 4- and 8-variable equations resulted in similar discrimination to the original equations (eTables 2-7 in the Supplement). There was no significant difference in the overall C statistics of the pooled and the original kidney failure risk equations (eg, 4-variable risk equation over 2 years: −0.0006; 95% CI, −0.0020 to 0.0008). When
2-year risk in all 31 cohorts was assessed individually, the pooled 4-variable equation performed significantly better than the original 4-variable equation in 5 cohorts, and in 5 cohorts it performed significantly worse (P < 0.05 for each comparison).

Discrimination of the 8-variable risk equation was slightly better than the 4-variable equation in cohorts that had the necessary components for both equations (eTables 8 and 9 in the Supplement). This was true using either the original or the pooled risk equations and in nearly all subgroupsof interest.

**Calibration**

Plots of the observed vs predicted risk demonstrated differences in calibration, with suboptimal performance in some of the non-North American cohorts (eFigures 3-6 for the North American cohorts; eFigures 7-10 for the non-North American cohorts, both in the Supplement). Baseline risk varied by region, with higher levels in North America compared with non-North America using the 4-variable equation (Figure 4). There was slightly less variation in baseline risk by region using the 8-variable equation (eFigure 11 in the Supplement).

In non-North American studies, use of a regional calibration factor that lowered the baseline risk by 32.9% at 2 years and 16.5% at 5 years decreased the root mean-squared distance of the observed to expected risk from 0.237 to 0.228 at 2 years and 0.299 to 0.287 at 5 years for the 4-variable equation and improved performance in 12 out of 15 studies at 2 years (P = .04) and 10 out of 13 studies at 5 years (P = .02) (eTable 10 in the Supplement). In contrast, use of a regional calibration factor in North American cohorts, the region where the kidney failure risk equations were developed, did not significantly improve performance. For example, the root mean-squared distance of the observed to expected risk at 2 years only minimally changed from 0.152 to 0.151 with the addition of the calibration factor and increased from 0.264 to 0.272 at 5 years for the 4-variable equation. eAppendix 2 in the Supplement shows all equations.

**Discussion**

In this collaborative meta-analysis involving 721,357 patients across 31 cohorts and over 30 countries, the kidney failure risk equations accurately predict the 2-year and 5-year probability of kidney failure in patients with CKD with a wide range of variation in age, sex, race, and in the presence or absence of diabetes.

The original equations reported by Tangri et al demonstrated excellent discrimination and appropriate calibration in the majority of the North American cohorts, and the addition of a recalibration factor optimized performance in non-North American populations. The 4-variable equation (age, sex, eGFR, and albuminuria) can be easily implemented in electronic
medical records and laboratory information systems. The use of this equation is consistent with the Kidney Disease Improving Global Outcomes (KDIGO) guideline, which recommends integration of risk prediction in the evaluation and management of CKD and is in agreement with a strong body of evidence demonstrating the importance of eGFR and albuminuria in predicting progression.13,15,22-35

Previous investigators developed alternative risk prediction models for progression of CKD to kidney failure,36 but most have not been externally validated. The kidney failure risk equations developed by Tangri et al16 were externally validated in a cohort of Canadian CKD patients referred for nephrology care, but their accuracy in nonreferred patients and regions outside Canada remained unknown. Thus, current clinical practice guidelines recommended the use of risk equations for predicting progression and planning dialysis access, but with appropriate caution regarding their external validity.37 The current validation study addresses these concerns, and more widespread clinical assessment can now be recommended. Similar to previous work, an incremental improvement in performance was observed with an 8-variable risk equation, which additionally includes serum albumin, phosphate, bicarbonate, and calcium levels over the 4-variable equation. The magnitude of improvement was smaller than in the original study but may be meaningful for patients for whom data for both equations is readily available. These findings suggest that the 4-variable risk equation might be adopted more widely, but the 8-variable equation should be made available if the additional variables are obtained and increased precision is desired.

The risk associations observed in the pooled validation sample were similar to those in the original kidney failure risk equation. In particular, younger age, male sex, lower eGFR, and higher albuminuria were associated with a higher risk of kidney failure defined by treatment with dialysis or transplant. The finding of lower risk of kidney failure with older age is consistent with the previous literature38 and is likely due to a combination of factors: 1) the same disease process (eg, diabetic nephropathy in a patient with type 1 diabetes with age of diagnosis at 15 years) is more likely to be indolent, if the patient has an eGFR of 30 mL/min/1.73 m² at age 75 years (60 years of exposure) vs age 45 years (30 years of exposure); 2) as patients age, they are more likely to die from a competing cause (malignancy, cardiovascular disease) than reach kidney failure; and 3) older patients may be more likely to choose conservative care for kidney failure rather than treatment with dialysis or transplant, our primary outcome.39 It is important to note that in the original development of the risk equation,4 competing risk models were evaluated and a threshold of eGFR of less than 10 mL/min/1.73 m² was tested as a secondary out-

---

**Table: Accuracy of Equations for Predicting Risk of Kidney Failure**

<table>
<thead>
<tr>
<th>Patients</th>
<th>2-Year predicted probability of kidney failure</th>
<th>5-Year predicted probability of kidney failure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No. of Patients</strong></td>
<td><strong>C Statistic (95% CI)</strong></td>
<td><strong>No. of Patients</strong></td>
</tr>
<tr>
<td><strong>Diabetes</strong></td>
<td></td>
<td><strong>Diabetes</strong></td>
</tr>
<tr>
<td>Yes 140947</td>
<td>0.897 (0.869-0.924)</td>
<td>Yes 105343</td>
</tr>
<tr>
<td>No 126336</td>
<td>0.918 (0.898-0.937)</td>
<td>No 118543</td>
</tr>
<tr>
<td><strong>Age, y</strong></td>
<td></td>
<td><strong>Age, y</strong></td>
</tr>
<tr>
<td>≥65 196626</td>
<td>0.903 (0.879-0.926)</td>
<td>≥65 162600</td>
</tr>
<tr>
<td>&lt;65 70847</td>
<td>0.898 (0.874-0.922)</td>
<td>&lt;65 61276</td>
</tr>
<tr>
<td><strong>Black</strong></td>
<td></td>
<td><strong>Black</strong></td>
</tr>
<tr>
<td>Yes 13125</td>
<td>0.910 (0.882-0.928)</td>
<td>Yes 8997</td>
</tr>
<tr>
<td>No 236463</td>
<td>0.896 (0.879-0.914)</td>
<td>No 199073</td>
</tr>
<tr>
<td><strong>Diabetes</strong></td>
<td></td>
<td><strong>Diabetes</strong></td>
</tr>
<tr>
<td>Yes 17770</td>
<td>0.890 (0.874-0.906)</td>
<td>Yes 16040</td>
</tr>
<tr>
<td>No 22223</td>
<td>0.902 (0.889-0.915)</td>
<td>No 22223</td>
</tr>
<tr>
<td><strong>Age, y</strong></td>
<td></td>
<td><strong>Age, y</strong></td>
</tr>
<tr>
<td>≥65 3311</td>
<td>0.892 (0.874-0.909)</td>
<td>≥65 2991</td>
</tr>
<tr>
<td>&lt;65 31420</td>
<td>0.898 (0.886-0.911)</td>
<td>&lt;65 30307</td>
</tr>
<tr>
<td><strong>Black</strong></td>
<td></td>
<td><strong>Black</strong></td>
</tr>
<tr>
<td>Yes 19979</td>
<td>0.892 (0.874-0.909)</td>
<td>Yes 24417</td>
</tr>
<tr>
<td>No 236463</td>
<td>0.896 (0.879-0.914)</td>
<td>No 23527</td>
</tr>
<tr>
<td><strong>Age, y</strong></td>
<td></td>
<td><strong>Age, y</strong></td>
</tr>
<tr>
<td>≥65 24336</td>
<td>0.905 (0.882-0.927)</td>
<td>≥65 14513</td>
</tr>
<tr>
<td>&lt;65 15678</td>
<td>0.891 (0.876-0.905)</td>
<td>&lt;65 15477</td>
</tr>
</tbody>
</table>
come; no differences in the performance of the kidney failure risk equations was observed.

Although recalibration was not needed in most North American cohorts, adding a regional calibration factor in non-North American cohorts improved calibration and would allow the risk equations to be used clinically in countries with different levels of baseline risk. This is similar to the Framingham Heart Study equation, which is used for estimating cardiovascular risk and has been recalibrated for use in multiple different populations. Different baseline risk between cohorts and regions may reflect different cohort inclusion criteria or treatment preferences for kidney failure rather than physiological differences in disease progression because risk relationships between the risk factors and kidney failure were fairly uniform across settings. Further studies examining additional causes of heterogeneity in higher- vs lower-risk populations are needed.

There are important clinical and research implications to this study’s findings. Clinicians can now use the 4- or 8-variable kidney failure risk equations, with the recalibration factor where applicable, that can inform patient-clinician communication and treatment decisions regarding the absolute risk of kidney failure, rather than the CKD stage alone. Decisions regarding access placement or transplant referral could be made once kidney failure risk thresholds are exceeded. Some kidney failure risk thresholds have been proposed on the basis of physician surveys and decision analyses (>3% or 5% risk for 5 years for nephrology referral, >20% or 40% risk over 2 years for vascular access planning), and should be evaluated further in cluster randomized trials or time series analyses. Routine reporting and clinical implementation is already under way in several centers, and its effect on patient care and health services is being studied. From a research perspective, the risk equation can be used to estimate event rates and statistical power for kidney failure outcomes in clinical trials and may be useful in selecting higher-risk patients for trial inclusion and identifying risk-treatment interactions.

This study has limitations. First, the risk equation does not assess kidney failure risk in patients with CKD stages G1 (GFR ≥90 mL/min/1.73m²) and G2 (GFR 60–89 mL/min/1.73m²). Previous studies have shown that patients with stages G1 to G2 and high levels of albuminuria should be considered as high risk. Second, due to the variables required, validation of the 8-variable equation was not possible in all cohorts. Therefore, nested comparisons between equations are limited to a subset. In some cohorts, proteinuria was converted to albuminuria. Although no meaningful differences in discrimination were observed in these populations, it is possible that risk relationships may differ slightly for the 2 measures. Furthermore, even with the inclusion of more than 700,000 participants in more than 30 countries, there was not significant representation from countries where there is limited access to renal replacement therapy. Validation in these countries with a combined end point of treated and untreated kidney failure should be performed. Third, there were missing data, particularly in the North American health systems. Missing data reduce the generalizability of our findings to North American health sys-
tems. However, results reflect data available in clinical health systems. Fourth, the risk equations provide the risk of kidney failure over 2 and 5 years. These time frames are important for decisions regarding nephrology referral, dialysis access planning, and preemptive transplant (i.e., kidney transplant prior to receiving dialysis), but they do not capture longer-term risk of kidney failure, which may affect other clinical decisions such as lifestyle modification.43 Fifth, the kidney failure risk equation incorporates routinely collected laboratory data. Accuracy of risk predictions may be enhanced in specific subpopulations by novel biomarkers of CKD; however, the incremental gain in predictive accuracy may not be justified by the cost of these newer assays for the entire CKD population.44 Sixth, there is no evidence that using the equation will improve outcomes. Well-designed pragmatic randomized trials are needed to definitively establish the evidence for efficacy.

Strengths of this study include the large patient population and accompanying diversity in age, sex, race, and etiology of kidney disease. In North America, the 4-variable original risk equation appears generalizable and highly accurate in most cohorts and can be easily implemented across multiple health care systems. Elsewhere, the recalibrated risk equation appears more accurate and can also be integrated into health care platforms. Partnerships with mobile technology developers and health care systems may ensure that knowledge translation occurs without long delays, which are common in biomedical research.

Conclusions

Kidney failure risk equations developed in a Canadian population showed high discrimination and adequate calibration when validated in 31 multinational cohorts. However, in some regions the addition of a calibration factor may be necessary.

**ARTICLE INFORMATION**

**Correction:** The wrong version of the article that was published has been corrected on January 25, 2016.

**Author Affiliations:** Department of Medicine, Seven Oaks General Hospital, University of Manitoba, Winnipeg, Canada (Tangri); Department of Community Health Sciences, Seven Oaks General Hospital, University of Manitoba, Winnipeg, Canada (Tangri); Johns Hopkins Medical Institutions, Baltimore, Maryland (Grams, Appel); Division of Nephrology at Tufts Medical Center, Boston, Massachusetts (Levey, Inker, Sarnak); Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland (Coresh, Appel, Woodward); Department of Medicine, University of Wisconsin School of Medicine and Public Health, Madison (Astor); Department of Population Health Sciences, University of Wisconsin School of Medicine and Public Health, Madison (Astor); Medical Division, Maccabi Healthcare Services, and Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel (Chodick); Chronic Disease Research Group, Minneapolis Medical Research Foundation, Minneapolis, Minnesota (Collins); Department of Medicine, University of Minnesota, Minneapolis (Collins); Department of Measurement & Reporting, Provincial Health Service Authority, Vancouver, British Columbia, Canada (Djurdjev); Department of General Practice and Primary Health Care, School of Population Health, University of Auckland, Auckland, New Zealand (Elley); Division of Renal Medicine, CLINTEC, Karolinska Institutet, Stockholm, Sweden (Evans); Departments of Medicine and Epidemiology and Biostatistics, Western University, and Institute for Clinical Evaluative Sciences, Ontario, Canada (Garg); Department of Cancer Research and Molecular Medicine, Faculty of Medicine, Norwegian University of Science and Technology, Trondheim (Hallan); Division of Nephrology, Department of Medicine, St Olav University Hospital, Trondheim, Norway (Hallan); Division of Nephrology, Endocrinology and Vascular Medicine, Department of Medicine, Tohoku University School of Medicine, Sendai, Japan (Ito); Department of Epidemiology and Health Promotion, Institute for Health Promotion, Graduate School of Public Health, Yonsei University, Seoul, Korea (Jee); Memphis Veterans Affairs Medical Center, Memphis, Tennessee (Kovesedy); University of Tennessee Health Science Center, Memphis, Tennessee (Kovesedy); Department of Medical Genetics, Molecular and Clinical Pharmacology, Division of Genetic Epidemiology, Medical University of Innsbruck, Innsbruck, Austria (Kronenberg); Department of Clinical Pharmacy and Pharmacology, University of Groningen, University Medical Center Groningen, the Netherlands (Heerspink); Division of Applied Health Sciences, University of Aberdeen, and NHS Grampian, Forrestherhill, Aberdeen, Scotland (Marks); Division of Nephrology, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, New York (Nadkarni); Division of Nephropathy and Hypertension, Cleveland Clinic, Cleveland, Ohio (Navaneethan); National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Phoenix, Arizona (Nelson); Department of Nephropathy and Hypertension, University of Erlangen-Nürnberg, Erlangen, Germany (Titze); CESP, INSERM, Villejuif, France (Stengel); Université Paris-Saclay, Université Paris-Sud, USVS, Villejuif, France (Stengel); The George Institute for Global Health, Nuffield Department of Population Health, University of Oxford, Oxford, England (Woodward); The George Institute for Global Health, University of Sydney, Sydney, Australia (Woodward); Dialysis Unit, University Hospital of the Ryukyu, Nishihara, Okinawa, Japan (Iseki).

**Author Contributions:** Dr Coresh had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Tangri, Grams, Levey, Coresh, Sarnak, Stengel, Woodward, Iseki. Acquisition, analysis, or interpretation of data: All authors. Drafting of the manuscript: Tangri, Grams, Levey, Coresh, Woodward, Iseki. Critical revision of the manuscript for important intellectual content: All authors. Statistical analysis: Grams, Coresh, Woodward. Obtained funding: Coresh. Administrative, technical, or material support: Grams, Coresh. Study supervision: Tangri, Grams, Levey, Coresh, Sarnak, Stengel, Woodward, Iseki.

**CKD-PC investigators/collaborators:** (Study acronym/abbreviations are listed in eAppendix 3 in the Supplement.)

**AASK:** Jackson T. Wright, Jr, MD, PhD, Case Western Reserve University; Lawrence J. Appel, MD, MPH, Johns Hopkins University; Tom Greene, PhD, University of Utah; and Brad C. Astor, PhD, MPH, University of Wisconsin. **ARIC:** Josef Coresh, MD, PhD, Kunihiro Matsushita, MD, PhD, Morgan E. Grams, MD, PhD, and Yingying Sang, MS, Johns Hopkins University. **British Columbia CKD:** Aedera Levin, MD, FRCP, BC Provincial Renal Agency and University of British Columbia and Ogijnjka Djurdjev, MSc, BC Provincial Renal Agency and Provincial Health Services Authority. **CCF:** Sankar D Navaneethan, MD, MPH, Joseph V. Nally, Jr, MD, and Jesse D. Schold, PhD, Cleveland Clinic. **CRIB:** David C. Wheeler, MD, FRCP, University College London; Jonathan Emerson, PhD, University of Oxford; Jonathan N. Townsend, MD, FRCP, Queen Elizabeth Hospital Birmingham; and Martin J. Landray, PhD, FRCP, University of Oxford. **CRIC:** Lawrence J. Appel, MD, MPH, Johns Hopkins University; Harold Feldman, MD, MScE, University of Pennsylvania; Chi-yuan Hsu, MD, MSc, University of California-San Francisco. **GCKD:** Kai-Uwe Eckardt, MD, University of Erlangen-Nürnberg; Anna Kottgen, MD, MPH, University of Freiburg; Florian Kronenberg, MD, Medical University of Innsbruck; and Stephanie Titze, MD, MSc, University of Erlangen-Nürnberg. **Geisinger:** Jamie Green, MD, MS, H. Lester Kirchner, PhD, and Robert Perkins, MD, and Alex R. Chang, MD, MS, Geisinger Medical Center, Danville, Pennsylvania. **GLOMMS-1 Study:** Corri Black, MBChB, MRCP, MSc, MFPH, FFPH, University of Aberdeen; Angharad Marks, MBChB, MRCP, MSc, PhD, University of Aberdeen; Nick Fluck, BSc, MBBCh, DPhil, FRCP, NHS Grampian, Aberdeen; Dr Laura Clark, MBChB, MRCP, NHS Grampian, Aberdeen; and Gordon J. Prescott, BSc, MSc, PhD, CS, University of Aberdeen. **Gonryo:** Sadayoshi Ito, MD, PhD, Tohoku University School of Medicine, Japan; Mariko Miyazaki, MD, Tohoku University School of Medicine; Masaki Nakayama, MD, Fukushima Medical University and Tohoku University School of Medicine; and Gen Yamada, MD, Tohoku University School of Medicine. **HUNT:** Stein Hallan, MD, PhD, Norwegian University of
Accuracy of Equations for Predicting Risk of Kidney Failure

REGARDS: David H. Warnock, MD, University of Alabama at Birmingham; Paul Muntner, PhD, University of Alabama at Birmingham; Suzanne Judd, PhD, University of Alabama at Birmingham; William McClellan, MD, MPH, Emory University; and Orlando Gutierrez, MD, MMSc, University of Alabama at Birmingham. Kaiser Permanente NW: David H. Warnock, MD, PhD, Mihal L. Thorp, DO, MPH, and Eric S. Johnson, PhD, Kaiser Permanente Northwest. KEEP: Allan J. Collins, MD, University of Minnesota and Minneapolis Medical Research Foundation, Shu-Cheng Chen, MS, MPH, and Suying Li, PhD, Minneapolis Medical Research Foundation. Maccabi: Gabriel Chodick, PhD, Maccabi Healthcare Services and Tel Aviv University, Israel; Isha Shalev, MD, Maccabi Healthcare Services and Tel Aviv University; Nachman Ash, MD, Maccabi Healthcare Services; and Bracha Shainberg, PhD, Maccabi Healthcare Services. MASTERPLAN: Jack F. M. Wetzels, MD, PhD, Raboud University Medical Centre, Nijmegen; Peter J. Blankstedt, MD, PhD, University Medical Center Utrecht; and Arjan D. van Zulien, MD, PhD, University Medical Center Utrecht. MDRD: Mark J. Sarnak, MD, MS, Andrew S. Levey, MD, Lesley A. Inker, MD, and Vandana Menon, MD, PhD, Tufts Medical Center and St Elizabeth’s Medical Center, Boston, and John Hopkins Medical Institutions, Columbia, MD. Medical University of Innsbruck: Barbara Kollerits, PhD, MPH, Medical University of Innsbruck; and Eberhard Ritz, MD, Ruprecht-Karls-University. Mt Sinai BioMe: Girish N. Nadkarni, MD, MPH, Erwin P. Bottinger, MD, Stephen B. Ellis, and Rajiv Nadkarni, Icahn School of Medicine at Mount Sinai. NephefTest: Marc Froissart, MD, PhD, Inserrm UI018, Benedicte Stengel, MD, PhD, Inserrm UI018 and University of Paris Sud-11; Marie Metzger, MD, Inserrm UI018 and University of Paris Sud-11; Jean-Philippe Haymann, MD, PhD, Sorbonne University; and Jean-Jacques Lebreton, MD, PhD, Assistance Publique-Hôpitaux de Paris. Hôpitaux Universitaires de Paris: Pascal Houillier, MD, PhD, Assistance Publique-Hôpitaux de Paris, Paris Descartes University, France; and Martin Flamant, MD, PhD, Assistance Publique-Hôpitaux de Paris. NZDCS: C. Raina Elley, MBChB, PhD, Timothy Kenealy, MBChB, PhD, and Simon A. Moyses, MSc, University of Auckland; and John F. Collins, MBChB, and Paul L. Drury, MA, MB, BCHIR, Auckland District Health Board. Okinawa 83/93: Kunishiti Iseki, MD, University Hospital of the Ryukyu. RENAAAL: Hidde J. L. Lambers Heerspink, PharmD, PhD, University of Groningen; Barry E. Brenner, MD, PhD, Brigham and Women’s Hospital and Harvard School of Medicine, and Dick de Zeeuw, MD, PhD, University of Groningen. Severance: Sun Ha Jee, PhD, Heejin Kimm, MD, PhD, and Yejin Mok, MPH, Yonsei University. SAPPHIRE: MD: Marie Evans, MD, PhD, Karolinska Institutet and Swedish Renal Registry and Maria Stendahl, MD, PhD, Swedish Renal Registry and Hospital of Ryhov. Sunnybrook: Navdeep Tangri, MD, PhD, FRPC, University of Manitoba; Maneesud Srid, MD, University of Toronto; David Naarmann, MD, MSc, FRPC, University of Toronto. VA: Dr. Claudia P. Kovesdy, MD, Memphis Veterans Affairs Medical Center and University of Tennessee Health Science Center; and Kamyar Kalantar-Zadeh, MD, MPH, PhD, University of California Irvine Medical Center. CKD-PC Steering Committee: Josef Coresh, MD, PhD (chair), Johns Hopkins University; Ron T. Gangevoort, MD, University Medical Center Groningen; Morgan E. Grams, MD, PhD, Johns Hopkins University; Paul E. de Jong, MD, PhD, University Medical Center Groningen; Kunitoshi Iseki, MD, University Hospital of the Ryukyu; Andrew S. Levey, MD, Tufts Medical Center; Kunihiro Matsushita, MD, PhD, Johns Hopkins University; Mark J. Sarnak, MD, MS, Tufts Medical Center; Benedicte Stengel, MD, PhD, Inserrm UI018 and University of Paris Sud; David Warnock, MD, University of Alabama at Birmingham; and Mark Woodward, PhD, George Institute and Johns Hopkins University.

CKD-PC Data Coordinating Center (Johns Hopkins University): Shoshana H. Balew, PhD (assistant project director), Josef Coresh, MD, PhD, (principal investigator), Morgan E. Grams, PhD (director of nephrology initiatives); Kunihiro Matsushita, MD, PhD (director); Yingying Sang, MS (lead programmer); and Mark Woodward, PhD (senior statistician) (also at George Institute).

Conflict of Interest Disclosures: All authors have completed and submitted the ICMEJ Form for Disclosure of Potential Conflicts of Interest. Dr Tangri reports receiving honoraria from Takeda Inc and serving on the advisory boards Takeda Inc and Otsuka. Dr Levey reports having a patent application pending for GFR estimation and serves as the principal investigator for CKD-EP and served as the principle coordinator for Chronic Kidney Disease Outcomes Quality Initiative (KDOQI) Guideline 2012. Dr Coresh reports receiving grants from the National Kidney Foundation and Kidney Disease Improving Global Outcomes and having a patent pending for GFR estimation. Dr Chodick reports receiving payments from AstraZeneca, GlaxoSmithKline, and Pfizer and has a patent pending for GFR estimation. Dr D’Agostino reports receiving grants from the National Heart, Lung and Blood Institute and having a patent pending for GFR estimation. Dr Tangri reports receiving grant from the National Health and Medical Research Council as the principal investigator for CKD-EPI, and US National Institute of Diabetes and Digestive and Kidney Diseases. Dr Marks reports receiving personal fees for funding on the adjudication committees of Boehringer-Ingelheim and AbbVie.

Funding/Support: The CKD-PC Data Coordinating Center is funded in part by a program grant from the US National Kidney Foundation (NIK funding sources include AbbVie, Amgen, and Merck) and grant K08DK092287 from the National Institute of Diabetes and Digestive and Kidney Diseases. 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 4 in the Supplement.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: Some of the data reported herein have been supplied by the US Renal Data System (USRDS). 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