Independent External Validation and Comparison of Death and Kidney Replacement Therapy Prediction Models in Advanced CKD

Susan J. Thanabalasingam, Eduard A. Iliescu, Patrick A. Norman, Andrew G. Day, Ayub Akbari, Gregory L. Hundemer, and Christine A. White

Rationale & Objective: The Kidney Failure Risk Equation (KFRE) is widely used to predict the risk of kidney replacement therapy (KRT) initiation in chronic kidney disease (CKD) stages G3-G5. The new Grams calculator developed for advanced CKD (stage G4+) predicts KRT initiation, cardiovascular events, and death by uniquely incorporating the competing risk of death. We aimed to validate this tool in a stage G4+ cohort for death and KRT.

Study Design: Retrospective cohort study.

Setting & Participants: 442 patients with CKD stage G4+ (mean ± SD age, 73 ± 12 years; mean ± SD estimated glomerular filtration rate, 20 ± 6.2 mL/min/1.73 m^2) who visited the multidisciplinary CKD clinic at Kingston Health Sciences Center in Ontario, Canada.

Outcomes & Analytical Approach: Discrimination and calibration were examined for the outcome of death using the 2- and 4-year Grams scores. The 2- and 5-year KFRE and 2- and 4-year Grams scores were compared in terms of discrimination and calibration for KRT.

Results: There were 91, 161, and 206 death events and 90, 145, and 159 KRT events in our cohort at 2, 4, and 5 years, respectively. The Grams model demonstrated modest discrimination for death at 4 years (area under the curve [AUC] 0.70; 95% CI, 0.65-0.75) and performed worse at 2 years (AUC, 0.63; 95% CI, 0.57-0.70). It only overpredicted death by approximately 10% across most of the predicted range. Both models had similar discrimination for KRT at 2 years (KFRE AUC, 0.83; 95% CI, 0.78-0.88 and Grams AUC, 0.8; 95% CI, 0.76-0.87), 4 years (Grams AUC, 0.82; 95% CI, 0.77-0.86), and 5 years (KFRE AUC, 0.81; 95% CI, 0.76-0.85). There was excellent calibration for KRT using the 2-year KFRE and Grams values for predicted risk thresholds of ≤15% and using the 5-year KFRE and 4-year Grams values for predicted risk thresholds of ≤20%. At higher risk ranges, KFRE overpredicts and Grams underpredicts the KRT risk.

Limitations: This is a single-center study with a primarily White cohort limited by smaller sample sizes at the higher ranges of the predicted risks, particularly for the Grams calculator.

Conclusions: The Grams model provides moderately accurate death predictions, and consideration should be given to its incorporation into patient education and advanced care planning. Both the Grams and KFRE models remain clinically useful for determining KRT risks in advanced CKD.

Visual Abstract included

Complete author and article information provided before references.

Correspondence to C.A. White (cw38@queensu.ca)

Kidney Med. 4(4):100440. Available online 7 March 2022.

doi: 10.1016/j.xkme.2022.100440

Chronic kidney disease (CKD) is associated with increased risks of cardiovascular events, mortality, and progressive decline in the glomerular filtration rate (GFR) leading to kidney failure.[1,2] However, outcomes among patients with CKD are quite variable. The rate of progression and other outcomes vary between individuals depending on clinical and demographic factors such as the etiology of CKD, degree of reduction of GFR and proteinuria, comorbid conditions, age, access to health care, ethnicity, and many others.[3] Consequently, interest in integrating risk prediction tools into clinical practice has grown so that patients at low risk may be spared from undue anxiety and costly medical testing, whereas patients at high risk can access timely, appropriate interventions such as nephrology referral, enrollment in multidisciplinary clinic, and kidney replacement therapy (KRT) preparation.[4-6]

Although numerous risk prediction tools have been developed in CKD cohorts worldwide, most have not been adequately validated or widely integrated into clinical practice.[7] The Kidney Failure Risk Equation (KFRE) was developed by Tangri et al[8] in a Canadian population (CKD stages 3-5) and are the most validated KRT prediction models.[8-15] The 4-variable iteration uses age, sex, estimated GFR (eGFR), and urine albumin-creatinine ratio (ACR) to predict the risks of CKD stages G3-G5 progressing to KRT initiation using 2-year and 5-year KFRE scores (KFRE-2 and KFRE-5, respectively).[8] Validation involving 31 multinational cohorts with a mean baseline eGFR of 46 mL/min/1.73 m^2 showed the KFRE to have high discriminatory ability and adequate calibration.[16]

In 2018, Grams et al[9] developed a tool (www.ckdpcrisk.org/lowgfrevents/) specifically for the CKD stage G4+ (eGFR < 30 mL/min/1.73 m^2) population to predict both the probabilities and order of KRT, cardiovascular disease (CVD) events, and death at 2 and 4 years. This tool incorporates 9 clinical and demographic characteristics previously shown to be important outcome predictors in CKD stage G4+.[17] This novel tool, which uniquely considered the competing endpoint of death in its development, has not been extensively externally validated. A recent study of 2 cohorts—the Swedish Renal Registry and the European Quality Study—examined its

PLAIN-LANGUAGE SUMMARY

Patients who have chronic kidney disease are at increased risk of both death and worsening kidney function requiring dialysis or transplant. Numerous risk prediction models have been developed to help estimate these risks and guide clinical decision-making. In our study, we examined the performance of the Kidney Failure Risk Equation—the most widely used prediction model in this setting—and that of the Grams calculator, which is a newer model that looks at the risks of both kidney failure and death. We found that the Grams model had moderate performance for predicting death in our advanced chronic kidney disease cohort. We also confirm that both models perform similarly well at predicting the risk of kidney replacement therapy initiation.

predictive ability for KRT but not for the other important outcome of death.[14] Our study aimed to evaluate this new tool's discrimination and calibration for death and KRT in a nephrology-referred CKD stage G4+ cohort and to compare its performance to that of the KFRE.

METHODS

Study Design and Cohort

We performed a retrospective cohort study of adults (≥18 years) who attended the multidisciplinary CKD clinic at least once at Kingston Health Sciences Center in Ontario, Canada, in 2013, and who had a documented albumin-creatinine ratio and an eGFR < 30 mL/min/1.73 m^2 as per the CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration) creatinine-based equation.[18] In 2013, the Kingston Health Sciences Center's CKD clinic was the only such clinic in the region that covered a catchment population of 500,000 residents in southeastern Ontario. Electronic medical records were used to extract demographic and clinical data, including systolic blood pressure, at the clinic visit associated with the blood work, smoking status, and diabetes and CVD history (defined as at least 1 instance of myocardial infarction, stroke, or heart failure). Ethics approval was obtained from the Health Sciences and Affiliated Teaching Hospitals Research Ethics Board (6004492). Informed consent was waived because all study participant information was deidentified. Our study follows Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) prediction model reporting.

Outcomes

Outcomes included observed incidences of death by 2 and 4 years from the 2013 index clinic visit and KRT initiation at 2, 4, and 5 years. Death was ascertained using records from the office of the Ontario Registrar General, which provided death dates, and *International Classification of Diseases* codes for the cause of death for all deceased patients.[19] Cardiovascular events were not included in this analysis, as we could not be confident that all events would be captured because patients could have been admitted to other regional hospitals. The Kingston Health Sciences Center, however, offers the only KRT program in our health region; therefore, KRT data were believed to be very robust, with little potential for loss to follow-up. The absence of CVD data precluded an analysis of timing of events relative to each other. Because of limited numbers of deaths after KRT initiation, we limited the analysis to total deaths independent of the timing of KRT.

Predictors

The predictors used in the study included the 2-year and 4-year Grams calculator risk scores (Grams-2 and Grams-4, respectively) for any KRT and any death and the 4-variable KFRE-2 and KFRE-5 scores for KRT at the time of the index clinic visit in 2013. The developers of the Grams model provided us with the model's code and coefficients, and these were used to generate our Grams model predictions. Predicted risks of any death and any KRT, independent of the order in which they occurred, were calculated by summing all the probabilities, including these outcomes. For example, the probability of any death comprised the sum of the probabilities of death after KRT, after CVD, after KRT and CVD, and death only.

Analyses

There were no missing data for the entire cohort with respect to all baseline data, with the exception of patient ethnicity. Where ethnicity data were unavailable (n = 3), Grams predictions were generated imputing White race. For the 7 patients whose albumin-creatinine ratio levels were below detection thresholds, we imputed the lowest model input (10 mg/g).

Discrimination

Discrimination describes a model's ability to separate those who experience the event from those who do not. This was assessed using receiver operating characteristic (ROC) curves, area under the curve (AUC) point estimates, and 95% confidence intervals (CIs). ROC curves and AUCs were generated to assess the discriminative performance of the 2- and 4-year Grams models at predicting death. AUCs were also calculated for the outcome of KRT for the 2- and 5-year KFRE models and the 2- and 4-year Grams models. Comparisons of the AUCs between the 2 models at 2 years for the observed risk of KRT were performed using the Delong-Delong-Clarke Pearson method.[20] We interpreted AUCs as follows: <0.7 as poor, 0.7-0.79 as fair, 0.8-0.89 as good and >0.9 as excellent.[21] Additionally, Harrell's C indexes (HCIs) were calculated to examine the discriminatory capacity, taking the time-to-event into account.

Calibration

Calibration examines the agreement between predictions and observed outcomes. This was captured in calibration

Table 1. Baseline Patient Characteristics (N = 442)

Characteristic	n (%)
Mean age (SD), y	73 (12)
Male, n (%)	246 (56)
Race, n (%)	
White	427 (97)
Black	2 (0.45)
Indigenous	6 (1.4)
Asian	4 (0.90)
Other	3 (0.68)
Median urine albumin-creatinine (Q1, Q3), mg/g	262 (46, 1,062)
<30, n (%)	220 (50)
30-300, n (%)	177 (40)
≥300, n (%)	43 (10)
Mean eGFR-EPI (SD), mL/min/1.73 m^2	20 (6.2)
CKD etiology	
Diabetic nephropathy, n (%)	204 (46)
Hypertensive nephrosclerosis, n (%)	120 (27)
Glomerulonephritis, n (%)	18 (4.1)
Polycystic kidney disease, n (%)	12 (2.7)
Other, n (%)	88 (20)
Comorbid conditions	
SBP, mean (SD)	131 (18)
Diabetes mellitus, n (%)	248 (56)
Cardiovascular disease history, n (%)	242 (55)
Smoking history, n (%)	73 (17)

Abbreviations: CKD, chronic kidney disease; eGFR, estimated glomerular filtration rate; Q, quartile; SBP, systolic blood pressure; SD, standard deviation.

plots that graphically compared predicted and observed risks of death at 2 and 4 years for the Grams model and compared 2- and 4-year versus 5-year KRT risks for the Grams and KFRE models. The plots were generated using locally weighted scatterplot smoothing across predicted risk values, with rug plots along the x-axis to visualize the distribution of the predicted risk data, as well as the average observed probability of an outcome among groups of patients defined by deciles of predicted risk. Brier scores, ranging from 0 (most accurate) to 1 (least accurate), measured prediction accuracy.

RESULTS

Patient Characteristics

There were 444 patients with an eGFR < 30 mL/min/1.73 m^2 and a confirmed CKD clinic visit. Two patients were excluded for inaccessible outcome data because they were known to have moved away from our health region, leaving a final cohort of 442 patients (Table 1). The mean ± SD age was 73 ± 12 years, 56% of the cohort were men, and 97% were White. The mean eGFR was 20 ± 6.2 mL/min/1.73 m^2, and the median (quartile 1, quartile 3) urine ACR was 262 mg/g (46 mg/g, 1,062 mg/g). The most common CKD etiology was diabetic kidney disease (46%), followed by hypertensive nephrosclerosis (27%). The mean systolic blood pressure was 131 ± 18 mm Hg. The majority of the cohort had a recorded history of diabetes mellitus (56%) and CVD (55%). A minority of patients were active smokers (17%).

Table 2. Predicted and Observed KRT and Death at 2, 4, and 5 Years

Predicted or Observed KRT and Death	Median (IQR) or n (%)
2-year Grams, %	
Probability of death, median (IQR)	23 (15-34)
Probability of KRT, median (IQR)	15 (2-25)
4-year Grams, %	
Probability of death, median (IQR)	45 (31-60)
Probability of KRT, median (IQR)	25 (14-40)
KFRE	
2-year KFRE, %, median (IQR)	15 (5-35)
5-year KFRE, %, median (IQR)	40 (16-74)
Observed death, n (%)	
2 years	91 (20)
4 years	161 (36)
5 years	206 (47)
Observed KRT, n (%)	
2 years	87 (20)
4 years	143 (32)
5 years	158 (36)

Abbreviations: IQR, interquartile range; KFRE, Kidney Failure Risk Equation; KRT, kidney replacement therapy.

Baseline Predictors

The median 2- and 4-year Grams predictions for death were 23% (interquartile range [IQR], 15%-34%) and 45% (IQR, 31%-60%), respectively. The median 2- and 4-year Grams predictions for KRT were 15% (IQR, 8%-25%) and 25% (IQR, 14%-40%), respectively. The median 2- and 5-year KFRE scores were 15% (IQR, 5%-35%) and 40% (IQR, 16%-74%), respectively (Table 2).

KRT and Mortality Events

There were 91, 161, and 206 death events and 90, 145, and 159 KRT events at 2, 4, and 5 years, respectively (Table 2; Fig 1). There were 78, 131, and 155 deaths before KRT, whereas deaths after KRT initiation numbered 13, 30, and 51, at 2, 4, and 5 years, respectively. Causes of death were unavailable for 19 of 206 (9.2%) observed deaths. Of the remaining 187 deaths, the leading cause was CVD (27%; Fig 1).

Model Performance

Discrimination

Figure 2 depicts ROC curves and corresponding AUCs, HCIs, and Brier scores for the Grams and KFRE models at 2, 4, and 5 years. The Grams-4 had fair discrimination for the outcome of death (AUC, 0.70; 95% CI, 0.65-0.75), but the Grams-2 had poor discrimination (AUC, 0.63; 95% CI, 0.57-0.70). HCIs accounting for the time-to-event discriminative capacity were similarly poor at the 2-year (HCI, 0.62; 95% CI, 0.56-0.67) and 4-year (HCI, 0.61; 95% CI, 0.55-0.67) time points for the outcome of death.

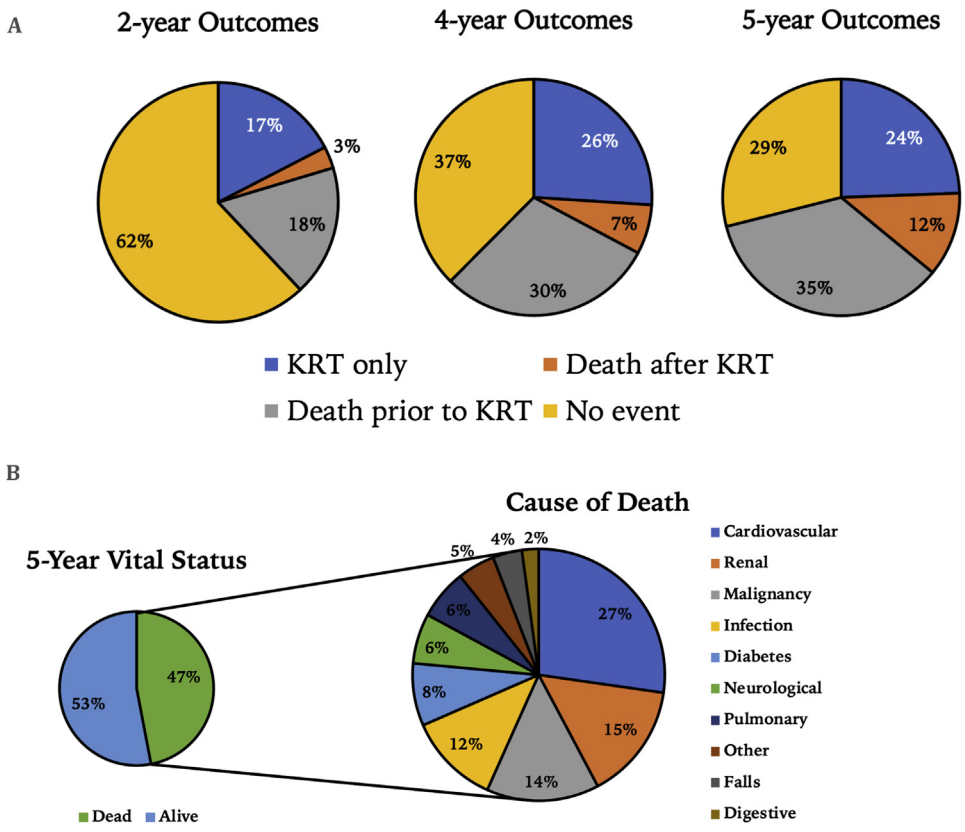


Figure 1. (A) Observed incidences of KRT, death, and no KRT or death event at the 2-, 4-, and 5-year time points; and (B) etiology of death at the 5-year time point. Abbreviations: KRT, kidney replacement therapy.

Both KFRE and Grams models demonstrated good discrimination for KRT initiation at 2, 4, and 5 years, with AUCs from 0.81 to 0.83 (Grams-2 AUC, 0.81 [95% CI, 0.76-0.87]; KFRE-2 AUC, 0.83 [95% CI, 0.78-0.88]; Grams-4 AUC, 0.82 [95% CI, 0.77-0.86]; KFRE-5 AUC, 0.81 [95% CI, 0.76-0.85]). A comparison of the KFRE-2 and Grams-2 KRT AUCs demonstrated no significant difference ($P = 0.21$). The KFRE-2 model demonstrated good time-to-event discrimination for the outcome of KRT initiation (HCI, 0.80; 95% CI, 0.75-0.85), whereas the Grams-2 (HCI, 0.78; 95% CI, 0.73-0.83), Grams-4 (HCI, 0.77; 95% CI, 0.73-0.80), and KFRE-5 (HCI, 0.77; 95% CI, 0.73-0.81) models were all fair.

Calibration

The Grams model is best calibrated for death at a <20% predicted risk and overpredicts death overall for a predicted risk of >20% at 2 and 4 years (Fig 3). The Grams-4 predicted death risk is approximately 10% higher than that observed in the ranges of >20%. Figure 4 demonstrates calibration plots comparing predicted versus observed KRT risks at 2, 4, and 5 years. There was excellent calibration using the KFRE-2 and Grams-2 for predicted KRT risk thresholds of <15% and using the KFRE-5 and Grams-4 for predicting KRT risk thresholds of <20%. At higher ranges of predicted risks, all models perform poorly, but the KFRE-2 and KFRE-5 appear to outperform the Grams-2 and Grams-4. The predicted KRT risk was higher than the observed risk using KFRE, whereas the observed risk was consistently higher using Grams beyond a threshold of 20% predicted risk at 2 years and a threshold of 30% at 4 and 5 years.

Brier scores assessing the predictive accuracy of death using the Grams model were 0.13 and 0.16 at 2 and 4 years, respectively. Brier scores assessing accuracy for predicting KRT initiation were all <0.20 at 2, 4, and 5 years (Fig 2).

DISCUSSION

This study in an advanced CKD cohort with high prevalences of diabetes and CVD is the first to externally validate the Grams prediction model for both death and KRT. We found that the Grams model demonstrated fair discrimination in identifying patients who experience death over a longer time frame (Grams-4 AUC, 0.70; 95% CI, 0.65-0.75) and performed inadequately over a shorter time frame (Grams-2 AUC, 0.63; 95% CI, 0.57-0.70). This modest discrimination may, in part, be because of the homogeneity of our cohort, which comprised patients with similar covariate values. Additionally, modest performance for the outcome of death overall is unsurprising,

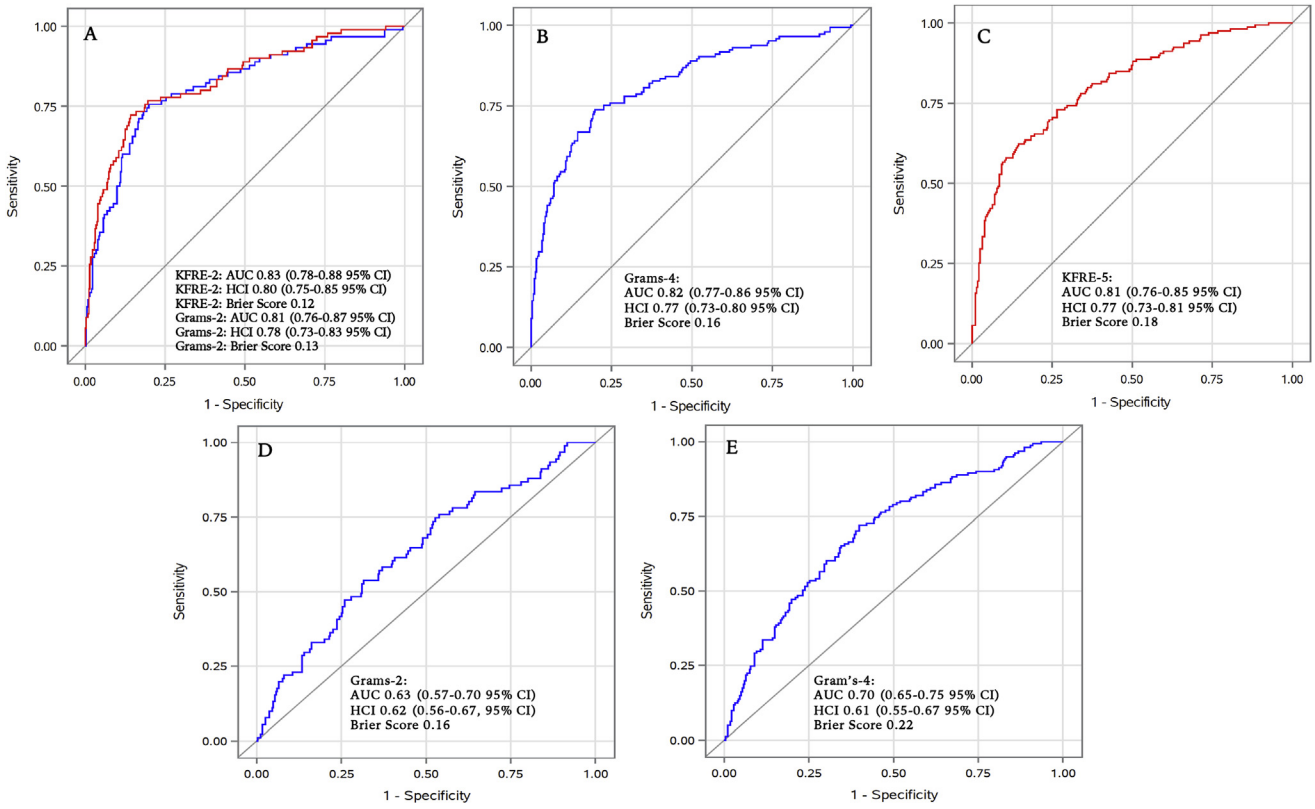


Figure 2. ROC curves for any KRT and any death. The Grams models are represented in blue, whereas the KFRE models are represented in red. The AUCs and HCIs are included in each pane and are interpreted as follows: <0.7 is considered poor, 0.7-0.79 is considered fair, 0.8-0.89 is considered good, and >0.9 is considered excellent. Brier scores are also included in each pane. These range from 0 (most accurate) to 1 (least accurate). (A) Grams-2 and KFRE-2 KRT ROC curves. (B) Grams-4 KRT ROC curve. (C) KFRE-5 KRT ROC curve. (D) Grams-2 death ROC curve. (E) Grams-4 death ROC curve. Abbreviations: AUC, area under the curve; CI, confidence interval; HCI, Harrell's C index; KFRE, kidney failure risk equation; KRT, kidney replacement therapy; ROC, receiver operating characteristic.

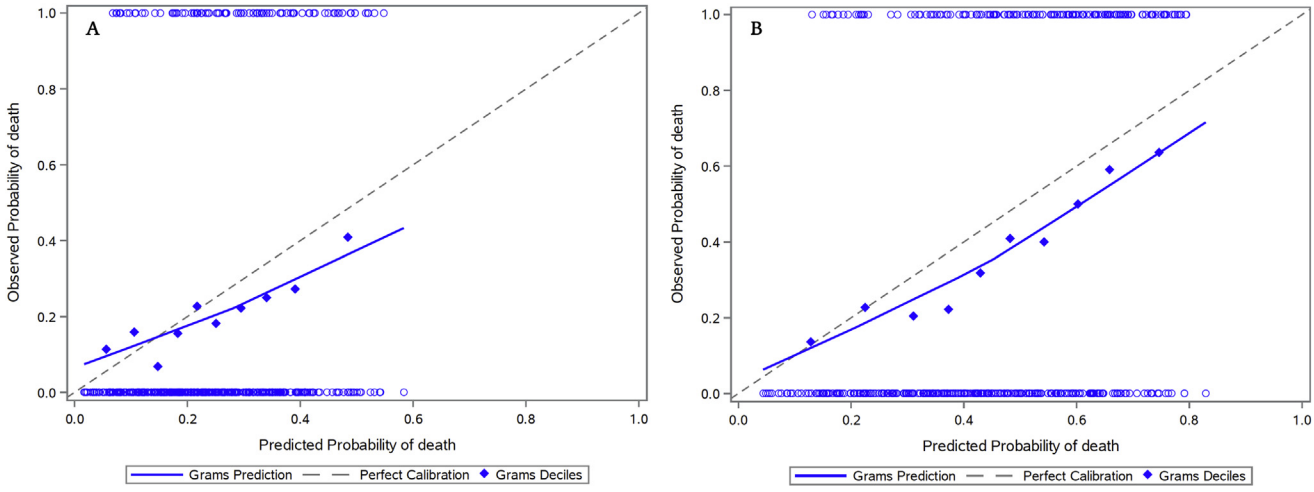


Figure 3. Calibration plots for the Grams model. The predicted probability of any death is shown on the x-axis, and the observed death rate is given on the y-axis. The dotted 45° line represents perfect agreement between the predicted and observed probabilities. The smoothed line is a locally weighted scatter plot smoothing curve across all predicted risk values and their corresponding observed risks. The dots represent the average observed probability of an outcome among each decile of the validation population. (A) Grams-2 death calibration plot; and (B) Grams-4 death calibration plot.

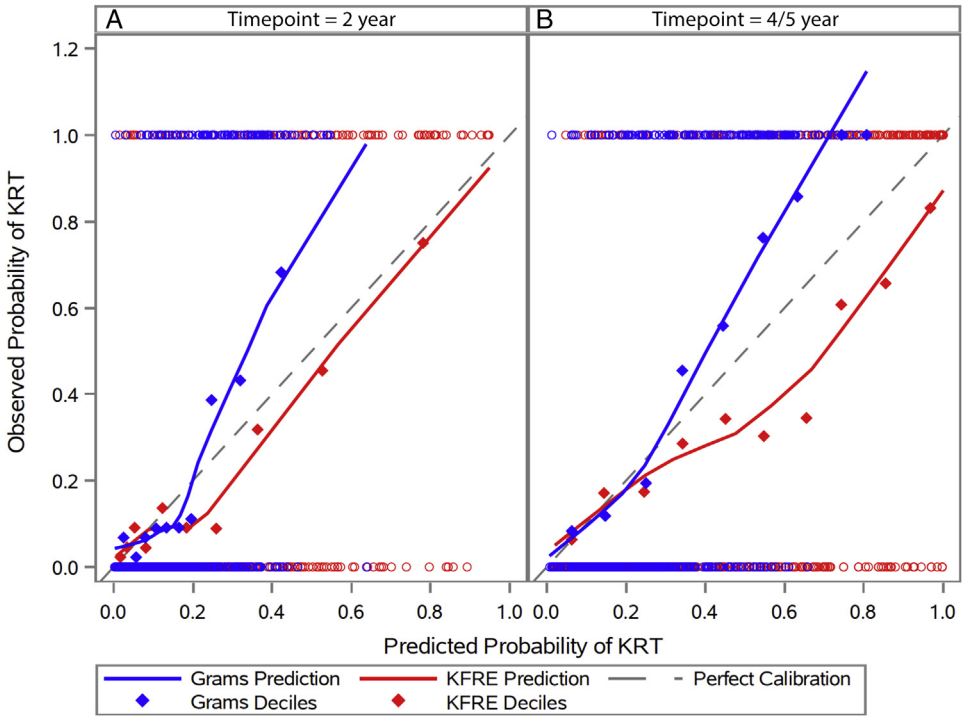


Figure 4. Grams-2 and KFRE-2 and Grams-4 and KFRE-5 KRT calibration plots. The 2-year KFRE and Grams models are depicted in (A), whereas the 4-year Grams and 5-year KFRE values are depicted together in (B). Blue lines and markers represent the Grams models, and the KFRE values are represented in red. The predicted probability of any KRT is shown on the x-axis, and the observed KRT rate is given on the y-axis. The dotted 45° line represents perfect agreement between the predicted and observed probabilities. The smoothed lines are a locally weighted scatterplot smoothing curve across all predicted risk values and their corresponding observed risks. The dots represent the average observed probability of an outcome among each decile of the validation population. Abbreviations: KFRE, kidney failure risk equation; KRT, kidney replacement therapy.

as death remains a difficult outcome to predict. Even widely used mortality prediction tools have been demonstrated to have only modest accuracy, with significant variability across various clinical settings.[22] We remain uncertain as to why the Grams-4 outperforms the Grams-2 for death discrimination.

Discrimination data specific to the outcome of death were not available for comparison in the development study, and ours is the first study to externally validate for this outcome. The Grams model was fairly well calibrated for the outcome of death, and while it tended to overpredict death throughout most of the prediction range, it only did so by approximately 10%. The observed death rates were higher in the Grams derivation cohorts (47%; mean follow-up, 3.5 years) than in our cohort (36% at 4 years; 47% at 5 years). These differences in baseline hazard rates may, in part, explain the approximately 10% overprediction.[9] We cannot comment on differences in patient characteristics between the 2 cohorts, as the development cohorts' characteristics are not provided.[9] The ability of the Grams-4 model to predict mortality with reasonable accuracy suggests that it could have an important role in augmenting clinical discussions around goals of care and patient education.

The Grams KRT models displayed good ability to discriminate between those who did and did not initiate KRT at 2 and 4 years, with AUCs of 0.81 (95% CI 0.76-0.87) and 0.82 (95% CI 0.77-0.86), respectively, which are very similar to those in the Swedish Renal Registry cohort (0.84 [95% CI, 0.83-0.85] vs 0.83 [95% CI, 0.82-0.83], respectively).[14] The KFRE KRT AUC results are also very similar or slightly higher in studies that also focused on advanced CKD populations (stage G4+).[11,14,15] Time-to-event discriminations using HCIs were fair to good, ranging from 0.77 to 0.80, with the KFRE-2 performing best (HCI, 0.80; 95% CI, 0.75-0.85). Our KFRE HCIs were comparable to those of the Swedish European Quality Study cohort (2-year HCI, 0.76 [95% CI, 0.72-0.80]; 5-year HIC, 0.75 [95% CI, 0.72-0.79]).[14]

Both the Grams and KFRE models were similarly well calibrated to KRT initiation at predicted risks of <15% at 2 years and predicted risks of <20% at 4 and 5 years but were poorly calibrated at higher ranges of predicted risk. The Grams model underestimates risk at risk thresholds of >20%, which was also demonstrated in the Swedish Swedish Renal Registry and the European Quality Study cohorts.[14] This diverges from the Grams' internal validation results, where the tendency was to overpredict risk,

although the degree of this was highly variable between cohorts.[9] The observed KRT rates were lower in the Grams' derivation cohort (12%; mean follow-up, 3.5 years) than our cohort (33% at 4 years), which likely contributes to the observed discrepancies.[9] The absence of summative data of patient characteristics in the Grams development cohort precludes a comparison of patient characteristics. Conversely, the KFRE models overpredict risks at higher thresholds. In both the original multinational KFRE validation study[16] and in the multinational stage G4+ KFRE validation study,[9] both marked underestimation and overestimation were seen in individual cohorts at all risk levels. Most other calibration studies have also demonstrated a tendency for the KFRE models to overestimate risk in advanced CKD[11,15] and in CKD stages G3-G5.[2] In the European Quality Study and Swedish Renal Registry cohorts, the KFRE-5 also overestimated risks, although, unlike in the current study, the KFRE-2 did not.[14] The KRT rate in the KFRE development cohort was only 11% (mean follow-up, 2.1 ± 2.0 years),[8] which is lower than the KRT rate of 20% observed in this study. This is not surprising given that our cohort had more advanced CKD with lower eGFRs and higher ACRs than the KFRE development cohort (Table S1). The differences in model performance between different cohorts underscore the importance of extensive model external validation in a variety of different settings and populations.

The overestimation of KRT initiation by the KFRE models likely stems from their lack of accounting for death as a competing risk. The Grams model accounts for the competing risk of death; it accordingly underestimates the risk of KRT initiation. The recent study by Ramspek et al[14] considered the competing risk of death in their discrimination and calibration analyses and found very similar findings to our own. Given this consistent overprediction and underprediction of KRT in terms of calibration for KFRE and Grams, respectively, and their similarly good discriminatory abilities, our findings suggest a potential role for using these models in tandem. In our population, the average of the KFRE and Grams scores appears to offer a more accurate clinical prediction score when the predicted risk scores are >20%. This should be further explored in other patient populations.

It has been appropriately advocated that risk prediction should be used to help guide patient education, clinical decision-making, and resource management in the care of patients with CKD.[4,5,23,24] In the design of a risk-based approach to clinical care, test performance at any given threshold, the outcome prevalence, and the potential ramifications of misclassifications on patient outcomes and resource utilization all need to be considered in the context of each center's population characteristics. Using provider discretion and taking into account the individual patient's clinical setting, these risk prediction tools can be used to augment discussions on prognostication. Thresholds for nephrology, multidisciplinary CKD care, and transplant or fistula planning referrals have been proposed and are used by some centers using the KFRE model.[5] Patients below proposed thresholds would still require monitoring and risk factor modifications, and those that do progress could then be appropriately referred. Tools and education to assist primary care providers in managing lower KRT risk patients with CKD need to be developed and easily available. These should particularly focus on cardiovascular disease– and kidney disease–modifying agents such as renin angiotensin system blockade and sodium-glucose cotransporter-2 inhibitors. Expensive resources required to manage nephrology clinics, multidisciplinary clinics, and KRT planning can then be more appropriately allocated to higher-risk patients. Risk predictions at the 2-year time point would be more reasonable for transplant, KRT education, and vascular access referrals, whereas the 4- and 5-year predictions may be more reasonable for nephrology referrals. To incorporate this risk-based approach into clinical decision-making, proposed thresholds need to be extensively externally validated in general nephrology and multidisciplinary CKD care referred populations.

Our study has several strengths. We included patients with CKD stage G4+ who were followed up in the only regional multidisciplinary CKD clinic, with minimal loss to follow-up and missing data. We also are the only regional KRT center; thus, it is highly likely that all KRT outcomes were captured. Vital status data were provided by the Ontario Registrar General, limiting the possibility of missing death events. Our validation cohort is similar to those of many other published CKD cohorts and is therefore representative of the population in whom the tools are used.[9] We acknowledge several limitations. Foremost, our analysis is limited by smaller sample sizes at the higher ranges of predicted risk, particularly for the Grams calculator. The observed KFRE scores were better spread out across the range of probabilities. Second, ours was a single-center study and, similar to the development cohorts, had a predominantly White cohort. Whether the findings can be extrapolated to racially diverse patient populations remains unknown. Additionally, we did not have a minimum eGFR at baseline for our cohort; thus, we included patients in our sample who would likely be opting for conservative care. However, the proportion of these patients with an eGFR < 6 mL/min/1.73 m^2 in our cohort was quite low overall (n = 4; 0.9%). Finally, the lack of integrated electronic medical records to access regional hospital records prevented an assessment of the cardiovascular event component of the Grams tool. This model offers the ability to predict the risks of cardiovascular events in patients with advanced CKD, which has been understudied.[25] This remains an important research priority, further highlighted by CVD being the leading cause of mortality (27%) in our cohort. Finally, we considered all deaths together instead of examining these separately pre- and post-KRT initiation, due to low numbers. Further study of the timing of death in larger cohorts is required.

In conclusion, in our cohort of patients with advanced CKD and high prevalences of diabetes and CVD, the Grams model for death provides a reasonably accurate estimation of death risk at the 4-year time point. Its incorporation into clinical care, in particular with respect to decisions around goals of care, should help both patients and their health care providers gain a better understanding of the likelihood of the crucial health outcome of death, which is often underappreciated and underdiscussed. In our cohort, the Grams model underestimates risks for KRT, whereas the KFRE model overestimates it at higher-risk thresholds. Further external validation studies, along with impact studies using model thresholds for clinical decision-making, are still required to confirm the usefulness of these prediction models.

SUPPLEMENTARY MATERIAL

Supplementary File (PDF)

Table S1. Comparison of the Kingston Health Sciences Center cohort to the KFRE development cohort.

ARTICLE INFORMATION

Authors' Full Names and Academic Degrees: Susan J. Thanabalasingam, MD, Eduard A. Iliescu, MD, Patrick A. Norman, MSc, Andrew G. Day, MSc, Ayub Akbari, MD, Gregory L. Hundemer, MD, and Christine A. White, MD, MSc.

Authors' Affiliations: Division of Nephrology, Department of Medicine, Queen's University, Kingston, Canada (SJT, EAI, CAW); Kingston General Health Research Institute, Kingston Health Sciences Center, Kingston, Canada (PAN, AGD); Department of Public Health Sciences, Queen's University, Kington, Canada (PAN, AGD); and Division of Nephrology, Department of Medicine, The University of Ottawa, Ottawa, Canada (AA, GLH).

Address for Correspondence: Christine A. White, MD, MSc, Division of Nephrology, Queen's University, Etherington Hall, 94 Stuart St., Kingston, Ontario, Canada, K7L 3N6. Email: cw38@queensu.ca

Authors' Contributions: Research idea and study design: CAW, SJT; data acquisition: SJT, EAI; data analysis/interpretation: PAN, AGD, CAW, SJT, AA, GLH; statistical analysis: PAN, AGD; supervision or mentorship: CAW, EAI. Each author contributed important intellectual content during manuscript drafting or revision and accepts accountability for the overall work by ensuring that questions pertaining to the accuracy or integrity of any portion of the work are appropriately investigated and resolved.

Support: None.

Financial Disclosure: The authors declare that they have no relevant financial interests.

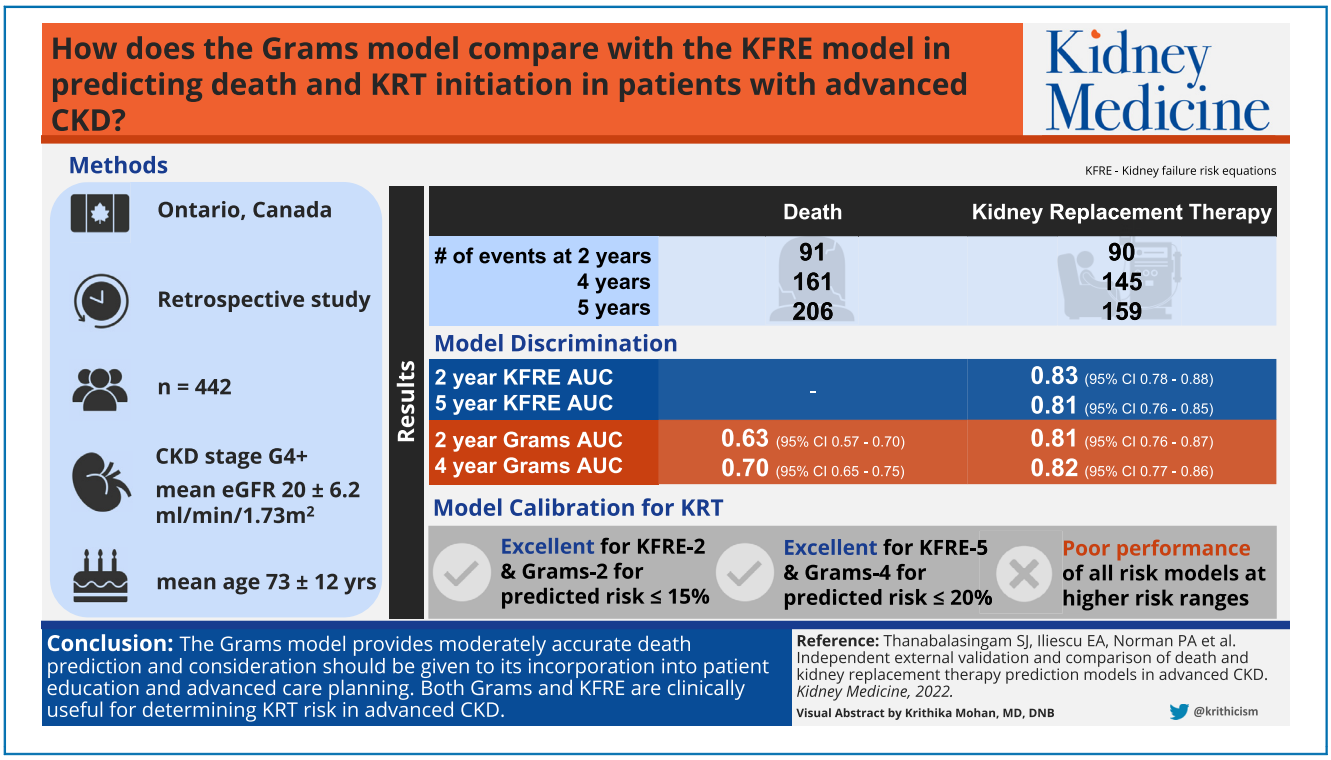
Acknowledgements: The authors acknowledge Andrew Goss, David Bruce, Frances MacLeod, and the Office of the Registrar General.

Peer Review: Received August 27, 2021. Evaluated by 3 external peer reviewers, with direct editorial input by the Statistical Editor and the Editor-in-Chief. Accepted in revised form December 29, 2021.

REFERENCES

1. Chronic Kidney Disease Prognosis Consortium; Matsushita K, van der Velde M, et al. Association of estimated glomerular filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: a collaborative meta-analysis. *Lancet*. 2010;375(9731):2073-2081.
2. Marks A, Fluck N, Prescott GJ, et al. Looking to the future: predicting renal replacement outcomes in a large community cohort with chronic kidney disease. *Nephrol Dial Transplant*. 2015;30(9):1507-1517.
3. Go AS, Yang J, Tan TC, et al. Contemporary rates and predictors of fast progression of chronic kidney disease in adults with and without diabetes mellitus. *BMC Nephrol*. 2018;19(1):146.
4. Farrington K, Covic A, Nistor I, et al. Clinical practice guideline on management of older patients with chronic kidney disease stage 3b or higher (eGFR <45 mL/min/1.73m^2): a summary document from the European Renal Best Practice Group. *Nephrol Dial Transplant*. 2017;32(1):9-16.
5. Tangri N, Ferguson T, Komenda P. Pro: risk scores for chronic kidney disease progression are robust, powerful and ready for implementation. *Nephrol Dial Transplant*. 2017;32(5):748-751.
6. Eckardt KU, Bansal N, Coresh J, et al. Improving the prognosis of patients with severely decreased glomerular filtration rate (CKD G4+): conclusions from a Kidney Disease: Improving Global Outcomes (KDIGO) Controversies Conference. *Kidney Int*. 2018;93(6):1281-1292.
7. Dekker FW, Ramspek CL, van Diepen M. Con: most clinical risk scores are useless. *Nephrol Dial Transplant*. 2017;32(5):752-755.
8. Tangri N, Stevens LA, Griffith J, et al. A predictive model for progression of chronic kidney disease to kidney failure. *JAMA*. 2011;305(15):1553-1559.
9. Grams ME, Sang Y, Ballew SH, et al. Predicting timing of clinical outcomes in patients with chronic kidney disease and severely decreased glomerular filtration rate. *Kidney Int*. 2018;93(6):1442-1451.
10. Lennartz CS, Pickering JW, Seiler-Mußler S, et al. External validation of the Kidney Failure Risk Equation and re-calibration with addition of ultrasound parameters. *Clin J Am Soc Nephrol*. 2016;11(4):609-615.
11. Peeters MJ, van Zuilen AD, van den Brand JA, et al. Validation of the kidney failure risk equation in European CKD patients. *Nephrol Dial Transplant*. 2013;28(7):1773-1779.
12. Potok OA, Nguyen HA, Abdelmalek JA, Beben T, Woodell TB, Rifkin DE. Patients,' nephrologists,' and predicted estimations of ESKD risk compared with 2-year incidence of ESKD. *Clin J Am Soc Nephrol*. 2019;14(2):206-212.
13. Whitlock RH, Chartier M, Komenda P, et al. Validation of the Kidney Failure Risk Equation in Manitoba. *Can J Kidney Health Dis*. 2017;4:2054358117705372.
14. Ramspek CL, Evans M, Wanner C, et al. Kidney failure prediction models: a comprehensive external validation study in patients with advanced CKD. *J Am Soc Nephrol*. 2021;32(5):1174-1186.
15. Hundemer GL, Tangri N, Sood MM, et al. Performance of the Kidney Failure Risk Equation by disease etiology in advanced CKD. *Clin J Am Soc Nephrol*. 2020;15(10):1424-1432.
16. Tangri N, Grams ME, Levey AS, et al. Multinational assessment of accuracy of equations for predicting risk of kidney failure: a meta-analysis. *JAMA*. 2016;315(2):164-174.
17. Evans M, Grams ME, Sang Y, et al. Risk factors for prognosis in patients with severely decreased GFR. *Kidney Int Rep*. 2018;3(3):625-637.

18. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med*. 2009;150(9):604-612.
19. Province of Onatrio. Ministry of Government and Consumer Services. Accessed August 12, 2020. https://www.ontario.ca/page/ministry-government-and-consumer-services
20. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837-845.
21. Tape TG. Interpreting diagnostic tests. Accessed March 10, 2022. http://gim.unmc.edu/dxtests/Default.htm
22. Siontis GC, Tzoulaki I, Ioannidis JP. Predicting death: an empirical evaluation of predictive tools for mortality. *Arch Intern Med*. 2011;171(19):1721-1726.
23. Kidney Disease: Improving Global Outcomes (KDIGO) Work Group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney Int*. 2012;3(suppl 1):1-150.
24. Akbari A, Clase CM, Acott P, et al. Canadian Society of Nephrology commentary on the KDIGO clinical practice guideline for CKD evaluation and management. *Am J Kidney Dis*. 2015;65(2):177-205.
25. Taal MW. Chronic kidney disease: towards a risk-based approach. *Clin Med (Lond)*. 2016;16(suppl 6):s117-s120.


How does the Grams model compare with the KFRE model in predicting death and KRT initiation in patients with advanced CKD?
Kidney Medicine
Methods
Ontario, Canada
Retrospective study
n = 442
CKD stage G4+
mean eGFR 20 ± 6.2 ml/min/1.73m2
mean age 73 ± 12 yrs
KFRE - Kidney failure risk equations
Results
Death
Kidney Replacement Therapy
# of events at 2 years 91 90
4 years 161 145
5 years 206 159
Model Discrimination
2 year KFRE AUC - 0.83 (95% CI 0.78 - 0.88)
5 year KFRE AUC - 0.81 (95% CI 0.76 - 0.85)
2 year Grams AUC 0.63 (95% CI 0.57 - 0.70) 0.81 (95% CI 0.76 - 0.87)
4 year Grams AUC 0.70 (95% CI 0.65 - 0.75) 0.82 (95% CI 0.77 - 0.86)
Model Calibration for KRT
Excellent for KFRE-2 & Grams-2 for predicted risk ≤ 15%
Excellent for KFRE-5 & Grams-4 for predicted risk ≤ 20%
Poor performance of all risk models at higher risk ranges
Conclusion: The Grams model provides moderately accurate death prediction and consideration should be given to its incorporation into patient education and advanced care planning. Both Grams and KFRE are clinically useful for determining KRT risk in advanced CKD.
Reference: Thanabalasingam SJ, Iliescu EA, Norman PA et al. Independent external validation and comparison of death and kidney replacement therapy prediction models in advanced CKD. Kidney Medicine, 2022.
Visual Abstract by Krithika Mohan, MD, DNB
@krithicism