Implementing the European Renal Best Practice Guidelines suggests that prediction equations work well to differentiate risk of end-stage renal disease vs. death in older patients with low estimated glomerular filtration rate

OPEN

Stein I. Hallan[1,2], Dena E. Rifkin[3,4,5], O. Alison Potok[3,4], Ronit Katz[6], Knut A. Langlo[1,2], Nisha Bansal[6,7] and Joachim H. Ix[3,4,5]

[1]Department of Clinical and Molecular Medicine, Faculty of Medicine, Norwegian University of Science and Technology, Trondheim, Norway; [2]Department of Nephrology, St. Olav Hospital, Trondheim, Norway; [3]Division of Nephrology-Hypertension, University of California San Diego, San Diego, California, USA; [4]Nephrology Section, Veterans Affairs San Diego Healthcare System, San Diego, California, USA; [5]Division of Preventive Medicine, Department of Family Medicine and Public Health, University of California San Diego, San Diego, California, USA; [6]Kidney Research Institute, University of Washington, Seattle, Washington, USA; and [7]Department of Nephrology, University of Washington Medical Center, Seattle, Washington, USA

Recent European guidelines suggest using the kidney failure risk equation (KFRE) and mortality risk equation for kidney disease (MREK) to guide decisions on whether elderly patients with chronic kidney disease should be referred early for dialysis preparation. However, the concurrent use of the two risk equations has not been validated. To do so we evaluated 1,188 individuals over five years with estimated glomerular filtration rate (eGFR) under 45ml/min/1.73m^2 and age over 65 years from the Norwegian population based HUNT study. Forty-two patients started renal replacement therapy and 462 died as their first clinical event. The KFRE was well calibrated (mean risk estimate 4.9% vs observed 3.5%) with high diagnostic accuracy (C-statistics 0.93). The MREK underestimated death risk in those with lower risk (mean risk estimate 30.1% vs observed 38.9%) and had moderate diagnostic accuracy (C-statistics 0.71). Only 31 individuals had estimated end stage kidney disease (ESRD) risk greater than death risk, and most experienced ESRD before death. Only two of 598 patients over 80 years old, and ten of 1,063 with eGFR 25-45ml/min/1.73m^2 at baseline experienced ESRD. Decision curve analysis demonstrated that for risk adverse patients, deferring ESRD preparation may be appropriate until predicted ESRD risk exceeds predicted death risk. For those preferring a more aggressive approach, referral when eGFR is under 25 ml/min/1.73m^2 may be beneficial if age remains under 80 years. Thus, the risk of ESRD is low compared to the risk of death in many older patients with chronic kidney disease stage 3b or worse, and combination of predicted ESRD and death risks, eGFR levels, age, and the patient`s valuations of harm and benefit can be helpful for deciding when to start dialysis preparations.

Correspondence: *Stein Hallan, University of Science and Technology, Department of Clinical and Molecular Medicine, Faculty of Medicine, Prinsesse Kristinas Gate 1, Trondheim 7030, Norway. E-mail: stein.hallan@ntnu.no*

Received 4 January 2019; revised 10 April 2019; accepted 19 April 2019; published online 10 May 2019

Kidney International (2019) **96,** 728–737; https://doi.org/10.1016/j.kint.2019.04.022

KEYWORDS: competing risk; kidney disease; mortality; prediction

Management of older patients with chronic kidney disease (CKD) is a common and challenging clinical scenario driven by the competing risks of ESRD and death. For the patient who will ultimately develop ESRD, early referral to a nephrologist for preparation with interventions such as placement of an arteriovenous fistula, a peritoneal dialysis catheter, or transplant evaluation are likely to improve clinical outcomes and life expectancy.[1] On the other hand, for the patient who will die before developing ESRD, the unwarranted concerns about future ESRD and the costs and risks associated with preparation for dialysis are unnecessary. Prior studies suggest that, on average, the risk of death is higher than the risk of ESRD in earlier stages of CKD, but in practice, identifying risk in individual patients remains difficult.[2]

A joint initiative of the European Renal Association–European Dialysis Transplant Association (ERA–EDTA) and the European Union Geriatric Medicine Society (EUGMS) recently published a clinical practice guideline[3] addressing the clinical approach to patients aged >65 years with an eGFR <45 ml/min per 1.73m^2. These European Renal Best Practice (ERBP) guidelines leveraged 2 prediction equations that have been developed and validated for both ESRD (the KFRE, developed by Tangri *et al.*[4]), and for mortality risk (MREK, developed by our group[5]) in CKD patients. The

guidelines recommend an algorithm in which both equations are applied concurrently to the same patient. If the risk for ESRD is estimated to be high, the mortality risk is low, and the patient is not frail, the algorithm suggests shared decision-making, pre-dialysis counseling, and modality selection. On the other hand, if the competing risk of death is estimated to be high, the guidelines advocate not stressing the risk of ESRD or potential future need for renal replacement therapy, and instead focusing primarily on advanced care planning. However, the ERBP guidelines do not specify how frailty should be assessed, and formal frailty scores are so heterogeneous that results based on different scoring systems cannot be compared or pooled.[6] Furthermore, what constitutes high risk for either outcome is not defined by the guidelines. For example, one patient may consider preparation for ESRD to be justified only if their risk of ESRD is higher than their risk of death. Another might consider it reasonable to begin preparations for ESRD when the risk is only 20% as great as the risk of death. Currently, very few studies or guidelines provide guidance on these factors, although referral to a nephrologist has been suggested for a 5-year ESRD risk of >5%–10% and renal replacement preparation for a 1-year ESRD risk of 10%–20%.[1,7] Comparable risk thresholds for mortality risk, and how to incorporate mortality risk at any given risk level of ESRD, have not been addressed.

Aside from the issue of defining high risk, use of the 2 prediction equations concurrently in the same individual has never been tested. The KFRE and MREK were developed and validated in different study populations, so it is not known whether they are well calibrated in the same populations. For example, in developing and validating the MREK, we evaluated[4] patients with a higher eGFR than those proposed in the ERBP guideline (mean eGFR 48 ml/min per 1.73 m^2). Furthermore, shared decision-making and including the patient perspective are important, but not well studied, in the elderly CKD patient approaching the life-changing decisions of ESRD treatments.[8,9]

Our purposes for this analysis were several. First, we sought to validate the performance of the 2 equations individually in older persons with CKD stage 3b or worse. Second, in an effort to simulate the clinical scenarios in individual patients, we sought to employ both equations in the same population concurrently to determine how risk of death and ESRD compare with one another in a European cohort of older patients with eGFR <45 ml/min per 1.73 m^2, as suggested by ERBP guidelines. Finally, we sought to evaluate the clinical impact of nephrologist referral algorithms suggested in these and other guidelines, while incorporating a gradient of possible patient valuations of risk and benefits.

RESULTS

Study population characteristics

Among 78,960 participants from the general population-based Nord-Trøndelag Health Study (HUNT; Norway)-2 and HUNT-3 studies, 23,880 subjects were aged ≥65 years. Among these, 1188 (5.0%) had an eGFR of <45 ml/min per 1.73 m^2 and constituted the study sample for this analysis (Figure 1). Their mean age was 80 years (SD: 7); their mean eGFR was 36 ml/min per 1.73 m^2 (SD: 8); 57% were female; 17% had diabetes; and 4% had an albumin–creatinine ratio of >30 mg/mmol (SD: 300 mg/g). Additional study population characteristics are summarized in Table 1, while baseline characteristics by first future outcome (none, ESRD, or death) are depicted in Supplementary Table S1.

External validation of the 2 prediction equations in older Europeans with eGFR <45 ml/min per 1.73 m^2

Overall, the agreement between observed rates of end points relative to the predicted rates was good for both equations (see Supplementary Excel Calculator for the concurrent prediction of mortality and kidney failure risk). The mean 5-year ESRD risk based on the KFRE was 4.9%, and the observed rate was 3.5% (n = 42) over the 5-year observation period. The mean mortality risk by the MREK was 30.1%, and the observed death rate was 38.9% (n = 462; Table 2). Logistic regression of outcomes on individual predicted risks indicated that both the KFRE and MREK were well calibrated on average (regression intercepts 0.28 ($P = 0.28$) and 0.17 ($P = 0.08$), respectively), but calibration slopes were not equal, although close, to 1.0 (1.3 and 0.8, $P < 0.01$ for both). However, visual inspection of outcome versus mean predicted risk by 5 predefined risk categories indicated that the calibration slope was not linear (Figure 2). A nonlinear regression showed that the MREK slightly underestimated mortality risk in the lower range, while the KFRE was very well calibrated but with wider confidence intervals, likely due to the lower number of ESRD cases relative to deaths. The 2 equations explained 33% and 10%, respectively, of the variation in the 2 outcomes, based on R^2. The ability to discriminate between an elderly CKD patient with a future ESRD event and a patient without an ESRD event was excellent for the KFRE (area under the receiver operating curve: 0.93). The MREK performed moderately well for death prediction, with an area under the receiver operating curve of 0.71, that is, being able to correctly classify 71 of 100 pairs of cases and controls (Table 2). The performance of these equations was similar when evaluated in the subset of 241 individuals with baseline eGFR <30 ml/min per 1.73 m^2 (KFRE 0.87 [95% confidence interval 0.80–0.94] and MREK 0.75 [95% confidence interval 0.68–0.82]). Kaplan-Meier survival plots showed superimposed survival lines for quartiles 1–3 for the KFRE regarding ESRD risk, and higher risk in quartile 4 (log-rank test for quartile 3 vs. quartile 4 = 0.0002; Supplementary Figure S1). For the MREK, each quartile of mortality risk clearly separated the elderly CKD patients (log-rank test <0.05 for all comparisons).

Performance of jointly employed ESRD and mortality-risk equations in elderly Europeans with an eGFR <45 ml/min per 1.73 m^2

Figure 3 depicts the association between the predicted ESRD and mortality risks in elderly CKD patients by either

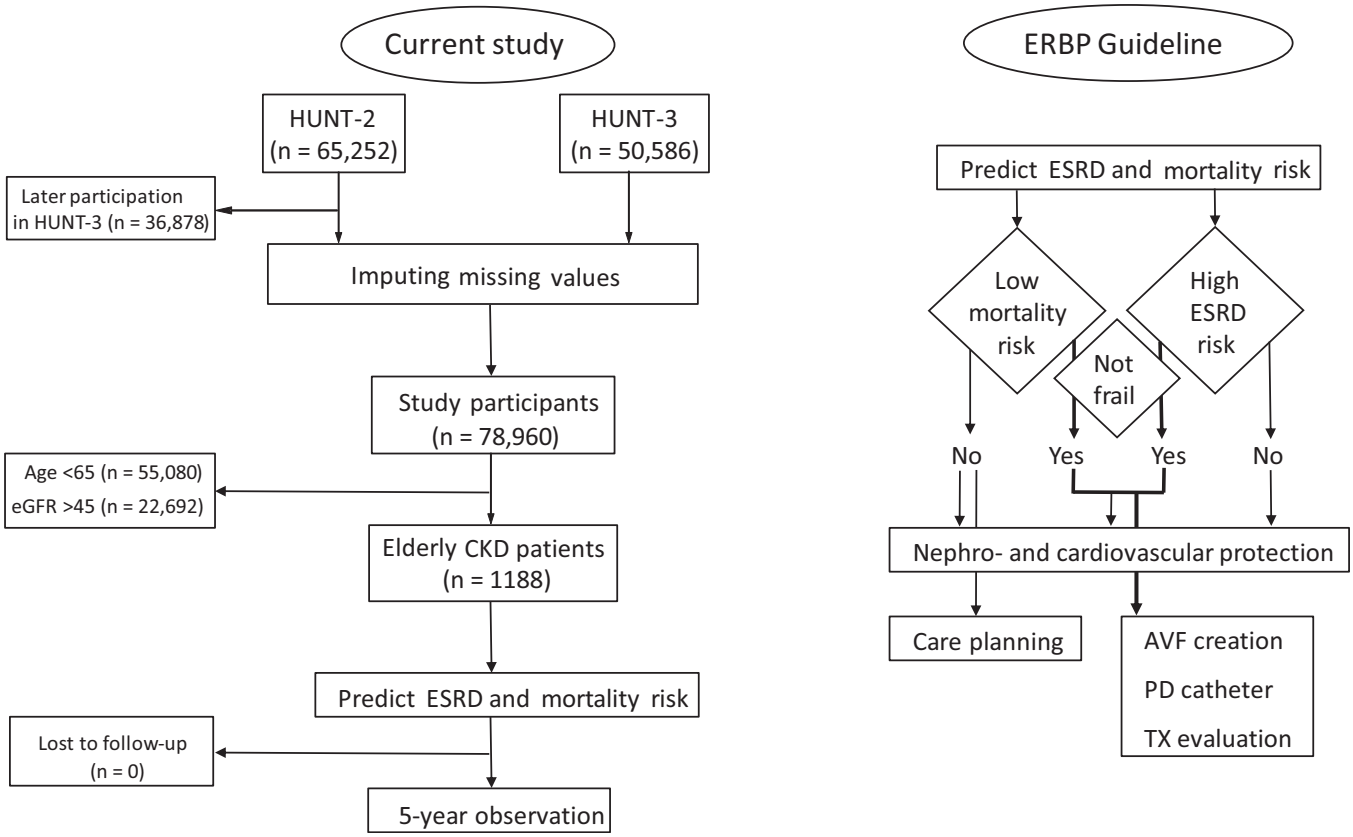

Figure 1 | Study design showing selection of study participants and an overview of the European Renal Best Practice (ERBP) guidelines for older chronic kidney disease (CKD) stage 3B+ patients. AVF, arteriovenous fistula; eGFR, estimated glomerular filtration rate; ESRD, end-stage renal disease; HUNT, Nord-Trøndelag Health Study, Norway; PD, peritoneal dialysis; TX, transplant.

prediction equation, and uses color-coding for the first observed actual clinical event during follow-up. On average, the 5-year mortality risk was approximately 10-fold higher than the ESRD risk (interquartile range: 0.16–0.40 vs. 0.01–0.04). Although very few individuals had predicted ESRD risk > mortality-risk estimates (datapoints above the line of identity, n = 31 [2.6%]), ESRD was the first clinical event for the majority in this group: 19 (61%) experienced ESRD before death, 5 (16%) died without ESRD, and 7 (23%) experienced neither event. Clinical characteristics for those with ESRD risk > mortality risk and vice versa are displayed in Table 3 by their first event. Participants with predicted ESRD > mortality risk who ultimately developed ESRD were more frequently male, had very low baseline eGFR, had higher blood pressure, and were more frequently receiving antihypertensive medications.

Implications of implementation of the ERBP guidelines and other referral algorithms

A partial least-squares–discriminant analysis identified age, body mass index, self-reported health status, systolic blood pressure, and eGFR as the most important baseline characteristics for discriminating between future ESRD, death, and event-free survival (Supplementary Figure S2). Given that age, eGFR, and general health are also used to define the population for the current ERBP guidelines, we displayed outcomes by these variables (Figure 4). The fraction experiencing ESRD before death decreased substantially with older age. Among the subset aged ≥80 years, only 2 of 598 experienced ESRD over the next 5 years. Both ESRD and mortality increased in individuals with lower baseline eGFR categories. However, ESRD was much less frequent than death in all eGFR categories, except for those with eGFR <15 ml/min per 1.73 m^2 at baseline, of whom approximately two thirds experienced ESRD before death. As expected, the mortality risk was higher at lower levels of self-reported health at baseline, but a large proportion of those who ultimately experienced ESRD before death were among those who reported poor health at baseline.

Finally, we used decision curve analysis to evaluate the clinical utility of implementing different algorithms that might be used to initiate preparation for dialysis by early referral to a nephrologist (Figure 5). Higher numbers on the y-axis suggest greater net benefit, and the patient`s perception of the harm:benefit ratio associated with referral is depicted across the x-axis. The best referral algorithm is represented by the highest level on the y-axis at any level of the x-axis in Figure 5. In this analysis, "benefit" comes from timely preparation for renal replacement therapy in those who ultimately experience ESRD as their first event, and "harm" comes from

Table 1 | Baseline characteristics of study participants

Characteristic	Total population (n = 78,960)	Elderly chronic kidney disease patients (n = 1188)
Male	46.8	43.2
Age (yr)	53.7 (17.8)	79.9 (6.8)
Higher education	18.2	6.2
General health		
"Poor" (Frail)	2.0	7.2
"Not so good"	27.2	56.9
"Good"	56.1	34.2
"Very good"	14.8	1.7
Diabetes	5.0	17.2
Myocardial infarction	4.2	22.8
Stroke	2.9	12.4
Heart failure	1.5	14.7
Smoker		
Never	46.4	53.0
Former	28.3	33.0
Current	25.3	14.0
Physically inactive	64.5	92.1
BP medication	20.2	67.1
Systolic BP (mm Hg)	134.8 (21.6)	147.3 (27.5)
Diastolic BP (mm Hg)	76.3 (12.7)	77.2 (16.2)
Body mass index (kg/m^2)	27.0 (4.4)	28.0 (4.5)
eGFR (ml/min per 1.73 m^2)	94.8 (20.3)	35.8 (7.8)
Stage (ml/min per 1.73 m^2)		
G3a (45–59)	4.2	Not included
G3b (30–44)	1.3	79.7
G4 (15–29)	0.3	18.3
G5 (<15)	0.04	2.0
ACR (mg/mmol)	1.9 (4.2)	7.1 (18.7)
Stage (mg/mmol)		
A2 (3–29)	8.9	53.0
A3 (≥30)	0.3	4.2

ACR, urine albumin–creatinine ratio; BMI, body mass index; BP, blood pressure; eGFR, estimated glomerular filtration rate.
Data are mean (1 SD) or percentage.

Table 2 | External validation of Kidney Failure Risk Equation (KFRE) and Mortality Risk Equation for Kidney Disease (MREK) in the subset of the HUNT population aged >65 years and with eGFR <45 ml/min per 1.73 m^2 at inception

Performance of risk prediction	5-year ESRD risk (observed, 3.5%)	5-year mortality risk (observed, 38.9%)
Risk equation	KFRE	MREK
Calibration		
Mean predicted risk (%)	4.9	30.1
Calibration curve intercept	0.28 ($P = 0.38$)	0.17 ($P = 0.08$)
Calibration curve slope	1.30 ($P = 0.001$)	0.84 ($P = 0.001$)
Hosmer-Lemeshow test	14.0 ($P = 0.08$)	5.2 ($P = 0.74$)
Discrimination		
Area under ROC curve	0.926	0.711
Overall goodness-of-fit		
Pseudo R^2	0.333	0.102
Clinical usefulness (%)		
Sensitivity	28.6	33.2
Specificity	99.5	89.4

eGFR, estimated glomerular filtration rate; ESRD, end-stage renal disease; HUNT, Nord-Trøndelag Health Study in Norway; ROC, receiver operating curve.
For good calibration, the calibration curve intercept should be close to 0.0 (indicating average calibration), and the slope should be close to 1.0. Hosmer-Lemeshow tests whether there is evidence that the model is not well calibrated (i.e., *P* values closer to 1.0 indicating better calibration). ROC describes discrimination (i.e., ability to differentiate a patient with a future outcome from a patient without the outcome of interest). Pseudo R^2 describes the goodness-of-fit for the model (i.e., explained variability). Sensitivity is the probability of a case testing positive. Specificity is the probability of a control testing negative (cutoff for probability of disease ≥0.50 for positive test).

the same interventions among those who experience death as their first event. An elderly CKD patient might believe that both substantial harms and benefits are associated with early referral (e.g., a harm:benefit ratio of 2:3). Such a patient would benefit from referral only if the predicted ESRD risk is higher than the mortality risk (blue line in Figure 5). However, some patients might wish to be more aggressive with their care if they believe that early referral is more advantageous (i.e., they have a lower harm:benefit ratio). Presuming a harm:benefit ratio of 1:4, this individual should be referred when their eGFR is <25 ml/min per 1.73 m^2 if age remains <80 years (i.e., the green line in Figure 5).

In this analysis, the ERBP guideline recommendation of referring those with "ESRD risk > mortality risk and not being frail" (thick red line) was never the best alternative in this model. The current Kidney Disease: Improving Global Outcomes (KDIGO) recommendation of referring individuals with a 5-year ESRD risk >50% (1-year >10%) (orange line in Figure 5) would be appropriate for only those seeking a conservative treatment plan (considering harm:benefit ratio worse than 1:1). We also tested other alternative referral algorithms derived by visual inspection of Figure 3 and clinical experience: ESRD risk higher than a specific fraction of the mortality risk (0.10/0.25/0.33/0.50 × mortality risk), or ESRD risk higher than a specific cutoff (5%/10%/20%). However, these algorithms were never the best choice, irrespective of the harm:benefit ratio (see Supplementary Figures S3 and S4).

Next, we performed sensitivity analyses to evaluate the impact of different ESRD risks; that is, inflating the ESRD risk 3-fold. The accuracy/performance of the risk equations are not influenced by the disease prevalence, but the competing mortality risk would be less overwhelming in those aged 65–80 years. In these analyses, the decision curve analysis (DCA) results were very similar to our primary analyses, but with higher net benefit and lines shifted to the right. This means that with a higher ESRD incidence, referral of those with an eGFR <25 and age <80 years still appeared to be the best option over an even wider range of harm:risk ratios. (See Supplementary Figures S5 and S6.) Overfitting can be a problem in DCA, but an analysis based on one hundred 10-fold cross-validations gave very similar results.

DISCUSSION

In a European cohort of 1188 individuals aged ≥65 years and with an eGFR <45 ml/min per 1.73 m^2 at inception, we evaluated the performance of 2 prediction equations recommended by the ERBP guidelines for evaluating competing risk of ESRD and death. The equations performed well overall. When deployed concurrently, only 2.6% had a predicted risk of ESRD higher than their risk of death over 5 years. However, among this high-risk subset, most experienced ESRD rather than death during follow-up. These findings have clinical implications for

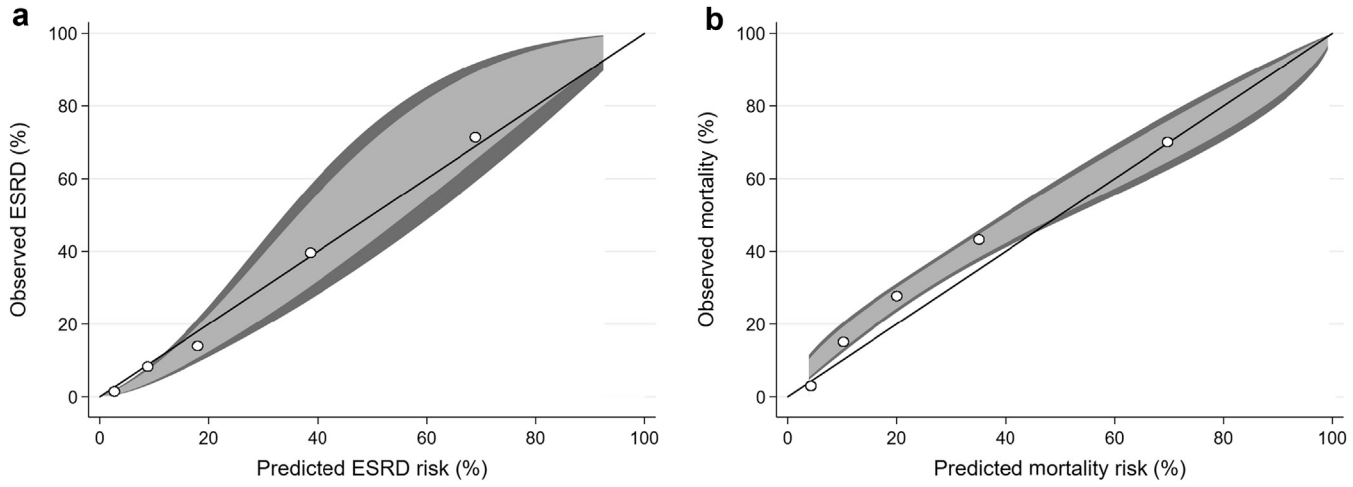

Figure 2 | Predicted versus observed 5-year risk of end-stage renal disease (ESRD) (a) and death (b), using the Kidney Failure Risk Equation and the Mortality Risk Equation for Kidney Disease, respectively, in elderly chronic kidney disease patients. White dots indicate observed versus mean predicted risk, by risk category (0.0%–4.9%, 5.0%–14.9%, 15%–24.9%, 25.0%–49.9%, and 50.0%–100.0%). Function is a nonlinear regression of outcome on predicted risk; 80% and 95% confidence intervals for the risk estimate are different from the line of perfect fit (the 45° diagonal).

guiding shared decision-making between patients and providers, and potentially in shaping public policy.

The American Geriatrics Association set forward guiding principles for treatment of older patients: (i) assess patient preferences; (ii) interpret the available evidence; (iii) estimate prognosis; (iv) consider treatment feasibility; and (v) optimize therapies and care plans.[10] The new ERBP guidelines address several of these topics, but they also recognized several important gaps in available data. Our study addresses several of these. First, the disease burden, as well as the burden of treatment, including preparations like vascular access creation,[11] is high for patients with CKD stage 3b or worse,[12] and studies have previously demonstrated that nephrologists unfortunately have limited knowledge of their patients' priorities.[13] We utilized DCA to illustrate how different patient preferences may influence the appropriate timing of referral to a nephrologist in the HUNT population. Second, our study generally validated the performance of both prediction equations in elderly patients with an eGFR <45 ml/min per 1.73 m^2. This validation is particularly important for the MREK because it had been developed and validated in US populations with higher baseline eGFR. Third, no prior study, to our knowledge, has evaluated utilization of both equations concurrently to test how the approach advocated by the ERBP guidelines might perform when used in clinical practice. Our findings quantify that the mortality risk is much higher than the ESRD risk in older community-living persons (10:1), despite an eGFR of <45 ml/min per 1.73 m^2 at baseline, a finding consistent with results of other studies.[14] Finally, our study evaluates the implications of nephrologist referral. The DCA analysis gives examples of how referral algorithms might differ by patient preferences, suggesting that patients should be referred to nephrologists based on various criteria, using ESRD and mortality-risk estimates, eGFR levels, and age, layered onto patient preferences regarding perceived benefits and harms.

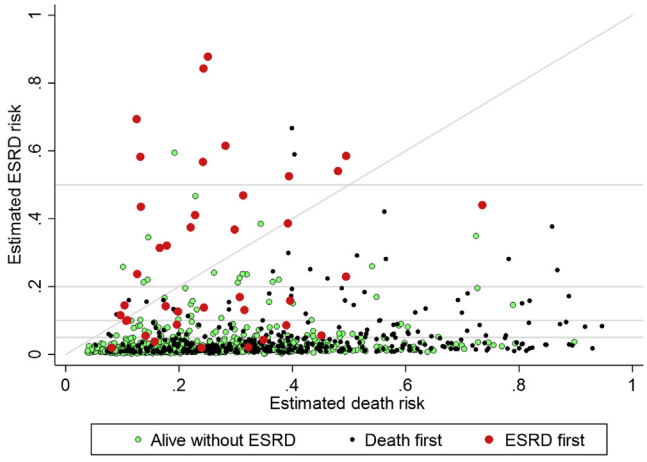

Figure 3 | Distribution of end-stage renal disease (ESRD) and mortality risk estimates in elderly chronic kidney disease patients with color codes for first future outcome.

The KFRE was developed in CKD clinic populations, and not in older community-living populations per se.[5] It has, however, been validated in a meta-analysis of 31 different cohorts.[15] The equation achieved excellent 5-year discrimination (C statistic 0.88; 95% confidence interval, 0.86–0.90), and discrimination was similar by age groups (<65 vs. ≥65 years). Calibration was also very good in European cohorts. The ERBP guidelines recommended that the KFRE should also be validated in elderly populations with an eGFR <45 ml/min per 1.73 m^2, which is accomplished by our study. We found that the KFRE equation performed well among those with lower kidney function, and that the C-statistic and calibration were similar to those in prior studies.[15]

The MREK equation was developed and validated in elderly populations in the US with only moderate CKD

Table 3 | Characteristics of participants with higher end-stage renal disease (ESRD) risk versus participants with higher mortality risk, and by the first clinical event experienced during follow-up

	Predicted ESRD risk > predicted mortality risk (n = 31)		Predicted mortality risk > predicted ESRD risk (n = 1157)	
Characteristic	ESRD first (n = 19)	Death first (n = 5)	Death first (n = 339)	ESRD first (n = 20)
Male	68.4	40.0	55.5	65.0
Age (yr)	72.3	73.5	82.2	74.2
Higher education	21.1	0.0	7.1	5.0
General health				
"Poor"	16.7	20.0	10.0	5.3
"Not so good"	55.6	80.0	63.8	79.0
"Good"	27.8	0.0	26.0	15.8
"Very good"	0.0	0.0	0.3	0.0
Diabetes	10.5	0.0	22.1	25.0
Myocardial infarction	26.3	0.0	29.8	15.0
Stroke	0.0	20.0	14.5	10.0
Heart failure	5.2	20.0	23.3	15.0
Smoker				
Never	36.8	60.0	48.4	30.0
Former	42.1	40.0	41.3	45.0
Current	21.1	0.0	10.3	25.0
Physically inactive	94.7	100.0	93.1	94.1
BP medication	84.2	80.0	61.1	75.0
Systolic BP (mm Hg)	149.8	140.6	144.7	147.6
Diastolic BP (mm Hg)	82.6	74.1	75.8	82.3
Body mass index (kg/m^2)	26.2	25.5	27.1	28.0
eGFR (ml/min per 1.73 m^2)	12.9	18.1	35.1	26.9
ACR (mg/mmol)	29.0	31.2	8.3	13.2
Predicted ESRD risk	47.9	33.8	4.5	12.7
Predicted mortality risk	23.7	22.9	40.0	29.0

ACR, albumin–creatinine ratio; BP, blood pressure; eGFR, estimated glomerular filtration rate.
Values are % unless otherwise indicated.

(eGFR 47 ± 11 ml/min per 1.73 m^2 and eGFR 50 ± 9 ml/min per 1.73 m^2 in the development and validation cohorts, respectively). The ERBP guidelines called for additional research to validate its performance in elderly patients with CKD stage 3b or worse.[3] We found identical discrimination (C statistic 0.722) relative to its validation in prior studies, despite evaluating a cohort with more advanced CKD (mean eGFR 36 ± 8 ml/min per 1.73 m^2). However, the MREK equation underestimated the mortality risk in the lower range, and this suboptimal calibration likely reflects its implementation in a population with lower baseline eGFR than that in which it was developed.

In clinical practice, if an older individual with CKD is at high risk of ESRD, and both the patient and provider agree that dialysis or transplantation would be in the patient's best interest, preparation for dialysis should be started in due time. Interventions might include placement of an arteriovenous fistula, which requires a surgical procedure and is not without risk. Should an older patient's risk of mortality be substantially higher than the risk of ESRD, avoidance of these procedures could spare the patient the financial, psychological, and physical costs associated with intensified nephrology care and ESRD preparation. Instead, clinical efforts to maximize functional status would be higher priority. The risk threshold at which ESRD planning should be initiated will differ for individual patients, practices, countries, and economies. However, as an example in our community-living Norwegian population, the vast majority of elderly persons with CKD had a higher 5-year mortality risk than ESRD risk. Initiating dialysis planning if 1-year ESRD risk is >10%, as suggested by Kidney Disease: Improving Global Outcomes may not be appropriate, except for patients who prefer the most conservative dialysis planning regimen. In our clinical experience, many elderly patients prefer a more aggressive approach (harm:benefit ratio better than 1:1), and referral when estimated ESRD > mortality risk or when eGFR <25 in patients <80 years old appeared to be the best algorithms. Results were similar in sensitivity analyses in which the ESRD incidence rate was inflated 3-fold. A rough estimate of the patient's perception of harm versus benefit can be elicited through an ordinary consultation or by using more formal techniques.[16] However, our study evaluated community-living individuals, and we recognize that risk of both ESRD and death may be higher in referral populations. Additional studies are required to evaluate DCA analyses in other settings.

Strengths of our study include the large sample size of the HUNT study, which allowed us to define a sufficiently large subset aged ≥65 years with eGFR <45 ml/min per 1.73 m^2 at baseline in order to test the recommendations put forward by ERBP. We also tested the implementation of these 2 risk equations concurrently for the first time. However, we only studied white Europeans, which limits the generalizability of our results to other ethnic groups. We note that the Norwegian mortality, morbidity, health care expenditure, accessibility, and dialysis treatment indications are all comparable to other European countries such as the United Kingdom, France, and Greece, and to some extent to the US (Supplementary Table S1). Future studies should be conducted in these settings to determine if the results generalize. Regions with a higher ESRD incidence will have somewhat reduced competing mortality risk, but sensitivity analysis showed that our main findings remained robust. By using DCA, we evaluated various referral algorithms over a wide range of harm-to-benefit ratios. We believe such analyses may be useful to patients and providers to develop individualized treatment approaches that can accommodate differences in patient preferences. Heart failure, which is part of the MREK equation, was not assessed in the HUNT-2, but we imputed this variable based on information from HUNT-3 participants. Most agree that modern imputation techniques improve precision and avoid selection bias,[17] especially when the imputations are based on information from tens of thousands of subjects having the actual variable directly measured, as was done here,[18,19] but we acknowledge the limitation of missing data. We may also have missed some kidney failure events in patients who were not offered dialysis or who chose conservative care (7%–16%),[20] and we did not assess frailty with a standard score.

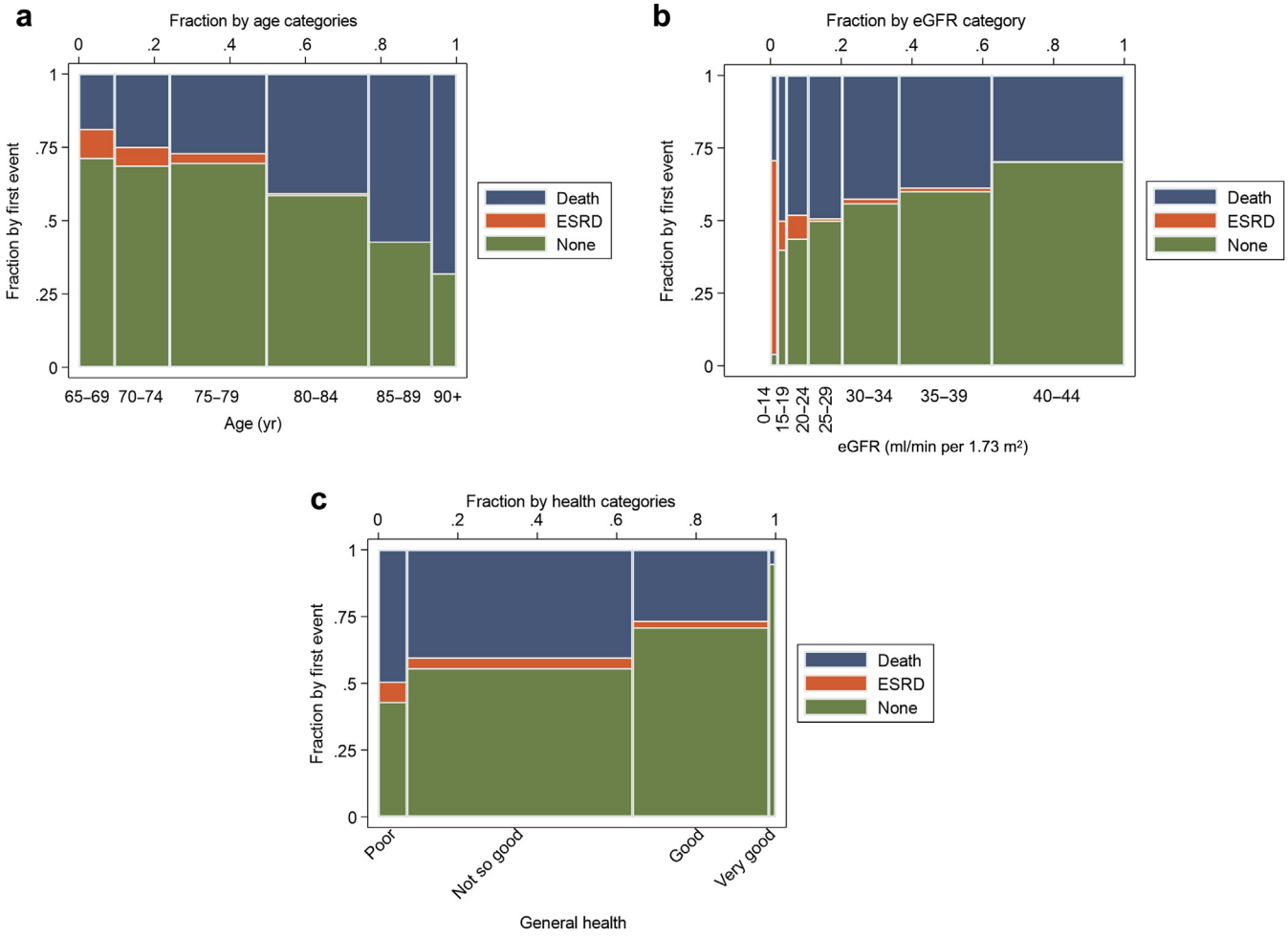

Figure 4 | Fraction by first event (end-stage renal disease [ESRD], death, or none) versus age categories (a), estimated glomerular filtration rate (eGFR) categories (b), and self-reported health categories (c).

In summary, the kidney failure and mortality risk equations put forward for concurrent use on older patients with an eGFR <45 ml/min per 1.73 m^2, by the ERBP guidelines (see Supplementary Excel Calculator), performed reasonably well in a large community-living population in Norway. Despite the low eGFR at baseline, the risk of death was more than 10-fold the risk of ESRD, on average. In clinical practice, our findings may assist clinicians as they decide which patients to refer for dialysis preparation, and which to manage without referral. In particular, in our study, the ERBP algorithm recommendation—that patients with "ESRD risk > mortality risk and not being frail" should be referred for dialysis preparation— did not appear to be the optimal referral algorithm, regardless of patient perspectives on harms and benefits. Overall, patients under the age of 80 years who seek a more aggressive mode of care should be referred when eGFR is < 25 ml/min per 1.73 m^2, and patients who prefer a more conservative mode should be referred based on ESRD versus mortality risk estimates. Our data may assist clinicians in communicating these trade-offs to their patients. The data may also assist public policy makers to avoid unnecessary costs, and simultaneously allow intensification of resources for patients who are at highest risk for ESRD. Overall, deploying both risk prediction equations concurrently in elderly Europeans with CKD stage 3b or greater led to good performance, but referral for dialysis preparation should also be based on consideration of eGFR, age, and the patient`s own perceptions of the trade-offs between potential harms and benefits.

METHODS

Study design and participants

The Nord-Trøndelag Health Study (HUNT) is a large general health study inviting all residents of Nord-Trøndelag County, Norway, to study visits every 10 years.[21] The county has a population of ~130,000 residents (>97% white), and is representative of Norway in regard to demographics, income, mortality, and morbidity, including ESRD risk. Furthermore, relevant aspects of Norwegian health care in general and kidney medicine in particular are not substantially different from that in the rest of Europe and the US (Supplementary Table S2[S1–S5]). Each survey comprised an extensive questionnaire on medical history and risk factors, and a clinical examination. We included subjects who had participated in either the HUNT-2 (1995–1997) or HUNT-3 (2006–2008) visits who were aged >65 years and had an eGFR <45 ml/min per 1.73 m^2. The Norwegian Death Registry is 100% complete for vital status and 98%

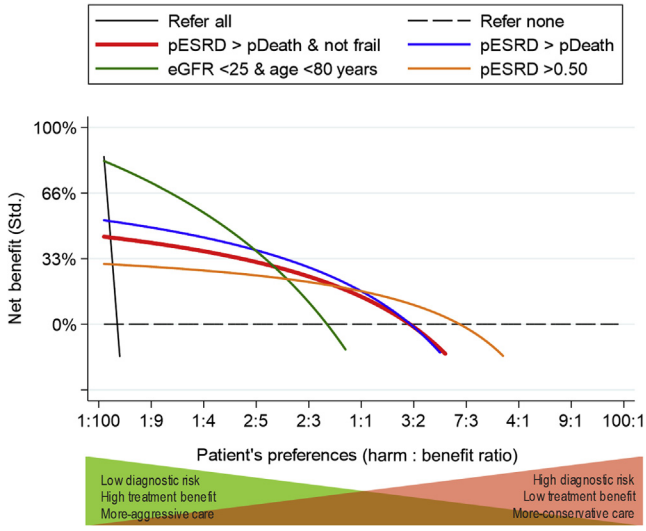

Figure 5 | Clinical utility of nephrology referral algorithms relative to patients' valuation of harm versus benefit. The European Renal Best Practice guideline algorithm (if "ESRD [end-stage renal disease] risk > mortality risk and not frail") is represented by the thick red line. In this modified decision curve analysis, utility (also known as net benefit) is the benefit minus the harm for the total group, adjusted for the individual patient's perception of the trade-off between harm and benefit (utility = [true positives/N] – [false positives/N] × harm:benefit ratio). Utility is expressed as a percentage of the maximal possible utility (i.e., divided by the prevalence of ESRD [0.035]). The maximum theoretical utility (100%) will, therefore, occur with a referral algorithm that selects all future ESRD cases (48/1188 = 0.035) and with a patient who feels that the harm:benefit ratio is very advantageous (e.g., 1:100), meaning that the negative impacts of referring patients, even if they are false positives for eventual risk of ESRD, is close to zero. For example, if the patient and provider both feel that the harm:benefit ratio is 1:4, the best option will be to refer the patient when estimated glomerular filtration rate (eGFR) is <25 and age is <80 years (green line), which will give a clinical utility of 48% of the maximum given current prevalence. If the harm:benefit ratio is 1:1, then referral when the predicted ESRD risk is greater than the predicted mortality risk would provide maximum clinical utility. p, probability.

for cause of death,[22] and each HUNT participant was linked to central registries utilizing the unique 11-digit identification number given to all Norwegian citizens at birth. Follow-up time was censored at 5 years, as both prediction equations were designed to estimate 5-year risk. Figure 1 depicts the study design and sampling for this study.

Definition of variables

Preexisting cardiovascular disease was defined as a physician's diagnosis of prior myocardial infarction, stroke, or heart failure. Participants provided information on general health status, educational level, and smoking history by questionnaire. The frequency, duration, and intensity of physical activity were reported, and individuals with <1 hour of light exercise per week were classified as inactive. Diabetes was defined by physician diagnosis or glucose >200 mg/dl. Blood pressure was measured 3 times after ≥5 minutes of rest, and the mean of the second and third measurement was reported. Systolic blood pressure >140 mm Hg, diastolic blood pressure >90 mm Hg, or use of antihypertensive medications defined hypertension. Glomerular filtration rate was estimated (eGFR) using the Chronic Kidney Disease–Epidemiology Collaboration (CKD-EPI) equation based on standardized serum creatinine values.[23] Urine albumin was measured with the immuno-turbidimetric method, and indexed to urine creatinine (albumin–creatinine ratio).[24] Frailty was not formally assessed at the HUNT examinations, so we used a self-reported 4-level general health score as a proxy ("How is your general health lately?"—Poor/Not so good/Good/Very good).[25] We defined those reporting "Poor" general health as frail. ESRD was defined as initiation of dialysis or receiving a kidney transplant; treatment options were only those provided by the government-funded health care system requiring mandatory reporting to the Norwegian Renal Registry.

The proportion of missing data in HUNT was very low for most variables included in this study (1%–2%), and moderate for self-reported lifestyle variables (smoking 7% and physical activity 17%). Data on prevalent heart failure were not obtained in HUNT-2, and albumin–creatinine ratio measurements in 3 urine samples were done only in patients with diabetes or hypertension, or a 5% random subsample, that is, these data were missing by design for 35% and 85% of participants. We used multiple imputation with chained equations, creating 20 datasets of the HUNT database, a technique well suited for imputation of such large proportions of missing data because our data were "missing at random" and we had a very high number of complete cases available for analysis (>5–10,000 for all variables).[18,19,26,27] Missing data were predicted with relevant multivariate regression analyses using available information from individuals with non-missing data. Datasets were then combined according to Rubin's rules and used in standard statistical analysis.

Prediction of death and ESRD

We predicted 5-year risk for ESRD and death using the KFRE and MREK equations, respectively.[4,5] The KFRE was initially developed and validated in 2 Canadian CKD clinic cohorts, but it was later validated in 750,000 patients with eGFR <60 ml/min per 1.73 m^2 from more than 30 countries (including Norway) and with similar results in subgroups by age, race, and diabetes status.[15] The MREK was developed in a subset of 828 participants in the Cardiovascular Health Study in the US who were selected based on being aged >65 years and having an eGFR of <60 ml/min per 1.73 m^2. The MREK was externally validated in 789 US individuals from the Health, Aging, and Body Composition study who had identical age and eGFR inclusion criteria.[28] In HUNT, we substituted "White" for race in the mortality prediction model for all participants, and we used the KFRE coefficients for "Non-American."

Statistical analysis

We used mean (1 SD) and percentages for descriptive analysis. We identified HUNT participants who experienced ESRD before death and compared them to participants who died before ESRD, as well as to those not experiencing either end point over 5 years. Each equation was individually validated using receiver operating curve analysis (C-statistics) for discrimination. Calibration curves evaluated observed rates across 5 categories of predicted risk (0%–4%, 5%–14%, 15%–24%, 25%–49%, 50%–100%) to visually depict calibration.[15] We also used logistic regression and polynomial functions where intercept and beta-coefficients can be used to test if calibration deviates from perfect calibration.[29,30] We used partial least squares–discriminant analysis to sharpen separation between groups and to understand which variables carried most of the class-separating information.[31] In line with recent recommendations on

reporting prediction models,[32] we used DCA to evaluate the clinical impact of various early nephrologist referral algorithms using the KFRE and MREK, as well as other suggested criteria.[33,34] DCA is a statistical method that evaluates whether a referral algorithm is useful in supporting clinical decisions, and which of the models leads to the best decisions. Early referral implies extra testing and treatment to prepare for dialysis or transplantation. The likely alternative is continued care in general practice and referral when there is a clear indication to start dialysis. DCA describes the relationship between disease prevalence, predictive characteristics of the test (e.g., a referral algorithm), and costs and benefits of the intervention. Only the effects of false-positive and false-negative results are evaluated, as costs of true decisions are assumed to be null (i.e., a "misclassification-cost" term).[35] Net Benefit (clinical utility) = (true positives/N) – (false positives/N) × (Pt/1-Pt), where Pt (probability threshold) is the level of diagnostic certainty above which the patients would choose to have the intervention. The (Pt/1-Pt) factor is equivalent to the harm:benefit ratio, that is, a factor used to incorporate the patient's perception of harms and benefits associated with early referral to a nephrologist. No specific harm:benefit ratio is assumed for individual patients, rather the DCA visulalizes the clinical utility of all referral algorithms of the full range of harm:benefit ratios, because different individuals have different risk thresholds.

Statistical analyses were performed using Stata 14.0 (Stata Corp., College Station, TX). All participants gave informed consent, including linkage to central national registries. The HUNT study was approved by the Regional Committee for Medical Research Ethics, the Norwegian Data Inspectorate, and the Ministry of Health.

DISCLOSURE

All the authors declared no competing interests.

ACKNOWLEDGMENT

The HUNT Study is a collaboration between the HUNT Research Center, Faculty of Medicine, Norwegian University of Science and Technology; The Norwegian Institute of Public Health; the Nord-Trøndelag County Council; and the Central Norway Regional Health Authority. The authors thank the health service and people of Nord-Trøndelag for their endurance and participation.

SUPPLEMENTARY MATERIAL

Figure S1. Kaplan-Meier plots displaying end-stage renal disease (ERSD)-free survival and overall survival in elderly chronic kidney disease patients. CI, confidence interval.

Figure S2. Partial least squares–discriminant analysis (PLS-DA) used to pick the most important variables for discriminating among outcomes. PLS-DA is a versatile algorithm that can be used for predictive and descriptive modelling, as well as discriminative variable selection, and it is especially useful for models with many predictors and with multi-collinearity. The PLS-DA decomposes both the X and Y data (i.e., a "supervised" method) into a set of scores and loadings in order to maximize the correlation between the scores for both the X and Y variables (i.e., not selected based on the direction of maximum variation as in principal components analysis). The analysis provides several statistics, such as loading weight, variable importance on projection (VIP), and a regression coefficient, which can be used to choose the most important discriminatory variables. PLS-DA components also enable a good graphical representation of the partition. ACR, albumin–creatinine ratio; BMI, body mass index; BP, blood pressure; DBP, diastolic blood pressure; eGFR, estimated glomerular filtration rate; Gen., general; SBP, systolic blood pressure.

Figure S3. Decision curve analysis comparing the major referral algorithms from the Results section (colored lines) with referral when 5-year end-stage renal disease (ESRD) risk is higher than a fraction (0.10, 0.25, 0.33, and 0.50) of the mortality risk (black lines). eGFR, estimated glomerular filtration rate.

Figure S4. Decision curve analysis comparing the major referral algorithms from the Results section (colored lines) with referral when 5-year end-stage renal disease (ESRD) risk is higher than a specific cutoff (5%, 10%, 20%, or 50%; black lines). eGFR, estimated glomerular filtration rate.

Figure S5. Fraction by first event (none [0; green], end-stage renal disease [ESRD; 1; red], or death [2; blue]) versus fractions by age categories (agekat; **A**), and eGFR categories (gfrkat; **B**) in a population with 3-fold higher ESRD risk compared with Norway.

Figure S6. Decision curve analysis for different referral algorithms in a population with 3-fold higher end-stage renal disease (ESRD) risk compared with Norway. eGFR, estimated glomerular filtration rate; p, probability.

Table S1. Baseline characteristics including predicted risk in patients by first-experienced future outcomes.

Table S2. Information on population mortality and morbidity, health care availability, and organization, tradition for when to start/offer renal replacement therapy in Norway versus other European countries and the US.

Supplementary Excel Calculator. A calculator for both the KFRE and MREK.

Supplementary material is linked to the online version of the paper at www.kidney-international.org.

REFERENCES

1. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. *Kidney Int Suppl*. 2013;3:1–150.
2. O'Hare AM, Choi AI, Bertenthal D, et al. Age affects outcomes in chronic kidney disease. *J Am Soc Nephrol*. 2007;18:2758–2765.
3. Farrington K, Covic A, Aucella F, et al. Clinical practice guideline on management of older patients with chronic kidney disease stage 3b or higher (eGFR <45 mL/min/1.73 m^2). *Nephrol Dial Transplant*. 2016;31(suppl_2):ii1–ii66.
4. Tangri N, Stevens LA, Griffith J, et al. A predictive model for progression of chronic kidney disease to kidney failure. *JAMA*. 2011;305:1553–1559.
5. Bansal N, Katz R, De Boer IH, et al. Development and validation of a model to predict 5-year risk of death without ESRD among older adults with CKD. *Clin J Am Soc Nephrol*. 2015;10:363–371.
6. Aguayo GA, Donneau AF, Vaillant MT, et al. Agreement between 35 published frailty scores in the general population. *Am J Epidemiol*. 2017;186:420–434.
7. Lerner B, Desrochers S, Tangri N. Risk prediction models in CKD. *Semin Nephrol*. 2017;37:144–150.
8. Ladin K, Lin N, Hahn E, et al. Engagement in decision-making and patient satisfaction: a qualitative study of older patients' perceptions of dialysis initiation and modality decisions. *Nephrol Dial Transplantat*. 2017;32:1394–1401.
9. Lim CE, Ng RW, Cheng NC, et al. Advance care planning for haemodialysis patients. *Cochrane Database Syst Rev*. 2016;7:CD010737.
10. American Geriatrics Society Expert Panel on the Care of Older Adults with Multimorbidity. Patient-centered care for older adults with multiple chronic conditions: a stepwise approach from the American Geriatrics Society. *J Am Geriatr Soc*. 2012;60:1957–1968.
11. Casey JR, Hanson CS, Winkelmayer WC, et al. Patients' perspectives on hemodialysis vascular access: a systematic review of qualitative studies. *Am J Kidney Dis*. 2014;64:937–953.
12. Roberti J, Cummings A, Myall M, et al. Work of being an adult patient with chronic kidney disease: a systematic review of qualitative studies. *BMJ Open*. 2018;8, e023507.
13. Ramer SJ, McCall NN, Robinson-Cohen C, et al. Health outcome priorities of older adults with advanced CKD and concordance with their nephrology providers' perceptions. *J Am Soc Nephrol*. 2018;29:2870–2878.

14. Grams ME, Sang Y, Ballew SH, et al., for the Chronic Kidney Disease Prognosis Consortium. Predicting timing of clinical outcomes in patients with chronic kidney disease and severely decreased glomerular filtration rate [published correction appears in *Kidney Int*. 2018;94:1025–1026]. *Kidney Int*. 2018;93:1442–1451.
15. 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;315:164–174.
16. Blinman P, King M, Norman R, et al. Preferences for cancer treatments: an overview of methods and applications in oncology. *Ann Oncol*. 2012;23:1104–1110.
17. Little RJ, D'Agostino R, Cohen ML, et al. The prevention and treatment of missing data in clinical trials. *N Engl J Med*. 2012;367:1355–1360.
18. Newgard CD, Haukoos JS. Advanced statistics: missing data in clinical research—part 2: multiple imputation. *Acad Emerg Med*. 2007;14:669–678.
19. Rubin DB. *Multiple Imputations for Nonresponse in Surveys*. New York, NY: Wiley; 1987.
20. Norwegian Renal Registry. Annual report 2017. Available at: http://nephro.no/nnr/AARSRAPPORT_NNR_2017.pdf. Accessed February 10, 2019.
21. Krokstad S, Langhammer A, Hveem K, et al. Cohort profile: the HUNT Study, Norway. *Int J Epidemiol*. 2013;42:968–977.
22. Pedersen AG, Ellingsen CL. Data quality in the Causes of Death Registry. *Tidsskr Nor Legeforen*. 2015;135:768–770.
23. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med*. 2009;150:604–612.
24. Hallan SI, Ovrehus MA, Romundstad S, et al. Long-term trends in the prevalence of chronic kidney disease and the influence of cardiovascular risk factors in Norway. *Kidney Int*. 2016;90:665–673.
25. Gonzalez-Pichardo AM, Navarrete-Reyes AP, Adame-Encarnacion H, et al. Association between self-reported health status and frailty in community-dwelling elderly. *J Frailty Aging*. 2014;3:104–108.
26. Siew ED, Peterson JF, Eden SK, et al. Use of multiple imputation method to improve estimation of missing baseline serum creatinine in acute kidney injury research. *Clin J Am Soc Nephrol*. 2013;8:10–18.
27. Marshall A, Altman DG, Royston P, Holder RL. Comparison of techniques for handling missing covariate data within prognostic modelling studies: a simulation study. *BMC Med Res Methodol*. 2010;10:7.
28. Harris TB, Visser M, Everhart J, et al. Waist circumference and sagittal diameter reflect total body fat better than visceral fat in older men and women: The Health, Aging and Body Composition Study. *Ann N Y Acad Sci*. 2000;904:462–473.
29. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J*. 2014;35:1925–1931.
30. Finazzi S, Poole D, Luciani D, et al. Calibration belt for quality-of-care assessment based on dichotomous outcomes. *PLoS One*. 2011;6, e16110.
31. Lee LC, Liong CY, Jemain AA. Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps. *Analyst*. 2018;143:3526–3539.
32. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21:128–138.
33. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006;26:565–574.
34. Fitzgerald M, Saville BR, Lewis RJ. Decision curve analysis. *JAMA*. 2015;313:409–410.
35. McNeil BJ, Keller E, Adelstein SJ. Primer on certain elements of medical decision making. *N Engl J Med*. 1975;293:211–215.