KI REPORTS
KIReports.org
CLINICAL RESEARCH

One- and 2-Year Mortality Prediction for Patients Starting Chronic Dialysis

Mikko Haapio[1], Jaakko Helve[1], Carola Grönhagen-Riska[2] and Patrik Finne[1,2]

[1]Nephrology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland; and [2]Finnish Registry for Kidney Diseases, Helsinki, Finland

Introduction: Mortality risk of patients with end-stage renal disease (ESRD) is highly elevated. Methods to estimate individual mortality risk are needed to provide individualized care and manage expanding ESRD populations. Many mortality prediction models exist but have shown deficiencies in model development (data comprehensiveness, validation) and in practicality. Therefore, our aim was to design 2 easy-to-apply prediction models for 1- and 2-year all-cause mortality in patients starting long-term renal replacement therapy (RRT).

Methods: We used data from the Finnish Registry for Kidney Diseases with complete national coverage of RRT patients. Model training group included all incident adult patients who started long-term dialysis in Finland in 2000 to 2008 (n = 4335). The external validation cohort consisted of those who entered dialysis in 2009 to 2012 (n = 1768). Logistic regression with stepwise variable selection was used for model building.

Results: We developed 2 prognostic models, both of which only included 6 to 7 variables (age at RRT start, ESRD diagnosis, albumin, phosphorus, C-reactive protein, heart failure, and peripheral vascular disease) and showed sufficient discrimination (c-statistic 0.77 and 0.74 for 1- and 2-year mortality, respectively). Due to a significantly lower mortality in the newer cohort, the models, to a degree, overestimated mortality risk.

Discussion: Mortality prediction algorithms could be more widely implemented into management of ESRD patients. The presented models are practical with only a limited number of variables and fairly good performance.

Kidney Int Rep (2017) **2,** 1176–1185; http://dx.doi.org/10.1016/j.ekir.2017.06.019
KEYWORDS: algorithm; end-stage renal disease; mortality; prediction; registry; risk
© 2017 International Society of Nephrology. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Patients with end-stage renal disease (ESRD) and on dialysis have significantly higher mortality compared with the non-ESRD population of same age.[1,2] Although we have seen progress in many areas of nephrology and in survival of some dialysis subpopulations,[3] and even in dialysis patients as a whole,[4,5] the prognosis of long-term dialysis patients is still pessimistic. At the same time, with a large number of patients starting long-term dialysis and with less kidney transplantations performed than anticipated, there has been a need for overall expansion of dialysis programs. This has caused growing demands on nephrology services and raised the question about which patients would benefit the most (or least) from long-term renal replacement therapy (RRT).[4–10] The most objective marker of benefit is survival time on RRT. Risk of death after RRT begins is mainly affected by factors such as age, comorbidities, and overall health status,[11–18] which are usually known factors. However, prediction of mortality risk for an individual ESRD patient is elusive.

Attempts to calculate patient-level risk have been made by constructing risk prediction algorithms or comorbidity indexes based on registry data derived from various ESRD patient populations.[19–28] At their best, risk algorithms could offer tools to guide individualized decision-making and sound management and use of RRT resources.[21] Unfortunately, previous algorithms have often shown deficiencies in their quality and in comprehensiveness of background data.[29,30]

The objective of this study was to develop 2 mortality prediction models, in the form of mathematical algorithms, by using the extensive data of the Finnish Registry for Kidney Diseases. The prediction models

Correspondence: Mikko Haapio, Helsinki University Hospital, P.O. Box 372, FI-00029 HUS, Finland. E-mail: mikko.haapio@hus.fi
Received 3 November 2016; revised 24 May 2017; accepted 20 June 2017; published online 26 June 2017

could then be used in everyday clinical work for determining patient-specific, all-cause mortality risk within 1 and 2 years from RRT start.

SUBJECTS AND METHODS

Data Source

All data were retrieved from the Finnish Registry for Kidney Diseases, which has 97% to 99% of records for all Finnish patients on long-term RRT since 1965.[31] This registry is maintained by the Finnish Kidney and Liver Association and financed by the Finnish government. All patients provide written informed consent and permission to use the data anonymously in registry reports and for research purposes. All Finnish dialysis units provide the Registry with specific information on patients starting long-term RRT. Information includes data on demographics, comorbidities, medications, and results of defined measurements and laboratory tests. Diagnoses are reported as *International Classification of Diseases-10th Revision* codes.

Study Population

To construct the 1- and 2-year all-cause mortality prediction algorithms, we used a training group that consisted of all incident patients 18 years or older who started long-term RRT (hemodialysis, peritoneal dialysis, and consequent kidney transplantation) from January 1, 2000 to December 31, 2008 (n = 4341) in Finland. Six patients were excluded due to pre-emptive kidney transplantation, the remaining 4335 patients were included in the final mortality prediction training group. Patients were followed from the first day of dialysis treatment until: (i) death within 1 (n = 597) or 2 (n = 1080) years, (ii) recovery of kidney function >3 months from RRT start (n = 63 at 1 year and n = 81 at 2 years), (iii) moving abroad (1 both at 1 and 2 years), (iv) loss to follow-up (none at 1 year and 1 at 2 years), or (v) end of the follow-up period at 1 (n = 3674) or 2 years (n = 3172) after start of dialysis. Patients who regained their kidney function in <3 months from dialysis start were excluded from the analyses because they were not considered long-term dialysis patients. However, long-term dialysis patients who died within 3 months were included. Patients were not censored at time of transplantation.

The validation cohort consisted of all incident patients 18 years or older who started long-term dialysis in Finland during a later time period, from January 1, 2009 to December 31, 2012 (n = 1770). Two patients with pre-emptive kidney transplantation were excluded, thus leaving 1768 patients in the validation group. Survival of these patients was followed until December 31, 2013. Therefore, follow-up of 1-year survival was possible for all the patients in the validation group, but the complete follow-up of 2-year survival was limited to patients who started dialysis before 2012 (n = 1341).

Outcome and Candidate Predictor Variables

Outcome was all-cause mortality within 1 or 2 years from RRT start. Based on literature, clinical importance, and availability in the Registry data, we originally selected 32 candidate variables that were tested for survival prognostication (Table 1).

In the training group, 0% to 23% of the laboratory values and 5% to 33% of the comorbidity data were missing. Of the final 1- and 2-year prediction model variables, data were unavailable as follows: age at RRT start 0%, ESRD diagnosis 0%, albumin 4%, phosphorus 3%, C-reactive protein 20%, heart failure 9%, peripheral vascular disease 6%, and peripheral vascular disease with limb amputation 7%. The percentage of missing data for all the original variables used in model construction is shown in Table 2.

Statistical Methods

We constructed the 2 mortality prediction algorithms using only the training group to maximize discrimination and calibration, and the final models were selected based only on the results in the training group. The validation group was used exclusively for validation of the final models. Prediction algorithms were developed using multivariable logistic regression with the binary outcome of death or not within 1 or 2 years from start of RRT. We used the Hosmer-Lemeshow (goodness-of-fit) test to assess calibration and the c-statistic (area under the receiver-operating characteristic curve [AUC]) to evaluate discrimination. Calibration of the predictive model was also assessed graphically.

To detect marked nonlinearity between continuous predictors and the outcome, we categorized continuous variables into 3 to 6 groups and modeled the categorical variables in univariable logistic regression analysis. If nonlinearity was observed, logarithmic transformation of the predictor was evaluated and compared with no transformation and the categorical variable. We calculated predicted probabilities and constructed graphs for continuous variables against probability of 1-year mortality and chose the best fitting transformation (either linear, logarithmic, or group variable) according to −2 log likelihood and the Hosmer-Lemeshow test. The predicted probabilities were calculated with the following equation:

$$\text{Predicted probabilities} = 1/\left(1 + e^{-\text{logit}}\right) \quad (1)$$

Table 1. The 15 candidate predictor variables for the development of 1-year risk prediction algorithm (training group)

	Preliminary analysis		
Significant ($P < 0.05$) variables after step-by-step regression analysis with 32 variables	Multivariable model OR (95% CI)	*P* value	Final model ($P < 0.001$ for all) Multivariable model OR (95% CI)
Age at RRT start (yr)	1.05 (1.04–1.06)	<0.001	1.05 (1.04–1.06)
Body mass index (kg/m^2)	—	0.003	—
<18.5	1.85 (1.07–3.19)	0.027	—
18.5–< 25	1 (reference)	—	—
≥25	0.82 (0.66–1.01)	0.059	—
ESRD diagnosis	—	<0.001	—
Glomerulonephritis	1 (reference)	—	1 (reference)
Polycystic disease	0.53 (0.27–1.07)	0.075	0.56 (0.28–1.11)
Diabetes type 1	2.29 (1.45–3.64)	<0.001	2.16 (1.37–3.39)
Diabetes type 2	1.72 (1.17–2.51)	0.005	1.63 (1.13–2.36)
Pyelonephritis	1.08 (0.56–2.09)	0.814	1.14 (0.60–2.19)
Amyloidosis	2.50 (1.57–3.98)	<0.001	3.10 (1.98–4.87)
Nephrosclerosis	1.36 (0.84–2.22)	0.216	1.48 (0.91–2.40)
Other	2.07 (1.40–3.06)	< 0.001	2.38 (1.63–3.49)
Unknown	1.29 (0.87–1.92)	0.209	1.49 (1.00–2.20)
Blood hemoglobin (g/l)	0.991 (0.984–0.997)	0.006	—
Serum creatinine (μmol/l)	0.999 (0.999–1.000)	0.006	—
Serum albumin (g/l)	0.966 (0.951–0.981)	< 0.001	0.959 (0.945–0.973)
Serum phosphorus (mmol/l)	—	< 0.001	—
<1.53	1 (reference)	—	1 (reference)
1.53–<2.0	0.84 (0.65–1.09)	0.190	0.75 (0.59–0.97)
≥2.0	1.37 (1.06–1.77)	0.016	1.15 (0.92–1.45)
Serum C-reactive protein, logarithmic	1.15 (1.07–1.24)	< 0.001	1.16 (1.07–1.24)
Systolic blood pressure, logarithmic	0.31 (0.15–0.64)	0.002	—
Diastolic blood pressure	1.01 (1.00–1.02)	0.006	—
Left ventricular hypertrophy	1.37 (1.11–1.68)	0.004	—
Heart failure	1.79 (1.38–2.34)	< 0.001	2.10 (1.65–2.69)
Peripheral vascular disease	1.63 (1.27–2.09)	< 0.001	1.66 (1.30–2.11)
Present or previous cancer	1.48 (1.14–1.92)	0.003	—
Medication for hypertension	0.73 (0.55–0.95)	0.021	—

CI, confidence interval; ESRD, end-stage renal disease; OR, odds ratio; RRT, renal replacement therapy.
Regression equation constant: 1.337 (preliminary analysis), −4.624 (final model)
The 32 original variables tested: age at RRT start, sex, body mass index, ESRD diagnosis, initial dialysis modality (hemodialysis or peritoneal dialysis), laboratory test variables (each separately: blood hemoglobin, serum creatinine, serum albumin, serum ionized calcium, serum urea, serum phosphorus, serum total cholesterol, serum high-density lipoprotein cholesterol, serum triglycerides, serum C-reactive protein), systolic blood pressure, diastolic blood pressure, comorbidities (each separately: angina pectoris, myocardial infarction, ischemic heart disease with coronary artery bypass grafting, left ventricular hypertrophy, heart failure, peripheral vascular disease, peripheral vascular disease with surgical operation, peripheral vascular disease with limb amputation, stroke, present, or previous cancer), medication for hypertension, lipid-lowering medication, lipid-lowering diet, smoking status (both separately: ex-smoker, present smoker).

where e is the base of the natural logarithm 2.71828, and logit is defined as:

$$\text{logit} = \beta 0 + \beta 1\chi 1 + \beta 2\chi 2 + \beta m\chi m \quad (2)$$

where $\beta 0$ is the constant of the logistic regression equation and $\beta 1$ to βm represent regression coefficients of the variables $\chi 1$ to χm. The regression coefficients can be calculated by taking the natural logarithm of the odds ratios presented in Tables 1 and 3. The equation can be used to calculate predicted probability of death within 1 year (or 2 years) for each patient. The regression equation constants are shown in Tables 1 and 3.

The data set was complete with regard to the outcome and the explanatory variables of age, sex, ESRD diagnosis, and dialysis modality, whereas data were missing to varying degrees (0%–33%) for the other variables. To be able to include all patients in the initial modeling with 32 explanatory variables, we performed multiple imputation in which the explanatory variables were used to impute missing values. The multivariable model was pooled from 5 imputed data sets. Missing data were not imputed for in the validation group. Two-sided *P* values <0.05 were considered statistically significant, and *P* values <0.001 were considered highly significant.

Forward and backward stepwise procedures were used to make a selection from the original 32 variables (Table 1). The 18 variables in the 1-year model and the 23 variables in the 2-year model were significant in either the forward or backward procedure. These variables were entered into logistic regression without a stepwise procedure, and only variables with a *P* value ≤0.05 were taken to the next step (leaving 15 variables in the 1-year model and 13 in the 2-year model) (Tables 1 and 3). These variables were again

Table 2. Percentage of patients with missing data for the variables originally used for model construction

Variable	Percentage missing within training group ($n = 4335$)
Age at RRT start	None
Sex	None
Body mass index	6.2
ESRD diagnosis	None
Initial dialysis modality	None
Blood hemoglobin	4.0
Serum creatinine	1.4
Serum albumin	4.4
Serum ionized calcium	5.1
Serum urea	1.8
Serum phosphorus	2.9
Serum total cholesterol	20.3
Serum HDL cholesterol	22.8
Serum triglycerides	22.2
Serum C-reactive protein	20.3
Systolic blood pressure	3.4
Diastolic blood pressure	3.5
Angina pectoris	6.7
Myocardial infarction	5.8
Ischemic heart disease with coronary artery bypass grafting	6.0
Left ventricular hypertrophy	14.6
Heart failure	8.9
Peripheral vascular disease	6.3
Peripheral vascular disease with surgical operation	6.8
Peripheral vascular disease with limb amputation	6.9
Stroke	5.8
Present or previous cancer	[a]
Medication for hypertension	5.2
Lipid-lowering medication	11.0
Lipid-lowering diet	33.1
Ex-smoker	22.8
Present smoker	18.8

ESRD, end-stage renal disease; HDL, high-density lipoprotein; RRT, renal replacement therapy.
[a]A total of 10.9% of patients in the training group had cancer at some point (and 89.1% did not, or the data were lacking).

assessed in a logistic regression model, and based on the order of statistical significance in these preliminary analyses, variables were further evaluated in 4 to 15 variable combinations that were compared for predictive performance (discrimination and calibration) within the training group. We also tested calibration of the predictive model graphically. Our aim was to build 2 comparably well-performing models that consisted of reasonably few variables to expand the practicality of the models. When selecting variables, we also considered the clinical importance and availability of the variables. The final predictive models contained only highly significant ($P < 0.001$) variables: 7 variables in the 1-year model and 6 in the 2-year model (Tables 1 and 3). To estimate whether the effect of a predictor was dependent on values of another predictor, we added, one at a time, all first-degree interactions (21 in the 7-variable model and 15 in the 6-variable model). Interactions with a P value <0.01 were considered significant and evaluated for clinical performance in the training group.

Validation of the final models was done by applying the final models to the patient cohort that started long-term RRT in 2009 to 2012 (2009–2011 for the 2-year model).

For statistical analyses, we used PASW Statistics 18 and 20 (IBM Inc., Armonk, New York, USA).

RESULTS

Study Populations

There were some significant differences between training and validation groups (Table 4). Patients in the validation group were older, more often men, and had a higher body mass index. The groups also differed with regard to several laboratory test results; the most significant results are shown in Table 4. Although statistically significant, the differences in laboratory values between groups were minor. The groups also differed in the etiologies of ESRD. Pyelonephritis and amyloidosis (usually secondary to rheumatoid arthritis) are decreasing as causes of ESRD, whereas ESRD from nephrosclerosis or from an unknown cause is becoming more common mainly because of the increasing age of those who start dialysis. Frequency of most comorbidities was rather similar between the patient groups, but angina pectoris, left ventricular hypertrophy, and cancer were more frequent in the more recent patients.

Mortality rates were significantly higher in the training group with 597 (13.8%; $P = 0.006$) and 1080 (24.9%; $P = 0.005$) deaths within 1 and 2 years from start of RRT, respectively, compared with 197 (11.1%) and 283 (21.2%) deaths in the validation group (Table 5). However, the proportion of patients who received a kidney transplant decreased in the more recent cohort. Few patients recovered kidney function after 3 months on RRT. The number of patients lost to follow-up was low (Table 5).

Predictors of Mortality

Our preliminary analyses resulted in 15 (1-year model) and 13 (2-year model) variables with P values <0.05 (Table 1 and 3). As explained in the Concise Methods section, we tested several combinations of these variables to find a clinically applicable model with sufficient predictive capability. The combination with the fewest variables tested (age at RRT start, ESRD diagnosis, and serum albumin) showed clearly minor predictive performance compared with combinations with a larger number of variables. Finally, we chose combinations with 7 and 6 variables (1- and 2-year final

Table 3. The 13 candidate predictor variables for the development of 2-year risk prediction algorithm (training group)

Significant ($P < 0.05$) variables after step-by-step regression analysis with 32 variables	Preliminary analysis		Final model ($P < 0.001$ for all)
	Multivariable model OR (95% CI)	*P* value	Multivariable model OR (95% CI)
Age at RRT start (yr)	1.05 (1.04–1.06)	<0.001	1.06 (1.05–1.06)
Body mass index (kg/m^2)	—	0.002	—
<18.5	1.39 (0.83–2.34)	0.210	—
18.5–<25	1 (reference)	—	—
≥25	0.77 (0.65–0.93)	0.005	—
ESRD diagnosis	—	< 0.001	—
Glomerulonephritis	1 (reference)	—	1 (reference)
Polycystic disease	0.75 (0.47–1.21)	0.242	0.73 (0.46–1.17)
Diabetes type 1	2.58 (1.81–3.69)	<0.001	2.81 (1.98–3.99)
Diabetes type 2	2.14 (1.59–2.89)	<0.001	2.17 (1.62–2.90)
Pyelonephritis	0.84 (0.48–1.47)	0.546	0.79 (0.46–1.38)
Amyloidosis	3.54 (2.41–5.21)	<0.001	3.72 (2.54–5.43)
Nephrosclerosis	1.47 (0.99–2.18)	0.055	1.62 (1.10–2.40)
Other	2.26 (1.64–3.11)	< 0.001	2.32 (1.70–3.17)
Unknown	1.42 (1.03–1.95)	0.031	1.51 (1.10–2.07)
Serum creatinine (μmol/l)	0.999 (0.999–1.000)	0.006	—
Serum albumin (g/l)	0.96 (0.94–0.97)	< 0.001	0.96 (0.94–0.97)
Serum phosphorus (mmol/l)	—	0.001	—
<1.53	1 (reference)	—	—
1.53–<2.0	1.04 (0.85–1.27)	0.712	—
≥2.0	1.42 (1.14–1.76)	0.002	—
Serum C-reactive protein, logarithmic	1.11 (1.05–1.18)	< 0.001	1.11 (1.05–1.18)
Angina pectoris	1.38 (1.13–1.68)	0.001	—
Heart failure	2.12 (1.66–2.71)	< 0.001	2.48 (1.98–3.10)
Peripheral vascular disease with limb amputation	1.79 (1.27–2.51)	< 0.001	1.90 (1.36–2.65)
Stroke	1.30 (1.03–1.65)	0.030	—
Present or previous cancer	1.33 (1.06–1.67)	0.013	—
Present smoker	1.31 (1.03–1.66)	0.030	—

CI, confidence interval; ESRD, end-stage renal disease; OR, odds ratio; RRT, renal replacement therapy.
Regression equation constant: −3.776 (preliminary analysis), −4.073 (final model).
The 32 original variables tested: please see Table 1.

models, respectively) (Tables 1 and 3), based on comparatively good performance in calibration (Hosmer-Lemeshow test) and discrimination (c-statistic) (Table 6).

In our interaction analyses, we found 3 significant first-degree interactions in the 1-year model: between ESRD diagnosis and age at start of RRT, between serum albumin and age at start of RRT, and between ESRD diagnosis and heart failure. We found 1 interaction in the 2-year model: between ESRD diagnosis and heart failure. Addition of these interactions to the models did not improve AUC in the training group, and they were therefore not included in the final predictive models.

Validation of the Model

Predictive ability of the final models was assessed in the validation group, and performance was found to be equal in the 1-year model and only slightly lower in the 2-year model compared with the training group (AUC: 0.768 vs. 0.769 and AUC: 0.764 vs. 0.740, respectively) (Table 6).

Calibration of the models was graphically assessed in the validation group by categorizing predicted probabilities into deciles (i.e., 10 equally sized groups with increasing magnitude of predicted probability) and comparing average predicted probabilities with observed mortality in the deciles (Figure 1). Both the 1- and 2-year models showed suboptimal calibration by overestimating death risks in the validation group, as also indicated by the significant *P* values in Hosmer-Lemeshow test (Table 6). In other words, our models predicted mortality to be higher in the validation group than what was actually seen in reality.

DISCUSSION

We developed prediction algorithms for mortality up to 2 years after start of long-term RRT. We used a less recent patient cohort to create the models, which were applied on a more recent validation population. The models showed comparably good predictive ability with regard to discrimination as assessed with the c-statistic, especially for prediction of mortality during the first year after the start of RRT. However, the models overestimated risk of death because survival had improved in the more recent patient cohort. This distorted calibration of the models. We chose to divide

Table 4. Characteristics of study patients according to data set

Characteristics	Training group ($n = 4335$)	Validation group ($n = 1768$)	*P* value
Age at RRT start (yr); median (IQR)	62.3 (21.2)	64.0 (19)	<0.001
Males, n (%)	2751 (63.5)	1198 (67.8)	0.001
Body mass index (kg/m^2); median (IQR)	25.7 (6.6)	26.6 (7.3)	<0.001
Initial dialysis modality, n (%)			0.685
Hemodialysis	3297 (76.1)	1,336 (75.6)	—
Peritoneal dialysis	1038 (23.9)	432 (24.4)	—
Laboratory measurements, median (IQR)			
eGFR (CKD-EPI) at RRT start (ml/min/1.73 m^2)	7.8 (4.4)	7.4 (4.0)	<0.001
Serum creatinine (μmol/l)	568 (281)	588 (283)	<0.004
Blood hemoglobin (g/l)	107 (21)	106 (20)	<0.005
Serum albumin (g/l)	33.0 (9.0)	32.0 (9.8)	0.001
C-reactive protein (mg/l)	8 (17)	6 (17)	<0.001
Serum phosphorus (mmol/l)	1.79 (0.73)	1.82 (0.73)	0.030
ESRD diagnosis, % within group			<0.001
Glomerulonephritis	13.6	13.6	—
Polycystic disease	8.8	10.1	—
Diabetes type 1	15.1	13.5	—
Diabetes type 2	20.3	21.4	—
Pyelonephritis	3.7	2.3	—
Amyloidosis	4.9	2.7	—
Nephrosclerosis	5.6	6.4	—
Other	13.7	14.4	—
Unknown	14.2	15.6	—
Comorbidity,[a] % within group			
Angina pectoris	20.8	18.0	0.016
Myocardial infarction	15.6	15.0	0.555
Left ventricular hypertrophy	33.2	36.8	0.012
Heart failure	11.1	10.7	0.648
Peripheral vascular disease	12.9	12.8	0.946
Peripheral vascular disease with limb amputation	4.5	4.8	0.559
Stroke	11.0	11.7	0.455
Cancer	11.1	12.3	<0.001

CKD-EPI, chronic kidney disease epidemiology collaboration; eGFR, estimated glomerular filtration rate; ESRD, end-stage renal disease; IQR, interquartile range; RRT, renal replacement therapy.
[a]Of the subjects with data available.

patients into training and validation groups based on year of start of RRT because this realistically reflected how predictive algorithms could be applied in the future. In most of the earlier studies, validation sets were drawn randomly from the entire study population, thus almost inevitably resulting in a larger degree of similarity between the groups.

Life-long treatment is demanding both for individual patients and for the health care system. A baseline risk prediction is an important tool when establishing frames for patient management, such as choice of dialysis modality and whether to aim for kidney transplantation.[7,8] Risk prediction may help care givers and patients to reach the best treatment decisions. In many cases, clinical expertise and experience suffice, but to deliver uniform and justifiable patient care, standardized prediction tools are warranted. Furthermore, efforts to reach equal, fair, and useful sharing of limited health care resources require comparisons of practice patterns and of quality of treatment.

Table 5. Outcome according to data set.

Outcomes	Training group ($n = 4335$) During 1 yr	Training group ($n = 4335$) During 2 yrs	Validation group ($n = 1768$) During 1 yr	Validation group ($n = 1768$) During 2 yrs
Mortality, all patients, n (%)	597 (13.8)	1080 (24.9)	197 (11.1)	283 (21.2)
Recipients of kidney transplant, alive at 1 and at 2 years, n (%)	375 (8.7)	824 (19.0)	104 (5.9)	226 (12.8)
Recovery of kidney function after 3 mo from RRT start, n (%)	63 (1.5)	81 (1.9)	24 (1.4)	25 (1.4)
Loss to follow-up, n	0	1	0	0

RRT, renal replacement therapy.

For the preceding reasons, several prediction tools were constructed by many research teams in recent years. These predictive models varied with regard to their length of prediction time, but most were similar in terms of where the data to construct the models were derived: local or national registries of incident ESRD populations. An example of a short-term model is the one by Couchoud *et al.*, who used the French Renal Epidemiology and Information Network (REIN) registry data from 2005 to 2012.[28] In their study, which had >24,000 patients 75 years or older they built a model to predict 3-month mortality, differentiating patients among low-, intermediate-, and high-risk groups. A substantial number of the patients died (10.5%), 22% of them after dialysis withdrawal. Their model had 9 variables and reached an AUC of 0.76 in internal validation. A model focusing on 1-year mortality was developed by Mauri *et al*. in their study of 5738 incident hemodialysis patients in Catalonia, Spain from 1997 to 2003.[21] These patients were randomly divided into development (60%) and validation (40%) groups. The investigators identified risk factors and built a 10-variable predictive model, which showed adequate discrimination (AUC 0.78). In another study that investigated 1-year mortality, Quinn *et al*. investigated incident Canadian dialysis patients from 1998 to 2005 (n = 16,025).[24] In their study, the researchers used statistically diverse methods and developed both a prognostic index for 1-year mortality and a summary risk score using derivation, validation, and testing

Table 6. The predictive ability of the final models using the c-statistic (area under the curve)

Model	Training group ($n = 4335$) 1 yr	Training group ($n = 4335$) 2 yrs	Validation group ($n = 1768$) 1 yrs	Validation group ($n = 1768$) 2 yrs
No.	4335	4335	1418*	1101[a]
AUC	0.768	0.764	0.769	0.740
Hosmer-Lemeshow test *P* value	0.018	0.069	0.041	0.015

AUC, area under the curve.
[a]The subjects for whom data on all included variables were available.

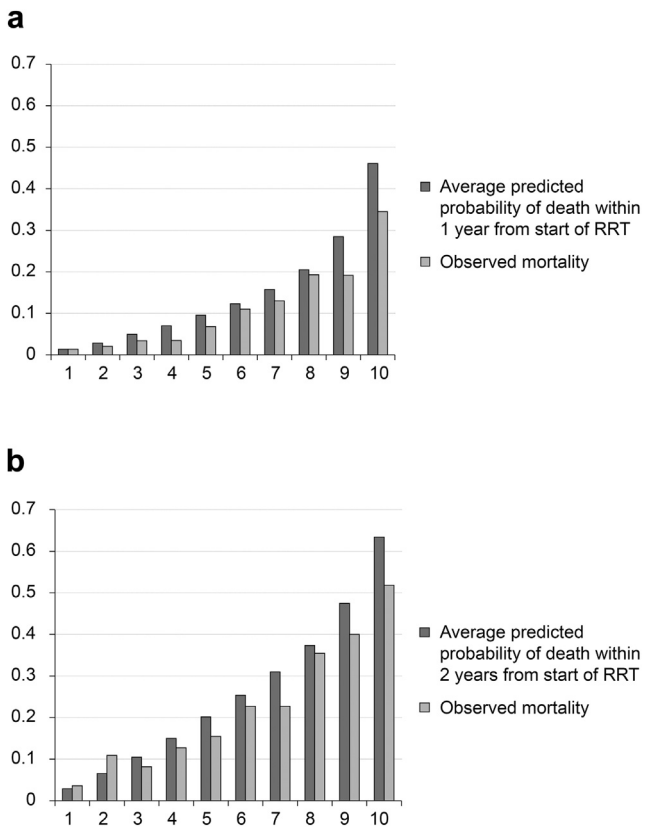

Figure 1. Calibration of the 1-year (a) and the 2-year (b) models. Average predicted probabilities of death (mean, dark gray bars) and observed mortality (proportion of patients who died, light gray bars) according to decile of predicted probability of death (x-axis). RRT, renal replacement therapy.

cohorts. Their final model was large, consisting of 15 variables, and might therefore be considered laborious in everyday clinical work. Of studies that aimed at model construction for >1-year mortality, a study by Liu *et al*. used data of US incident dialysis patients from 2000 for index development, and then validated the index with data from 1999 and 2001 incident and 2000 prevalent dialysis populations.[23] Each of the included comorbid conditions received a numerical weight and the comorbidity score of the patients was the sum of the weights. The number of patients in the cohorts was large (>240,000 in total), with a follow-up of 2.3 to 2.5 years. The performance of the index was better than that of Charlson Comorbidity Index, but the study was criticized for potential weaknesses of data sources and for including only patients who had survived >9 months on RRT. The discrimination ability of the index was rather low (0.67 in the validation cohort). Wagner *et al*. analyzed a nationwide patient cohort that started peritoneal or hemodialysis from 2002 to 2004 in the United Kingdom.[25] Both the training and the validation cohorts were from the same time period, the patients were censored at transplantation, and were followed for 3 years. The model could sufficiently discriminate among 4 patient groups according to level of mortality risk and had a c-statistic of 0.73. However, almost one-half of the incident dialysis patients were excluded because of missing data. Another prognostic model, which looked at long-term survival using Dutch registry data and a cohort from 1995 to 2005 (n = 13,868) was designed by Hemke *et al*., who developed a 10-year model with a c-statistic of 0.72 using internal validation.[26] In comparison to the national studies described previously, the multinational European Renal Association–European Dialysis and Transplant Association (ERA-EDTA) Registry study involved 793 centers from 37 countries and aimed both at developing predictive models for survival and comparing multivariable regression model with another model (''self-learning rule-based model'') and the modified Charlson comorbidity score.[20] As a result, 2 novel instruments were created, with a tendency for better performance in longer (5 year) than shorter (1 year) prediction.

There were many similarities between our study and those described previously, especially statistical methods and outcomes with regard to discrimination ability. Overall, AUC for 3-month to 1-year death prediction was between 0.70 and 0.84 and between 0.67 and 0.75 for 2-year prediction.[21–28,30,32] In most of the earlier studies, an AUC of approximately 0.75 was considered adequate. However, the outcome of most studies in the field might still be considered suboptimal. This was inevitable when using medical registry data; for example, death caused by other than registered comorbid factors, models that did not account for severity of an important comorbid condition, and factors that presented only after start of long-term RRT. Regarding our models, data for most of the included variables (e.g., age, serum albumin, and phosphorus) are usually well documented for patients starting chronic RRT. However, other variables, such as ESRD diagnosis and heart failure, may be less explicit.

Many studies that compared performance of different predictive models reported difficulties in reaching definite conclusions. These difficulties mainly resulted from the use of different statistical methods in building the models. Other problems were associated with models being constructed in different patient populations and within dissimilar health care systems.[20,23,30–33] There are also many other reasons why it will remain difficult to establish exact patient prognosis. The age of ESRD patients starting long-term RRT is increasing, and patients have more comorbid conditions compared with dialysis patients of the past. Both these factors increase risk of death. In contrast, general medical and dialysis treatment advancements are expected to improve the prognosis of future RRT populations. Therefore, performance of prediction

models constructed with previous data may be lower when applied to present or future patients. If real improvement in prognosis occurs, models are likely to overestimate mortality risk in newer populations, which the present study showed.

This study had some limitations to be addressed. First, the amount of missing data was reasonably high for some variables included in our prediction models, especially on C-reactive protein (but for the benefit of our models, <10% for all other variables (Table 2). Therefore, in order not to exclude a large proportion of the training group, multiple imputation for missing data was performed and this might have altered the results of our analyses compared with having analyzed a 100% comprehensive original data. In contrast, had we not imputed for missing data (and thus had excluded patients with missing data), we might have caused selection bias. In the final models with 6 to 7 variables, only 9% of the patients lacked data on ≥1 variables. Furthermore, in the validation group, missing data were not imputed. Second, we did not censor patients at kidney transplantation, and this might have affected our study results. However, censoring at transplantation could have altered the weight of model variables because many of the patients who later received a transplant had more favorable characteristics (e.g., higher serum albumin and fewer comorbidities) at RRT start compared with patients who would never receive a transplant. This approach was also justified because possibility for transplantation and its realization were often unknown at start of RRT. Third, as far as we know, there are no studies that showed the impact of prediction models like ours on patient outcome. It might also be questioned whether the most beneficial timeframe to evaluate mortality was the first year or the first 2 years on RRT, as in our models, or whether prediction should focus on earlier or later mortality.[28] We chose our time intervals mainly because they were used in several earlier studies,[19–21,23,24,27,30] and also due to certain characteristics of the Finnish nephrology care system: a relatively high percentage of ESRD patients are treated conservatively, that is, without entering dialysis, thus favoring less short-term prognostication. However, in other settings, a different temporal approach might be better justified. These uncertainties stress the importance of incorporating predictive models to clinical work in various settings to gather practical experience of the usefulness of models. Fourth, because we used data only derived from Finnish ESRD patients, the universality of our prediction models to non-Finnish ESRD populations might be open to doubt. Therefore, we see it as important to establish international collaboration among research teams to test the performance of national prediction models in foreign ESRD populations. Because data collection might be rather incoherent between different national registries, possibly hindering the bilateral applicability of models, model validation would likely be more reliable comparing models with only a limited number of variables and with a precise definition of the variables.

Our study had several strengths. In particular, the registry database we used was of exceptionally high quality in terms of national health care unit and patient coverage. Data from the registry originated directly from dialysis units and nephrologists treating dialysis patients. Furthermore, these data were prospectively collected for incident cohorts without exclusion. We used national administrative data in combination with patient characteristics and comorbidities, as well as laboratory data. Few patients were excluded. The validation group we used was recent, and the training group was also contemporary. Importantly, using an external and newer validation data set increased the credibility of the performance of our models in circumstances that better mimic the clinical situation. Finally, for a prediction model to be applicable for clinical practice, it should be convenient and as little time-consuming as possible to use. We were able to construct prediction models with only a small number of variables, thus increasing practicality of the models in clinical settings. We would be glad to send the exact equations of our models to those interested in applying them.

To conclude, mortality prediction algorithms might serve as additional determinants of treatment plans for ESRD patients. Testing of models should be encouraged to evaluate their objective benefits. Our 1- and 2-year prediction models showed adequate performance, and due to their small size in terms of variable number, they could be easily applied into clinical work.

DISCLOSURE

All the authors declared no competing interests. The results presented in this paper have not been published previously in whole or part.

ACKNOWLEDGMENTS

We acknowledge the support from the Board of the Finnish Registry for Kidney Diseases, and all the nephrologists and staff in all Finnish central hospitals who have reported to the Finnish Registry for Kidney Diseases: Helsinki University Hospital, Turku University Central Hospital, Satakunta Central Hospital, Kanta-Häme Central Hospital, Tampere University Central Hospital, Päijät-Häme Central Hospital, Kymenlaakso Central Hospital, Etelä-Karjala Central Hospital, Mikkeli Central Hospital, Itä-Savo Central Hospital,

Pohjois-Karjala Central Hospital, Kuopio University Central Hospital, Keski-Suomi Central Hospital, Etelä-Pohjanmaa Central Hospital, Vaasa Central Hospital, Keski-Pohjanmaa Central Hospital, Oulu University Central Hospital, Kainuu Central Hospital, Länsi-Pohja Central Hospital, Lappi Central Hospital, and Åland Central Hospital.

We acknowledge the grammatical help kindly provided by Professor Andrew A House, MD, PhD, MSc, London Health Sciences Centre, London, Ontario, Canada.

MH and JH were supported with grants from the Diabetes Research Foundation and Liv och Hälsa Foundation, and JH was supported by the Finnish Medical Foundation and the Finnish Kidney Foundation. The support consisted of general funding for the research project and monthly salary for the authors. MH received a special governmental subsidy for health sciences research. JH received a grant from Finska Läkaresällskapet. The sponsors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; or preparation, review, and approval of the manuscript.

MH and PF had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis, and responsibility for the decision to submit for publication.

MH acquired data and interpreted the results, drafted and revised the manuscript, and approved the final version. JH and CG-R revised the manuscript and approved the final version. PF conceived and designed the work, acquired data and interpreted the results, revised the manuscript, and approved the final version.

REFERENCES

1. Foley RN, Parfrey PS, Sarnak MJ. Clinical epidemiology of cardiovascular disease in chronic renal disease. *Am J Kidney Dis*. 1998;32:S112–S119.
2. Nordio M, Limido A, Maggiore U, et al. Survival in patients treated by long-term dialysis compared with the general population. *Am J Kidney Dis*. 2012;59:819–828.
3. Haapio M, Helve J, Groop PH, et al. Survival of patients with type 1 diabetes receiving renal replacement therapy in 1980-2007. *Diabetes Care*. 2010;33:1718–1723.
4. Collins AJ, Foley RN, Gilbertson DT, et al. The state of chronic kidney disease, ESRD, and morbidity and mortality in the first year of dialysis. *Clin J Am Soc Nephrol*. 2009;1(4 Suppl): S5–S11.
5. *ERA-EDTA Registry: ERA-EDTA Registry Annual Report 2012*. Amsterdam: Department of Medical Informatics, Academic Medical Center; 2014.
6. Brunori G, Viola BF, Maiorca P, et al. How to manage elderly patients with chronic renal failure: conservative management versus dialysis. *Blood Purif*. 2008;26:36–40.
7. Courtney AE, Maxwell AP. The challenge of doing what is right in renal transplantation: balancing equity and utility. *Nephron Clin Pract*. 2009;111:c62–c67. discussion c68.
8. Dusseux E, Albano L, Fafin C, et al. A simple clinical tool to inform the decision-making process to refer elderly incident dialysis patients for kidney transplant evaluation. *Kidney Int*. 2015;88:121–129.
9. Tonelli M, Wiebe N, Knoll G, et al. Systematic review: kidney transplantation compared with dialysis in clinically relevant outcomes. *Am J Transplant*. 2011;11:2093–2109.
10. O'Connor NR, Kumar P. Conservative management of end-stage renal disease without dialysis: a systematic review. *J Palliat Med*. 2012;15:228–235.
11. Udayaraj U, Tomson CR, Gilg J, et al. UK Renal Registry 11th Annual Report (December 2008): Chapter 6 Comorbidities and current smoking status amongst patients starting renal replacement therapy in England, Wales and Northern Ireland: national and centre-specific analyses. *Nephron Clin Pract*. 2009;111(Suppl 1):c97–c111.
12. Fotheringham J, Jacques RM, Fogarty D, et al. Variation in centre-specific survival in patients starting renal replacement therapy in England is explained by enhanced comorbidity information from hospitalization data. *Nephrol Dial Transplant*. 2014;29:422–430.
13. Helve J, Haapio M, Groop PH, et al. Comorbidities and survival of patients with type 1 diabetes on renal replacement therapy. *Diabetologia*. 2011;54:1663–1669.
14. Helve J, Haapio M, Groop PH, et al. Biochemical variables and survival of patients with type 1 diabetes on renal replacement therapy. *Blood Purif*. 2014;38:18–23.
15. Avram MM, Bonomini LV, Sreedhara R, et al. Predictive value of nutritional markers (albumin, creatinine, cholesterol, and hematocrit) for patients on dialysis for up to 30 years. *Am J Kidney Dis*. 1996;28:910–917.
16. Xue JL, Everson SE, Constantini EG, et al. Peritoneal and hemodialysis: II. Mortality risk associated with initial patient characteristics. *Kidney Int*. 2002;61:741–746.
17. Bradbury BD, Fissell RB, Albert JM, et al. Predictors of early mortality among incident US hemodialysis patients in the Dialysis Outcomes and Practice Patterns Study (DOPPS). *Clin J Am Soc Nephrol*. 2007;2:89–99.
18. van Manen JG, van Dijk PC, Stel VS, et al. Confounding effect of comorbidity in survival studies in patients on renal replacement therapy. *Nephrol Dial Transplant*. 2007;22: 187–195.
19. Merkus MP, Jager KJ, Dekker FW, et al. Predictors of poor outcome in chronic dialysis patients: The Netherlands Cooperative Study on the Adequacy of Dialysis. The NECOSAD Study Group. *Am J Kidney Dis*. 2000;35:69–79.
20. Geddes CC, van Dijk PC, McArthur S, et al. The ERA-EDTA cohort study–comparison of methods to predict survival on renal replacement therapy. *Nephrol Dial Transplant*. 2006;21: 945–956.
21. Mauri JM, Cleries M, Vela E. Catalan Renal Registry. Design and validation of a model to predict early mortality in hemodialysis patients. *Nephrol Dial Transplant*. 2008;23: 1690–1696.
22. Couchoud C, Labeeuw M, Moranne O, et al. A clinical score to predict 6-month prognosis in elderly patients starting dialysis for end-stage renal disease. *Nephrol Dial Transplant*. 2009;24: 1553–1561.

23. Liu J, Huang Z, Gilbertson DT, et al. An improved comorbidity index for outcome analyses among dialysis patients. *Kidney Int*. 2010;77:141–151.
24. Quinn RR, Laupacis A, Hux JE, et al. Predicting the risk of 1-year mortality in incident dialysis patients: accounting for case-mix severity in studies using administrative data. *Med Care*. 2011;49:257–266.
25. Wagner M, Ansell D, Kent DM, et al. Predicting mortality in incident dialysis patients: an analysis of the United Kingdom Renal Registry. *Am J Kidney Dis*. 2011;57:894–902.
26. Hemke AC, Heemskerk MB, van Diepen M, et al. Survival prognosis after the start of a renal replacement therapy in the Netherlands: a retrospective cohort study. *BMC Nephrol*. 2013;14:258.
27. van Diepen M, Schroijen MA, Dekkers OM, et al. Predicting mortality in patients with diabetes starting dialysis. *PLoS One*. 2014;9:e89744.
28. Couchoud CG, Beuscart JB, Aldigier JC, et al, REIN Registry. Development of a risk stratification algorithm to improve patient-centered care and decision making for incident elderly patients with end-stage renal disease. *Kidney Int*. 2015;88:1178–1186.
29. Grams ME, Coresh J. Assessing risk in chronic kidney disease: a methodological review. *Nat Rev Nephrol*. 2013;9:18–25.
30. Miskulin DC, Martin AA, Brown R, et al. Predicting 1 year mortality in an outpatient haemodialysis population: a comparison of comorbidity instruments. *Nephrol Dial Transplant*. 2004;19:413–420.
31. Finnish Registry for Kidney Diseases. *Finnish Registry for Kidney Diseases Report 2013*. Helsinki, Finland: Finnish Registry for Kidney Diseases; 2014.
32. van Manen JG, Korevaar JC, Dekker FW, et al. How to adjust for comorbidity in survival studies in ESRD patients: a comparison of different indices. *Am J Kidney Dis*. 2002;40:82–89.
33. Fried L, Bernardini J, Piraino B. Comparison of the Charlson comorbidity index and the Davies score as a predictor of outcomes in PD patients. *Perit Dial Int*. 2003;23:568–573.