OPEN

Development and validation of a predictive mortality risk score from a European hemodialysis cohort

Jürgen Floege[1], Iain A. Gillespie[2], Florian Kronenberg[3], Stefan D. Anker[4], Ioanna Gioni[5], Sharon Richards[6], Ronald L. Pisoni[7], Bruce M. Robinson[7], Daniele Marcelli[8], Marc Froissart[9], Kai-Uwe Eckardt[10] on behalf of the ARO Steering Committee (collaborators)[11]

[1]Nephrology, RWTH University of Aachen, Aachen, Germany; [2]Center for Observational Research (CfOR), Amgen Ltd, Uxbridge, UK; [3]Division of Genetic Epidemiology, Department of Medical Genetics, Molecular and Clinical Pharmacology, Medical University of Innsbruck, Innsbruck, Austria; [4]Department of Innovative Clinical Trials, University Medical Centre Göttingen, Göttingen, Germany; [5]On behalf of Amgen Ltd, Uxbridge, UK; [6]Global Biostatistics, Amgen Ltd, Uxbridge, UK; [7]Arbor Research Collaborative for Health, Ann Arbor, Michigan, USA; [8]EMEALA Medical Board, Fresenius Medical Care, Bad Homburg, Germany; [9]International Development Nephrology, Amgen Europe GmbH, Zug, Switzerland and [10]Nephrology and Hypertension, University of Erlangen-Nuremberg, Erlangen, Germany

Although mortality risk scores for chronic hemodialysis (HD) patients should have an important role in clinical decision-making, those currently available have limited applicability, robustness, and generalizability. Here we applied a modified Framingham Heart Study approach to derive 1- and 2-year all-cause mortality risk scores using a 11,508 European incident HD patient database (AROii) recruited between 2007 and 2009. This scoring model was validated externally using similar-sized Dialysis Outcomes and Practice Patterns Survey (DOPPS) data. For AROii, the observed 1- and 2-year mortality rates were 13.0 (95% confidence interval (CI; 12.3–13.8)) and 11.2 (10.4–12.1)/ 100 patient years, respectively. Increasing age, low body mass index, history of cardiovascular disease or cancer, and use of a vascular access catheter during baseline were consistent predictors of mortality. Among baseline laboratory markers, hemoglobin, ferritin, C-reactive protein, serum albumin, and creatinine predicted death within 1 and 2 years. When applied to the DOPPS population, the predictive risk score models were highly discriminatory, and generalizability remained high when restricted by incidence/prevalence and geographic location (C-statistics 0.68–0.79). This new model offers improved predictive power over age/comorbidity-based models and also predicted early mortality (C-statistic 0.71). Our new model delivers a robust and reproducible mortality risk score, based on readily available clinical and laboratory data.

Correspondence: Jürgen Floege, Division of Nephrology, RWTH University of Aachen, Pauwelsstrasse 30, 52057 Aachen, Germany. E-mail: juergen.floege@rwth-aachen.de

Chronic kidney disease (CKD), which has evolved as a global health burden,[1] affects up to 13% of United States (US)[2] and European[3] adults, who suffer a high incidence of comorbidities and an increased mortality risk[4]. Mortality rates in end-stage renal disease patients on chronic HD, relating mainly to cardiovascular complications and infections, remains higher than that of many cancers or heart failure, at up to 19.2 per 100 person-years versus only 1.2 in the general European population.[5]

An improved ability to identify those patients at an increased risk of death appears desirable for several reasons. Thus, identification of high-risk patients may help focus efforts on risk mitigation strategies. In addition, a valid, general, easy-to-use mortality risk score in HD patients could also be used in patient discussions or when scheduling transplants. In health-care economics, such a score may categorize patients in comorbidity-adjusted registries or reimbursement systems, and inform planning. Furthermore, it may also serve as a research tool—homogenizing the case mix entering clinical trials and targeting specific interventions to particular patient subgroups—thus reducing sample sizes without compromising statistical power.

Previously developed risk scores lack applicability, robustness, and generalizability. An early study by Wright,[6] which categorized patients as 'low', 'medium', and 'high' risk on the basis of age and comorbidities, was popularized by Khan[7] who examined the predictive power of this stratification (referred subsequently here as the Wright–Khan mortality index). A scoring system based on prediction model β-coefficients advanced methodologies, allowing objective assessment of contributory factors and their weighted impact.[8] Recent large and complex studies[9–15] used internal validation that contributes little to generalizability. Generalizability may be further limited by restricted patient populations,[9,13] geographic locations,[9,11,15] small sample sizes,[11] or insufficient variables.[9,11–14] The current study therefore aimed to develop, in a large European cohort of incident

HD patients, risk scores for 1- and 2-year all-cause mortality and to validate these scores externally in a similarly sized, predominantly prevalent HD population.

RESULTS

Study population

Between 1 January 2007 and 31 December 2009, 11,508 patients were recruited into the second Analyzing Data, Recognizing Excellence and Optimizing Outcomes (ARO) cohort (AROii; Figure 1). Thirty-seven percent of patients initiated HD within Fresenius Medical Care (FME) facilities; nevertheless, the overall median dialysis vintage was only 4 days upon admission. Nonchronic HD patients, those with no laboratory data, and/or those with a history of transplantation (alone or combined; $N = 773$) were excluded. In addition, 1013 patients left the study during baseline, leaving 9722 patients. During the first and second year of follow-up, 1060 (10.9%) and 654 (9.4%) deaths were reported, respectively, giving 1- and 2-year mortality rates of 13.0 (95% CI 12.3–13.8) and 11.2 (95% CI 10.4–12.1) per 100 person-years, respectively. In the first year, 344 (3.5%) patients left the study owing to a renal transplant, and 1338 (13.8%) patients were lost to follow-up (LTFU); in the second year, 288 patients (4.1%) received renal transplants and 600 (8.6%) patients were LTFU. Patients LTFU did not differ greatly from those who were not (Supplementary Table S1 online). Of the 1938 LTFU patients, 527 (27.2%) patients returned to FME after their follow-up stop date. Patients lost or not lost to transplantation are shown in Supplementary Table S2 online.

Table 1 shows baseline characteristics of the study populations. Although AROii and Dialysis Outcomes Practice Patterns III (DOPPS III) patients were similar in many aspects, we noted some differences. The baseline vascular access differences between AROii and the third Dialysis Outcomes Practice Patterns (DOPPS) cohort patients may be explained by the mix of incident and prevalent patients in DOPPS. Additional differences include geography, dialysis

vintage, smoking habits, diabetes, cancer, and cardiovascular disease history. Notably, the proportion of patients dying in each cohort was similar. Within the DOPPS III cohort, mean dialysis vintage differed by 'region' (Europe: 4.1 ± 5.5 years; Japan: 6.9 ± 7.1 years; North America: 3.4 ± 4.1 years; Australasia: 4.5 ± 5.0 years).

Predictors of mortality

In our main AROii analysis (based on a first 3-months on follow-up baseline), increasing age, low body mass index, and a cardiovascular disease or cancer history were independently associated with both 1- and 2-year mortality (Table 2). Former or current smokers were at a greater risk within 2 years but not at 1 year, as were patients with a CKD etiology of diabetic nephropathy or tubulo-interstitial disease. Of the dialysis quality parameters, baseline use of, or change to, vascular access via a catheter was associated with an increased risk for both time periods, as was lower actual blood flow.

Lower hemoglobin concentrations were associated with an increased risk for 1- and 2-year mortality; higher levels were linked with better survival. Baseline inflammation (increased C-reactive protein concentrations and high ferritin levels) was highly predictive of mortality at both 1 and 2 years. Malnutrition and/or inflammation, as evidenced by low concentrations of serum albumin, was also consistently predictive. Predialysis serum creatinine represented an additional risk marker, with lower values associated with higher risk, probably reflecting decreased muscular mass and potentially protein wastage in addition to low serum albumin. Finally, hypercalcemia was associated with a higher 1-year mortality risk.

The results obtained using a 90- to 180-day baseline were remarkably consistent with 0- to 90-day baseline observations, or when LTFU patients were coded as deceased (Supplementary Table S3 online). Of note, the relationship between predialysis serum creatinine and mortality was evident in both analyses, suggesting that any residual renal function at the time of HD initiation in this incident dialysis population could not fully explain this association when a 0- to 90-day baseline was applied.

Risk-score derivation and application

When hazard ratios (HRs) were converted to risk-score points, extreme age had the greatest risk contribution (Table 2). A cancer history was generally more disadvantageous than a cardiovascular disease history. Among laboratory parameters, elevated C-reactive protein concentrations contributed the greatest risk, followed by low albumin and creatinine values. Although lower hemoglobin contributed additive risk, higher hemoglobin values and lower ferritin concentrations contributed most to lowering the risk score.

The risk percentage attributable to risk-score totals differed by follow-up length (Figure 2). The contribution of modifiable risk markers increased as the risk score increased (Supplementary Figure S1 online), but only marginally around 50% of the total risk.
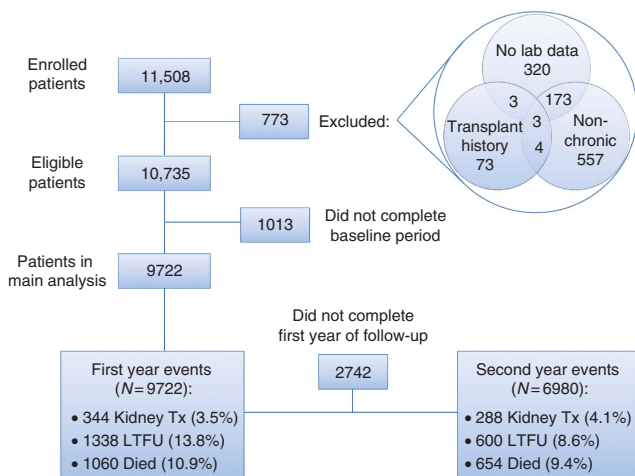

Figure 1 | Derivation of the AROii study population.

Table 1 | Baseline characteristics of the study populations and subpopulations

Parameters	Mª/NMᵇ	AROii (0–3 Mo)ᶜ (N = 9722)	AROii (3–6 Mo)ᵈ (N = 8783)	DOPPS III (0–3 Mo)ᵉ (N = 10,615)
Incident/prevalent on dialysis (%)	NM	100/0	100/0	16/84
Dialysis vintage (months)	NM	0.5 ± 1.1	3.4 ± 1.1	53.9 ± 67.9
Dialysis vintage (months) (median [IQRᶠ])	NM	0.1 [0.0, 0.5]	3.1 [3.0, 3.5]	27.9 [9.3, 70.9]
1-year all-cause deaths	—	1060 (10.9)	910 (10.4)	1048 (9.9)
2-year all-cause deaths	—	1714 (17.6)	1465 (16.7)	1736 (16.4)
Geography:				
Europe	—	9722 (100)	8783 (100)	4994 (47)
Japan	—	0	0	2743 (25.8)
North America	—	0	0	2190 (20.6)
Australasia	—	0	0	688 (6.5)
Age at baseline (years)	NM	64.4 ± 14.7	64.3 ± 14.7	63.4 ± 14.3
*Gender*ᵍ	NM			
Female		3904 (40.2)	3550 (40.4)	4420 (41.6)
Male		5817 (59.8)	5232 (59.6)	6195 (58.4)
*Smoking status*ᵍ	NM			
Nonsmoker		3608 (37.1)	3294 (37.5)	5078 (47.8)
Former		1659 (17.1)	1506 (17.1)	2552 (24.0)
Current		736 (7.6)	685 (7.8)	1505 (14.2)
Missing		3719 (38.3)	3298 (37.5)	1480 (13.9)
Body mass index (kg/m²)ᵍ	M	26.0 ± 7.8	26.0 ± 5.8	25.1 ± 6.1
Missing		683 (7.0)	474 (5.4)	1038 (9.8)
History of diabetes	NM	2740 (28.2)	2536 (28.9)	4189 (39.5)
History of cardiovascular disease	NM	2480 (25.5)	2430 (27.7)	6995 (65.9)
History of cancer	NM	557 (5.7)	504 (5.7)	1347 (12.7)
*Chronic kidney disease etiology*ᵍ	NM			
Hypertension/vascular		1548 (15.9)	1415 (16.1)	1764 (16.6)
Glomerulonephritis		895 (9.2)	811 (9.2)	2415 (22.8)
Diabetes		2335 (24.0)	2136 (24.3)	2901 (27.3)
Tubulo-interstitial		1062 (10.9)	949 (10.8)	968 (9.1)
Polycystic kidney disease		534 (5.5)	490 (5.6)	595 (5.6)
Miscellaneous/other		3106 (31.9)	2793 (31.8)	1220 (11.5)
Invalid/Missing		242 (2.5)	189 (2.2)	752 (7.1)
*Vascular access in the first 90 days*ᵍ	M			
No change: Fistula or graft		3154 (32.4)	3990 (45.4)	7802 (73.5)
No change: Catheter		2908 (29.9)	2062 (23.5)	1637 (15.4)
Change: Fistula/graft to catheter		211 (2.2)	193 (2.2)	99 (0.9)
Change: Catheter to fistula/graft		922 (9.5)	809 (9.2)	355 (3.3)
Other		0	0	158 (1.5)
Missing		2527 (26.0)	1729 (19.7)	564 (5.3)
Actual blood flow (ml/min)ᵍ	M	302.7 ± 131.3	327.1 ± 259.7	294.0 ± 90.0
Missing		920 (9.5)	794 (9.0)	925 (8.7)
Dialysis adequacy (equilibrated Kt/V)ᵍ	M	1.3 ± 0.5	1.4 ± 0.3	1.5 ± 0.3
Missing		718 (7.4)	458 (5.2)	1375 (13.0)
Intradialytic weight change (kg; post-pre)ᵍ	M	− 1.7 ± 0.8	− 1.9 ± 0.8	− 2.2 ± 2.7
Missing		75 (0.8)	36 (0.4)	301 (2.8)
*Hemoglobin (g/l)*ᵍ	M			
<100		2796 (28.8)	935 (10.6)	1248 (11.8)
100 to <120		4966 (51.1)	4015 (45.7)	5473 (51.6)
⩾120		1638 (16.8)	3613 (41.1)	3849 (36.3)
Missing		322 (3.3)	220 (2.5)	45 (0.4)
*Ferritin (μg/l)*ᵍ	M			
<500		3847 (39.6)	2386 (27.2)	3342 (31.5)
⩾500		4822 (49.6)	5383 (61.3)	6629 (62.4)
Missing		1053 (10.8)	1014 (11.5)	644 (6.1)
C-reactive protein (mg/l)ᵍ	M	15.6 ± 25.3	13.3 ± 23.2	13.7 ± 80.6
Missing		2393 (24.6)	2551 (29.0)	4150 (39.1)
*Serum albumin (g/l)*ᵍ	M			
<35		2518 (25.9)	1554 (17.7)	2516 (23.7)
⩾35		5948 (61.2)	5964 (67.9)	7713 (72.7)
Missing		1256 (12.9)	1265 (14.4)	386 (3.6)

Table 1 | (Continued)

Parameters	Mª/NMᵇ	AROii (0–3 Mo)ᶜ (N = 9722)	AROii (3–6 Mo)ᵈ (N = 8783)	DOPPS III (0–3 Mo)ᵉ (N = 10,615)
*Cholesterol (mmol/l)*ᵍ	M			
<3.6		1722 (17.7)	1231 (14.0)	2890 (27.2)
3.6 to <6.0		4589 (47.2)	3463 (39.4)	5385 (50.7)
⩾6.0		532 (5.5)	423 (4.8)	360 (3.4)
Missing		2879 (29.6)	3666 (41.7)	1980 (18.7)
*LDL-cholesterol (mmol/l)*ᵍ	M			
<2.6		3240 (33.3)	2385 (27.2)	3692 (34.8)
2.6 to <3.3		1281 (13.2)	859 (9.8)	789 (7.4)
3.3 to <4.1		653 (6.7)	467 (5.3)	341 (3.2)
4.1 to <4.9		182 (1.9)	143 (1.6)	82 (0.8)
⩾4.9		87 (0.9)	42 (0.5)	23 (0.2)
Missing		4279 (44.0)	4887 (55.6)	5688 (53.6)
*Creatinine (μmol/l)*ᵍ	M	565.4 ± 187.6	614.1 ± 201.7	777.9 ± 256.4
Missing		926 (9.5)	925 (10.5)	52 (0.5)
*Calcium (mmol/l)*ᵍ	M			
<2.1		2742 (28.2)	1884 (21.5)	1576 (14.8)
2.1 to <2.6		6368 (65.5)	6388 (72.7)	8552 (80.6)
⩾2.6		126 (1.3)	133 (1.5)	377 (3.6)
Missing		486 (5.0)	378 (4.3)	110 (1.0)
*Phosphate (mmol/l)*ᵍ	M			
<0.8		291 (3.0)	271 (3.1)	83 (0.8)
0.8 to <1.5		4583 (47.1)	3958 (45.1)	3390 (31.9)
⩾1.5		4527 (46.6)	4335 (49.4)	7064 (66.5)
Missing		321 (3.3)	219 (2.5)	78 (0.7)
*Parathyroid hormone (ng/l)*ᵍ	M			
< 150		2557 (26.3)	2762 (31.4)	3426 (32.3)
150 to <300		2571 (26.4)	2260 (25.7)	2964 (27.9)
300 to <600		1871 (19.2)	1327 (15.1)	2059 (19.4)
⩾600		731 (7.5)	463 (5.3)	946 (8.9)
Missing		1992 (20.5)	1971 (22.4)	1220 (11.5)

Abbreviations: AROii, second Analyzing Data, Recognizing Excellence and Optimizing Outcomes (ARO) cohort; DOPPS, Dialysis Outcomes and Practice Patterns Survey; LDL, low-density lipoprotein.
Categorical variables are reported using *n* (%). Continuous variables are reported using mean ± s.d.
ªFactors considered modifiable.
ᵇFactors considered non-modifiable.
ᶜAROii derivation data set using a 0- to 90-day baseline.
ᵈAROii derivation data set using a 90- to 180-day baseline.
ᵉDOPPS III validation data set using a 0- to 90-day baseline.
ᶠInter quartile range.
ᵍVariables where missing values were imputed.

Internal discrimination and calibration
The distribution of 1- and 2-year risk-score points for patients with and without events is shown in Figure 3, with the intersection point between patients—8 and 9 points, respectively—defining 'high-' and 'low-'risk patients. On applying these cutoffs, the risk score was highly sensitive (2- and 1-year sensitivity 70.7% (95% CI 68.5–72.8%) and 81.5% (95% CI 79.2–83.9%), respectively) but slightly less specific (2- and 1-year specificity 66.0% (95% CI 65.0–67.0%) and 56.4% (95% CI 55.3–57.4%), respectively; Table 3). By extending this risk categorization to tertile of increasing risk, our risk scores effectively separated patients in real-life clinical terms; the proportion of patients in AROii who actually died within 1 and 2 years increased significantly as tertile of risk increased from 'low' through 'medium' to 'high' (all chi-squared for trend *P* values <0.001, respectively; Table 4). Calibration curves—which

essentially answer the question 'do close to x of 100 patients with a risk prediction of x% have the outcome?'[16]—demonstrate a strong linear relationship between predicted and actual 1- and 2-year mortality (Figure 4). Greater calibration was observed for 2 years ($R^2 = 0.98$) than for 1 year ($R^2 = 0.94$), possibly reflecting fewer events in the latter; the consistently lower predicted versus observed mortality in both accords with the lower specificity described above.

Risk-score validation
The predictive 1- and 2-year risk scores were highly discriminatory when applied externally to the DOPPS population (Table 5). Although generalizability remained high when the DOPPS population was restricted to distinct geographic locations, small 'regional' differences were noted, with the predictive value being lower in North America and

Table 2 | Risk markers for 1- and 2-year all-cause mortality, with associated derived risk score points, in a European incident hemodialysis cohort

Parameter (unit) and values	2-year all-cause mortality		1-year all-cause mortality	
	HR[a] (95% CI[b])	Points[c]	HR (95% CI)	Points
Age—continuous (years)	1.04 (1.03–1.04)		1.03 (1.03–1.04)	
Age—categorical (years)				
≤39		−5		−5
40 to 49		−2		−2
50 to 59		0		0
60 to 69		2		2
70 to 79		4		4
≥80		6		6
Smoking status				
Nonsmoker	1	0		
Former	1.16 (1.03–1.31)	1		
Current	1.28 (1.02–1.59)	1		
Body mass index (kg/m^2)				
<18.5	1.66 (1.36–2.03)	3	1.49 (1.15–1.92)	2
18.5 to <25	1	0	1	0
25 to <30	0.88 (0.78–0.99)	−1	0.97 (0.84–1.13)	0
≥30	0.80 (0.69–0.92)	−1	0.84 (0.70–1.01)	−1
Cardiovascular disease history				
Yes	1.30 (1.17–1.44)	1	1.33 (1.17–1.51)	2
No	1	0	1	0
Cancer history				
Yes	1.75 (1.49–2.05)	3	1.93 (1.60–2.32)	4
No	1	0	1	0
Chronic kidney disease etiology				
Hypertension/vascular	1.00 (0.79–1.27)	0		
Glomerulonephritis	1	0		
Diabetes	1.43 (1.14–1.79)	2		
Tubulo-interstitial	1.30 (1.03–1.66)	1		
Polycystic kidney disease	0.83 (0.58–1.20)	−1		
Miscellaneous/other	1.12 (0.89–1.39)	1		
Vascular access				
No change: Fistula/graft	1	0	1	0
No change: Catheter	1.42 (1.25–1.61)	2	1.48 (1.23–1.76)	2
Change: Fistula/graft to catheter	1.32 (1.06–1.63)	2	1.46 (1.13–1.89)	2
Change: Catheter to fistula/graft	1.08 (0.88–1.34)	0	1.12 (0.84–1.50)	1
Actual blood flow (ml/min)				
<267.0	1	0	1	0
267.0–<298.7	0.89 (0.77–1.03)	−1	0.86 (0.72–1.03)	−1
298.7–<332.1	0.88 (0.76–1.02)	−1	0.82 (0.69–0.98)	−1
≥332.1	0.83 (0.71–0.97)	−1	0.82 (0.67–1.00)	−1
Hemoglobin (g/l)				
<100	1.24 (1.11–1.38)	1	1.30 (1.14–1.49)	2
100 to <120	1	0	1	0
≥120	0.82 (0.71–0.96)	−1	0.86 (0.71–1.06)	−1
Ferritin (μg/l)				
<500	0.85 (0.76–0.96)	−1	0.84 (0.74–0.97)	−1
≥500	1	0	1	0
C-reactive protein (mg/l)				
<2.6	1	0	1	0
2.6–<7.0	1.31 (1.08–1.59)	2	1.25 (0.97–1.61)	1
7.0–<18.2	1.58 (1.31–1.90)	3	1.63 (1.29–2.05)	3
≥18.2	2.11 (1.77–2.53)	4	2.21 (1.78–2.75)	5
Serum albumin (g/l)				
<35	1.42 (1.27–1.59)	2	1.66 (1.42–1.94)	3
≥35	1	0	1	0

Table 2 | (Continued)

Parameter (unit) and values	2-year all-cause mortality		1-year all-cause mortality	
	HR[a] (95% CI[b])	Points[c]	HR (95% CI)	Points
Creatinine (μmol/l)				
<431.1	1.46 (1.23–1.72)	2	1.45 (1.19–1.76)	2
431.1–<539.2	1.19 (1.01–1.41)	1	1.13 (0.91–1.40)	1
539.2–<672.9	1.09 (0.92–1.30)	0	1.09 (0.87–1.35)	0
≥672.9	1	0	1	0
Calcium (mmol/l)				
<2.1			1.11 (0.98–1.27)	1
2.1 to <2.6			1	0
≥2.6			1.68 (1.06–2.65)	3

Multivariate analysis. Parameters significant at the 5% level shown.
[a]HR, hazard ratio.
[b]CI, confidence interval.
[c]Risk-score points.

higher in Japan. Risk stratification capacity was also good, with observed mortality increasing with tertile of increasing predicted risk (chi-squared for trend *P* values <0.001; Table 4).

Additional discrimination over existing scores
When the previously published Wright–Khan[6,7] classification was applied, 3381 (35%), 4248 (44%), and 2093 patients (21%) were classified as low, medium, and high risk, respectively. Compared with medium-risk patients, low-risk patients experienced a lower event rate (HR 0.41; 95% CI 0.36–0.48), whereas high-risk patients experienced a higher rate (HR 1.80; 95% CI 1.63–2.00). In this dialysis population, the predictive power of the Wright–Khan classification was moderate (area under the curve (AUC) 0.66; Table 6). The addition of ARO score predictors improved the predictive power (AUC 0.74), with a net 24 and 27% of patients with and without events, respectively, correctly reclassified. Dialysis and laboratory parameters appeared to have the greatest impact.

Applying the Liu comorbidity index,[12] 5315 (55%), 1860 (19%), and 2547 (26%) patients were classified as low (0–3 points), medium (4 points), and high (≥5 points) risk, respectively, and this variable was predictive of mortality (low- vs. medium-risk HR 0.75; 95% CI 0.66–0.85; high vs. medium risk HR 1.55; 95% CI 1.36–1.77). Nevertheless, the addition of the ARO score variables improved the predictive power (AUC from 0.60 to 0.75), and a net 35 and 31% of patients with and without events, respectively, were correctly reclassified. Initially, the addition of age had the greatest effect, with the subsequent addition of medical and clinical history contributing little to correct reclassification. When dialysis and laboratory parameters were added, however, further correct reclassification was observed. An additional analysis, based on the Liu comorbidity index excluding CKD etiology (in their original study,[12] the score was more predictive when this parameter was removed), gave similar findings (Supplementary Table S4 online).

Risk prediction over shorter time periods
The 2-year score was highly predictive of 1-year death (c-index range 0.74–0.75), although less so than the 1-year score. Importantly, in the subset of patients who had not commenced HD (*N* = 4247), it effectively predicted mortality in the first 90 days (c-index = 0.71).

DISCUSSION
We describe a sensitive and discriminate mortality risk score developed using a large European cohort of incident HD patients. The model was robust, with similar performances in incident dialysis patients at 0–90 or 90–180 days into chronic treatment. Of note, our population started dialysis in 2007–2009: it reflects the current state of the art in medical therapy. In contrast, the most recent previous mortality risk model study included patients initiating dialysis in 2002–2004.[14]

Our aim was not to develop a risk score dedicated to incident patients on HD, but a versatile mortality risk prediction tool generalizable to the widest possible HD population, including both incident and prevalent dialysis patients. External validation in DOPPS confirmed this, with a high degree of discrimination observed when we validated the score against the incident subset and the predominantly prevalent component in DOPPS (Table 5). Generalizability to HD in other geographic areas was also apparent. The observed C-statistic generated (~0.73), although 'acceptable' rather than 'excellent,'[17] was comparable with the previous internally validated studies of Couchoud (0.70),[9] Cohen (AUC 0.77),[11] and van Walraven (0.75),[13] as well as in internal validation of the Framingham Risk Score (0.79[18]). The development of a mortality risk score in a large international database such as AROii, with external validation in another independent worldwide data set as DOPPS, goes significantly beyond previous risk prediction tools. Furthermore, we demonstrate that the use of simple clinical, dialysis, and laboratory routine parameters improved predictive ability over more parsimonious models based on comorbidities alone or age and comorbidities.

ARO All-cause mortality risk score for patients on chronic hemodialysis					
Parameter (unit) and values	1-Year risk points	2-Year risk points	Parameter (unit) and values	1-Year risk points	2-Year risk points
Age [years]			Actual blood flow [ml/min]		
≤39	−5	−5	< 267	0	0
40 to 49	−2	−2	267 to < 299	−1	−1
50 to 59	0	0	299 to < 332	−1	−1
60 to 69	2	2	≥ 332	−1	−1
70 to 79	4	4	Hemoglobin [g/dl]		
≥80	6	6	<10	2	1
Smoking status:			10 to <12	0	0
Current	–	1	≥ 12	−1	−1
Former	–	1	Serum ferritin [µ/l]		
Non smoker	–	0	< 500	−1	−1
CVD history			≥ 500	0	0
Yes	2	1	C-reactive protein [mg/l]		
No	0	0	< 2.6	0	0
Cancer history			2.6 to < 7.0	1	2
Yes	4	3	7.0 to < 18.2	3	3
No	0	0	≥ 18.2	5	4
CKD Etiology:			Serum albumin [g/l]		
Hypertension/vascular	–	0	<35	3	2
Glomerulonephritis	–	0	≥35	0	0
Diabetes	–	2	Serum creatinine [µmo/l]		
Tubulo-interstitial	–	1	< 431	2	2
Polycystic kidney disease	–	−1	431 to < 539	1	1
Unknown renal diagnosis	–	1	539 to < 673	0	0
BMI [kg/m²]			≥ 673	0	0
< 18.5	2	3	Serum total calcium [mmo/l]		
18.5 to < 25.0	0	0	<2.1	1	–
25.0 to < 30	0	−1	2.1 to <2.6	0	–
≥ 30	−1	−1	≥ 2.6	3	–
Vascular access					
No change: Fistula/Graft	0	0			
No change: Catheter	2	2			
Change: Fistula/Graft to Catheter	2	2	Total cumulated risk points		
Change: Catheter to Fistula/graft	1	0			
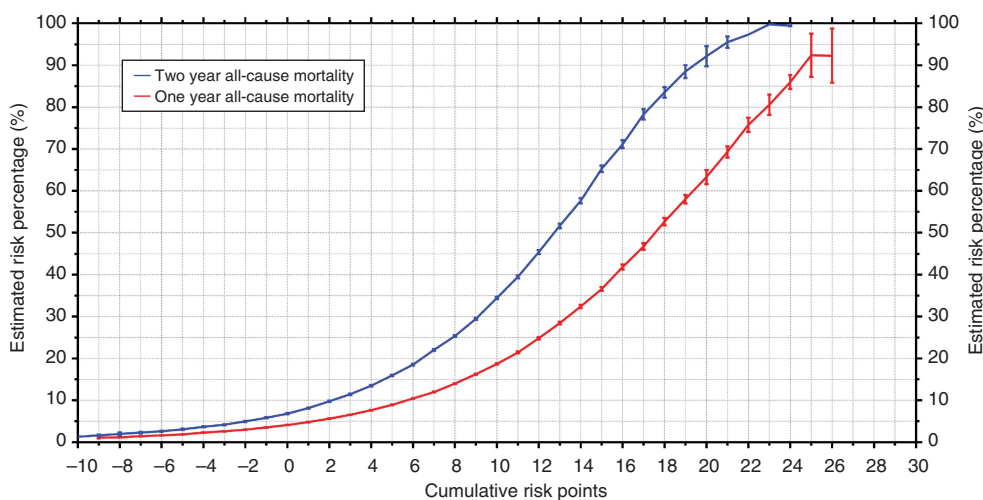
Risk stratification	Low risk	Intermediate risk	High risk
1-Year	< 9%	9 to <19%	≥19%
2-Year	< 15%	15 to <29%	≥29%

Figure 2 | Convenient risk-point calculator printout, including conversion from risk points to estimated all-cause mortality and subsequent categorization in low-, intermediate-, and high-risk groups.

Five aspects particularly distinguish this from previously developed risk scores. First, prior attempts were based on often-small patient populations confined to one geographic area.[8,9,11,12,14,19] Second, in contrast to other studies,[9–14] we focused exclusively on incident HD patients, thus minimizing survival bias. Third, we studied patients from various countries, socioeconomic groups, and health-care systems, and prospectively collected data without exclusions. Other recent scores have focused on older dialysis patients,[9] transplant wait-listed HD patients,[13] or particular socioeconomic groups.[12] Fourth, we demonstrate the improved discrimination of our risk predictors over older models[6,7] comprising age and/or comorbidities alone, and reinforce the clinical meaningfulness of our score through risk stratification capacity analyses. Age and comorbidities should both be integrated in a risk prediction tool as main drivers for mortality;[20] retaining only comorbidities may limit predictive power.[12,21] This is apparent in the current study, in which the addition of age alone to the Liu comorbidity index correctly reclassified a net 35 and 11%
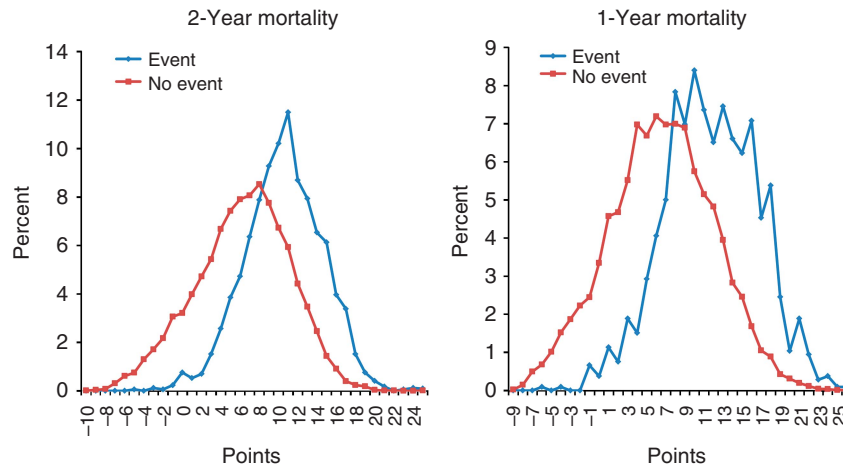
Figure 3 | The distribution of 1- and 2-year risk score points for European incident hemodialysis patients with and without the event of interest.

Table 3 | Sensitivity and specificity of the mortality risk scores among European hemodialysis patients when applied to high- and low-risk groups

Risk Period	Risk	Died (%)	Survived (%)	Total
2-year	Low	503 (8.7)	5285 (91.3)	5788
	High	1211 (30.8)	2723 (69.2)	3934
	Total	1714	8008	9722
		Sensitivity: 70.7% [95% CI 68.5–72.8%]		
		Specificity: 66.0% [95% CI 65.0–67.0%]		
1-year	Low	196 (3.9)	4883 (96.1)	5079
	High	864 (18.6)	3779 (81.4)	4643
	Total	1060	8662	9722
		Sensitivity: 81.5% [95% CI 79.2–83.9%]		
		Specificity: 56.4% [95% CI 55.3–57.4%]		

Abbreviation: CI, confidence interval.

of patients with and without events (Table 6). Laboratory parameters provided the most discriminatory advantage in our score, in line with previous observations.[22,23] Finally, and potentially most relevant clinically, we apply our score to a population selected for HD but who had not yet initiated HD, and effectively predict early mortality in this group.

From a methodological viewpoint, our approach lies toward the simplistic end of the analytic spectrum. Advantageously, the methods are easily replicable and a simple risk calculator could be implemented easily (e.g., in smartphone applications). Like others,[9,13,14] we used imputation to deal with missing data; our choice of Cox regression is also customary.[12–14] Other approaches range from simplistic Kaplan–Meier plots[6] to complicated fractional polynomial[13] or bootstrapped logistic regression models.[9] More complicated models may improve prognostic ability,[24] but additional computational complexity may well outweigh potential benefits.

Generic indices such as Charlson Comorbidity Index (CCI)[25] or the Index of Co-Existent Disease (ICED),[21] although applied to dialysis patients, were not designed for

bedside use and required adaptation.[26,27] Although such scoring systems may be useful in administrative and economical decision-making, they appear to perform less well for mortality risk prediction in HD populations.[20] Our score, in which ∼50% of the variables were potentially modifiable, may also lay the ground for specific intervention studies. In contrast to others,[10] our score includes notably few comorbidities. Another recent study supplemented data on four variables (age, dementia, peripheral vascular disease, albumin) with a 'surprise' question ('Would I be surprised if this patient died within the next 6 months?').[11] However, only 6-month mortality and 500 patients derived from five US dialysis centers were assessed. Performance of the 'surprise' question in other countries, especially on long-term outcomes, is unknown and so far not externally validated.

Clinically, a mortality risk score should be based on routinely measured parameters, which are easy to derive and calculate. For example, the ICED evaluation can take trained people up to 1 h to complete.[28] It should also be based on accurate, objectively measured variables. This excludes, for example, dementia and congestive heart failure (difficult to define in dialysis patients), or subjective parameters such as self-rated health.[29] This practical aspect must be balanced against potential uncontrolled confounding. Highly subjective parameters—such as the 'surprise' question,[11,30] however complementary—will invariably depend on physician background/training and patient knowledge. Although we do not advocate for a binary measure of risk, we would argue that the observed high sensitivity (the ability of the score to correctly identify high-risk patients) is clinically advantageous in the dialysis setting, as the false positive rate (those considered low-risk based on their score but who died nonetheless) will be low.

Our approach has limitations. First, our study is based on data generated from a single commercial dialysis provider, and therefore it could be considered less generalizable to the

Table 4 | Risk stratification capacity of the risk score: estimated versus actual all-cause mortality in hemodialysis cohorts for patients classed as 'Low', 'Medium', and 'High' risk based on their risk score

Dataset	Model-based estimated risk of all-cause mortality	Actual all-cause mortality			
		2-year		1-year	
		Died (%)	Did not die (%)	Died (%)	Did not die (%)
AROii	Low[a]	112 (3.9)	2728 (96.16)	69 (2.2)	3077 (97.8)
	Medium[b]	391 (13.3)	2557 (86.7)	284 (8.6)	3009 (91.4)
	High[c]	1211 (30.8)	2723 (69.2)	707 (21.5)	2576 (78.5)
DOPPS	Low[d]	195 (5.8)	3151 (94.2)	69 (2.4)	2804 (97.6)
	Medium[e]	537 (15.2)	2990 (84.8)	322 (8.3)	3571 (91.7)
	High[f]	1004 (26.8)	2738 (73.2)	657 (17.1)	3192 (82.9)

Abbreviations: AROii, second Analyzing Data, Recognizing Excellence and Optimizing Outcomes (ARO) cohort; DOPPS, Dialysis Outcomes and Practice Patterns Survey.
[a]Lower tertile of risk (<15% for 2 years; <9% for 1 year).
[b]Intermediary tertile of risk (15% to <29% for 2 years; 9% to <19% for 1 year).
[c]Upper tertile of risk (≥29% for 2 years; ≥19% for 1 year).
[d]Lower tertile of risk in the DOPPS population (<15% for 2 years; <9% for 1 year).
[e]Intermediary tertile of risk in the DOPPS population (15% to <29% for 2 years; 9% to <19% for 1 year).
[f]Upper tertile of risk in the DOPPS population (≥29% for 2 years; ≥19% for 1 year).
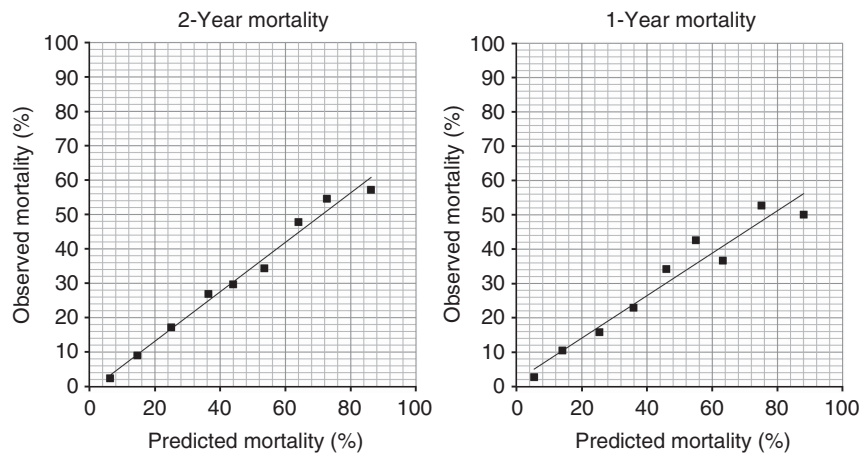
Figure 4 | The relationship between predicted and observed 1- and 2-year mortality in a European incident hemodialysis cohort.

Table 5 | External validation for the AROii mortality risk score at 1 and 2 years in DOPPS, C-statistic

DOPPS population	C-statistic by mortality[a]	
	2-year	1-year
Incident on hemodialysis	0.73–0.75	0.75–0.76
Prevalent on hemodialysis	0.72	0.73
All	0.72	0.72–0.73
Europe	0.72	0.73–0.74
North America	0.68–0.69	0.69–0.70
Australia	0.71–0.73	0.72–0.74
Japan	0.77–0.79	0.77–0.79

[a]Range over 10 iterations.

wider HD population. Our risk score performed favorably when applied externally to the DOPPS population, however, suggesting that it can be applied to a large community of HD patients, although we acknowledge that it may be less generalizable to peritoneal dialysis patients. Focusing on one

large provider allowed us a clinical database, which indicates that recorded outcomes will reflect patients' diagnoses rather than health-care providers' claims for reimbursement, as might be observed in administrative databases.[10,12]

Second, we only assessed 90-day, 1-year, and 2-year mortality, whereas 3-[14] or 5-year[24] mortality may also be important. However, where reported, a remarkably similar C-statistic of 0.75 was obtained in longer analyses.[14]

Third, patients with no laboratory data were excluded, but the loss of 500 of ~11,000 patients in AROii contrasts favorably with a recent registry analysis[14] in which ~5500 of 11,000 patients were excluded owing to missing data.

Fourth, a relevant portion of patients were LTFU, especially in the first year, partly reflecting our stringent definition. When we assumed that nonreturning patients had died, however, this had no major bearing on our findings.

Fifth, comorbidity severity was not considered. However, this often introduces subjectivity, and including severity grade did not improve the model in another analysis.[26]

Table 6 | Additional 2-year all-cause mortality discriminatory ability, conferred by different risk predictors, in a European incident hemodialysis cohort

Analysis	Original model	New model	AUC	Δ AUC	Abs IDI	NRI$_{Events}$	NRI$_{Non-events}$
Wright–Khan variable analysis							
Independent addition of variables							
	Wright–Khan	—	0.661			—	—
	Wright–Khan	+ medical history	0.679	0.018	0.005	0.23	− 0.07
	Wright–Khan	+ clinical factors	0.677	0.016	0.007	0.08	0.07
	Wright–Khan	+ dialysis	0.687	0.026	0.012	0.33	− 0.08
	Wright–Khan	+ labs	0.722	0.061	0.041	0.20	0.24
Cumulative addition of variables							
	Wright–Khan	—	0.661			—	—
	Wright–Khan	+ medical history	0.679	0.018	0.005	0.23	− 0.07
	Wright-Khan + medical history	+ clinical	0.687	0.008	0.006	0.08	0.07
	Wright–Khan + medical history + clinical	+ dialysis	0.701	0.015	0.010	0.27	0.01
	Wright–Khan + medical history + clinical + dialysis	+ labs	0.738	0.036	0.034	0.19	0.21
All vs. Wright–Khan alone							
	Wright–Khan	+ medical history + clinical + dialysis + labs	0.738	0.077	0.056	0.24	0.27
Liu variable analysis							
Independent addition of variables							
	Liu	—	0.601	—	—	—	—
	Liu	+ age	0.696	0.094	0.046	0.35	0.11
	Liu	+ medical history	0.610	0.009	0.005	− 0.65	0.80
	Liu	+ clinical factors	0.622	0.021	0.007	− 0.01	0.15
	Liu	+ dialysis	0.642	0.041	0.013	0.27	0.02
	Liu	+ labs	0.699	0.098	0.051	0.20	0.28
Cumulative addition of variables							
	Liu	—	0.601	—	—	—	—
	Liu	+ age	0.696	0.094	0.046	0.35	0.11
	Liu + age	+ medical history	0.700	0.004	0.003	− 0.62	0.60
	Liu + age + medical history	+ clinical	0.709	0.009	0.007	0.05	0.16
	Liu + age + medical history + clinical	+ dialysis	0.721	0.012	0.010	0.32	− 0.04
	Liu + age + medical history + clinical + dialysis	+ labs	0.750	0.029	0.030	0.17	0.23
All vs. Liu alone							
	Liu	+ age + medical history + clinical + dialysis + labs	0.750	0.149	0.096	0.35	0.31

Abbreviations: Abs IDI, Absolute Integrated Discrimination Improvement; AUC, area under the curve; NRI, (category-free) net reclassification improvement. Wright–Khan: patients classified as low, medium, and high risk according to the score of Wright *et al.*[6] and Khan *et al*[7]; Liu: patients classified into tertile of increasing risk according to the comorbidity index of Liu *et al*[12]; Medical history: CKD etiology (Wright–Khan variable analysis only), history of cancer and/or cardiovascular disease; Clinical: Body mass index, smoking status; Dialysis: Vascular access change, actual blood flow; Labs: serum albumin, C-reactive protein, hemoglobin, ferritin, and creatinine. NRI$_{Events}$ and NRI$_{Non-events}$ correspond, respectively, to the proportion of events/nonevents reclassified correctly minus the proportion of events/nonevents reclassified incorrectly. For ΔAUC, and Abs IDI, a positive number corresponds to more events/nonevents being reclassified correctly.

Sixth, potential predictors of death particularly in elderly dialysis patients—such as late referral, dependency for transfers, severe behavioral disorders, health-related quality of life, frailty assessment, and unplanned dialysis[9,31]—could not be assessed in our analysis of routinely captured data. Other predictive parameters in dialysis patients[32–35] might conceivably improve our score.

Seventh, the inclusion of patients receiving kidney transplants during follow-up may have selected a healthy cohort, as transplant-listed patients tend to be younger and healthier. By treating these events as censored observations, however, we were in accordance with the analytical

recommendations of a recent study focusing on the issue of renal transplantation as a competing event in survival analysis in nephrology.[36] We acknowledge, however, that our score may be less generalizable to HD populations with excessively higher transplant rates than ours (∼8%).

Finally, although country-specific predictions might provide further insights, many subgroups would be too small for meaningful analyses. Of note, our score yielded slightly lower C-statistics in the US DOPPS patients compared with patients in Europe and, in particular, with patients in Japan. This suggests potentially unmeasured confounding in the US analysis. Within Europe, the almost

identical score performance in the AROii and DOPPS cohort strongly suggests that the AROii population is representative for a larger European population.

In conclusion, we describe a novel mortality risk score, potentially applicable to all incident or prevalent HD patients, with improved predictive power over age-/comorbidity-based models. Such a tool is now available for research purposes, allowing either correction for imbalanced mortality risk within groups or the selection of high-risk patients; the data may also prove to be useful when communicating with patients, for example, through highly visual heat-maps (Supplementary Figure S2 online). Although risk scores may help generate hypothesis, they can only describe associations and not establish causality.

MATERIALS AND METHODS
Study population
The ARO research initiative[37] comprises European FME HD patients. Anonymized patient-level medical history data, plus longitudinal laboratory and medication data are captured quarterly, as are ICD-10-coded hospitalizations and deaths. The study is based on AROii, which includes consecutive incident (<6 months on dialysis) adult patients without renal transplantation history collected from >300 FME facilities in 14 European countries in 2007–2009. Data on chronic dialysis patients (⩾10 contiguous HD sessions) with available laboratory data were restricted further to patients remaining in the study for ⩾3 months (Figure 1). Patient data were anonymized, and informed consent was obtained from all patients by FME.[37] Local ethics committees' approvals were collected.

Follow-up
Follow-up commenced on the date of patients' first FME HD session; time at risk was accrued from the end of baseline (first 3 months of follow-up) until patients experienced the event of interest or were censored: undergoing a renal transplant, being LTFU (>45 days without continuous FME dialysis treatment; LTFU), or end of follow-up (30 September 2011).

Statistical analysis
Statistical analyses, reproduced independently by a second biostatistician, were performed using SAS (Windows version 9.2; SAS, Cary, NC, USA). Baseline demographic, clinical, dialysis, and laboratory data were categorized into quartiles or biologically relevant groups (Table 1). Continuous variables were described using means and standard deviations; categorical data were reported as counts and frequencies. Rates per 100 person-years with accompanying 95% CIs were calculated.

Risk-score derivation
Risk scores were derived using a modified Framingham Heart Study[38] approach. The predictive effect of exposures was determined using baseline Cox regression models. Ten imputed data sets were created using Monte-Carlo Markov Chain multiple imputation to deal with missing observations, with all covariates described in Table 1 fitted initially. Generated coefficients were combined using Rubin's rules,[39] and variables with at least one significant (5% level) stratum were retained. Coefficients were converted to 'points', with 1 point indicating the risk equivalent of 5 years additional age; a total score was computed for each patient by summing all points.

Estimated absolute risk was calculated by substituting the points total for each risk category in the original Cox model. An algorithm describing risk score point derivation in detail and including 1- and 2-year risk equations is included in the Supplementary material online.

Internal discrimination and calibration
A number of analyses were performed to quantify the relationship between risk score predictions and actual mortality. Traditional measures of discrimination—sensitivity and specificity—were calculated, with the predicted low-/high-risk cutoff determined graphically (Figure 3). This discrimination was extended to further assess the risk stratification capacity: patients were categorized as 'low', 'medium', or 'high' risk on the basis of tertile, and this was compared with their actual 1- and 2-year death status. Changes in the proportion of deaths across these strata were assessed using the chi-squared test for trend. Finally, calibration was assessed graphically: patients' predicted survival was binned into 10% risk groups, and these data were plotted against observed death, calculated as the proportion of patients in each group who actually died.

Contribution of modifiable and non-modifiable risk markers
Further analyses were performed to demonstrate potential clinical utility. By choosing, for illustrative purposes, modifiable and non-modifiable risk markers, a risk 'heat-map' was constructed. Risk markers were stratified in a nested 2×2 arrangement, and the product of the points contributing to strata was calculated. A low-/medium-/high-risk 'traffic-light' color scheme was applied on the basis of tertiles of all calculated points.

By assigning risk scores to patient-level risk markers, the relative contribution of potentially modifiable and non-modifiable (Table 1) risk markers was calculated: if a patient had a score of 15 and 6 of these points were from modifiable factors, then it was assumed that 40% of their risk was modifiable. By applying these proportions to the actual risk percentage of death, the relative contribution of modifiable with increasing risk score was estimated.

Risk-score validation
External validation was performed using the third DOPPS[19,40,41] cohort. This was chosen for reasons of contemporaneousness (DOPPS: 2005–2008; AROii: 2007–2009) and data comparability. The DOPPS study design dictates randomly selected incident and prevalent HD patients from randomly selected dialysis units within each country, with chosen facilities representing the different facilities and regions within each country. The seven-group ICD-10-based CKD etiology category defined in AROii was applied in DOPPS using ICD-9 diagnostic codes. DOPPS data were also limited to patients with ≥3 months' follow-up. The discriminatory ability of the predictive risk-score model was assessed using Harrell's C-statistic.[42] These scores range from 0 to 1, with 1 indicating a perfect prediction and 0.5 indicating a chance prediction; values of 0.7–0.8 could be considered acceptable and those of 0.8–0.9 could be considered excellent.[17]

Sensitivity analyses
Additional analyses were conducted to assess the impact of our methodological choices. It was possible that a proportion of patients LTFU were lost because they withdrew from dialysis as their condition deteriorated, with these patients censored instead of

counted as having a mortality event. To assess the potential impact of such a misclassification, we considered all nonreturning LTFU patients as deceased on their last observation date and repeated the analysis. In addition, to test whether the first 90 days of HD, which represents a high-risk period,[43] was less predictive of longer-term mortality, the AROii baseline was reset to 90–180 days and the predictor analysis was repeated.

Additional discrimination

Analyses were conducted to assess the additional discriminatory power of our 2-year risk score over existing risk scores. Those described by Wright–Khan *et al.*[6,7] and Liu *et al.*[12] were chosen to represent European and North American patients, respectively. Patients were classified according to each score and grouped into low, medium, and high risk (implicit in the Wright–Khan paper; tertile of patients' points total (including points for CKD etiology) for the Liu paper). Separately, and starting with a model containing this variable, we added independently and sequentially the medical history, clinical, dialysis, and laboratory parameters. At each stage, we calculated the AUC, the Integrated Discrimination Improvement, IDI ('the increment in the predicted probabilities for the subset experiencing an event and the decrement for the subset not experiencing an event'), and the category-free NRI$_{events}$ and NRI$_{non-events}$ (the 'net' proportion of individuals with/without events reclassified correctly using the new model over the original model'), as proposed by Pencina *et al.*,[44] and as calculated by Kennedy.[45][46] For the Liu analysis, the inclusion of age was also assessed, as this comorbidity index omits age.

Risk prediction over shorter time periods

The discriminatory ability of the 2-year risk-score model over shorter time periods was assessed using Harrell's C-statistic. First, the 2-year score was applied to 1-year mortality. Subsequently, the ability to predict early death (in the first 90 days of dialysis) was assessed using predialysis data for the subpopulation of patients who initiated HD in FME facilities. Where data for any risk-score parameter was missing (e.g., 100% for actual blood flow), the neutral risk-score point was applied—i.e., zero.

SUPPLEMENTARY MATERIAL

Supplementary Appendix. Algorithm for deriving points and risk for 2-year all-cause mortality and algorithm for deriving points and risk for one year all-cause mortality.

Table S1. Characteristic of AROii Patients by Lost-to-Follow-Up Status.

Table S2. Characteristic of AROii Patients by Kidney Transplant During Follow-Up Status.

Table S3. Risk Markers for 1- and 2-Year All-Cause Mortality in a European Incident Hemodialysis Cohort. Sensitivity analyses using a 3- to 6-month baseline and assuming that lost-to-follow-up patients had died. Parameters significant at the 95% level shown.

Table S4. Additional 2-Year All-Cause Mortality Discriminatory Ability, Conferred by Different Risk Predictors, in a European Incident Hemodialysis Cohort. Applying the comorbidity index of Liu *et al.*[12] excluding CKD etiology points.

Figure S1. Contribution of modifiable risk markers (as defined in Table 1) to the risk score according to cumulated risk points.

Figure S2. Two-Year All-Cause Mortality Risk "Heat-Map" for a European Incident Dialysis Population, Based on Selected Modifiable and Non-Modifiable Risk Markers.

Supplementary material is linked to the online version of the paper at http://www.nature.com/ki

REFERENCES

1. Eckardt KU, Coresh J, Devuyst O *et al.* Evolving importance of kidney disease: from subspecialty to global health burden. *Lancet* 2013; **382**: 158–169.
2. Coresh J, Selvin E, Stevens LA *et al.* Prevalence of chronic kidney disease in the United States. *JAMA* 2007; **298**: 2038–2047.
3. Meguid El Nahas A, Bello AK. Chronic kidney disease: the global challenge. *Lancet* 2005; **365**: 331–340.
4. Kalantar-Zadeh K, Abbott KC, Kronenberg F *et al.* Epidemiology of dialysis patients and heart failure patients. *Semin Nephrol* 2006; **26**: 118–133.
5. de Jager DJ, Grootendorst DC, Jager KJ *et al.* Cardiovascular and noncardiovascular mortality among patients starting dialysis. *JAMA* 2009; **302**: 1782–1789.
6. Wright LF. Survival in patients with end-stage renal disease. *Am J Kidney Dis* 1991; **17**: 25–28.
7. Khan IH, Catto GR, Edward N *et al.* Influence of coexisting disease on survival on renal-replacement therapy. *Lancet* 1993; **341**: 415–418.
8. Foley RN, Parfrey PS, Hefferton D *et al.* Advance prediction of early death in patients starting maintenance dialysis. *Am J Kidney Dis* 1994; **23**: 836–845.
9. 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.
10. Miskulin D, Bragg-Gresham J, Gillespie BW *et al.* Key comorbid conditions that are predictive of survival among hemodialysis patients. *Clin J Am Soc Nephrol* 2009; **4**: 1818–1826.
11. Cohen LM, Ruthazer R, Moss AH *et al.* Predicting six-month mortality for patients who are on maintenance hemodialysis. *Clin J Am Soc Nephrol* 2010; **5**: 72–79.
12. Liu J, Huang Z, Gilbertson DT *et al.* An improved comorbidity index for outcome analyses among dialysis patients. *Kidney Int* 2010; **77**: 141–151.
13. van Walraven C, Austin PC, Knoll G. Predicting potential survival benefit of renal transplantation in patients with chronic kidney disease. *CMAJ* 2010; **182**: 666–672.
14. 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.
15. Chen JY, Tsai SH, Chuang PH *et al.* A comorbidity index for mortality prediction in Chinese patients with ESRD receiving hemodialysis. *Clin J Am Soc Nephrol* 2014; **9**: 513–519.
16. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating:* New York, 2010.
17. Hosmer DW Jr, Lemeshow S. *Applied logistic regression*. John Wiley & Sons, 2004.
18. D'Agostino RB Sr, Grundy S *et al.* Validation of the framingham coronary heart disease prediction scores: Results of a multiple ethnic groups investigation. *JAMA* 2001; **286**: 180–187.

19. 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.

20. 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.

21. Greenfield S, Apolone G, McNeil BJ *et al.* The importance of co-existent disease in the occurrence of postoperative complications and one-year recovery in patients undergoing total hip replacement. Comorbidity and outcomes after hip replacement. *Medical care* 1993; **31**: 141–154.

22. Takahashi R, Ito Y, Takahashi H *et al.* Combined values of serum albumin, C-reactive protein and body mass index at dialysis initiation accurately predicts long-term mortality. *Am J Nephrol* 2012; **36**: 136–143.

23. Usvyat LA, Barth C, Bayh I *et al.* Interdialytic weight gain, systolic blood pressure, serum albumin, and C-reactive protein levels change in chronic dialysis patients prior to death. *Kidney Int* 2013; **84**: 149–157.

24. 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.

25. Charlson ME, Pompei P, Ales KL *et al.* A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987; **40**: 373–383.

26. 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.

27. Hemmelgarn BR, Manns BJ, Quan H *et al.* Adapting the Charlson Comorbidity Index for use in patients with ESRD. *Am J Kidney Dis* 2003; **42**: 125–132.

28. Levey AS, Martin AA, Miskulin DC *et al.* Comorbidity assessment in ESRD: Methods and results in a national dialysis chain. *J Am Soc Nephrol* 1999; **10**: 169A–170A.

29. Thong MS, Kaptein AA, Benyamini Y *et al.* Association between a self-rated health question and mortality in young and old dialysis patients: a cohort study. *Am J Kidney Dis* 2008; **52**: 111–117.

30. Moss AH, Ganjoo J, Sharma S *et al.* Utility of the 'surprise' question to identify dialysis patients with high mortality. *Clin J Am Soc Nephrol* 2008; **3**: 1379–1384.

31. Joly D, Anglicheau D, Alberti C *et al.* Octogenarians reaching end-stage renal disease: cohort study of decision-making and clinical outcomes. *J Am Soc Nephrol* 2003; **14**: 1012–1021.

32. Kalantar-Zadeh K, Kopple JD, Block G *et al.* A malnutrition-inflammation score is correlated with morbidity and mortality in maintenance hemodialysis patients. *Am J Kidney Dis* 2001; **38**: 1251–1263.

33. Zoccali C, Tripepi G, Mallamaci F. Predictors of cardiovascular death in ESRD. *Semin Nephrol* 2005; **25**: 358–362.

34. Andreozzi F, Presta I, Mannino GC *et al.* A functional variant of the dimethylarginine dimethylaminohydrolase-2 gene is associated with insulin sensitivity. *PloS One* 2012; **7**: e36224.

35. Postorino M, Marino C, Tripepi G *et al.* Prognostic value of the New York Heart Association classification in end-stage renal disease. *Nephrol Dial Transplant* 2007; **22**: 1377–1382.

36. Noordzij M, Leffondre K, van Stralen KJ *et al.* When do we need competing risks methods for survival analysis in nephrology? *Nephrol Dial Transplant* 2013; **28**: 2670–2677.

37. de Francisco AL, Kim J, Anker SD *et al.* An epidemiological study of hemodialysis patients based on the European Fresenius Medical Care hemodialysis network: results of the ARO study. *Nephron Clin Pract* 2011; **118**: c143–c154.

38. Sullivan LM, Massaro JM, D'Agostino RB Sr. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. *Stat Med* 2004; **23**: 1631–1660.

39. Rubin DB. *Multiple imputation for nonresponse in surveys,* vol. 519. Wiley Online Library, 1987.

40. Canaud B, Tong L, Tentori F *et al.* Clinical practices and outcomes in elderly hemodialysis patients: results from the Dialysis Outcomes and Practice Patterns Study (DOPPS). *Clin J Am Soc Nephrol* 2011; **6**: 1651–1662.

41. Pisoni RL, Gillespie BW, Dickinson DM *et al.* The Dialysis Outcomes and Practice Patterns Study (DOPPS): design, data elements, and methodology. *Am J Kidney Dis* 2004; **44**: 7–15.

42. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996; **15**: 361–387.

43. Chan KE, Maddux FW, Tolkoff-Rubin N *et al.* Early outcomes among those initiating chronic dialysis in the United States. *Clin J Am Soc Nephrol* 2011; **6**: 2642–2649.

44. Pencina MJ, D'Agostino RB Sr., D'Agostino RB Jr. *et al.* Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008; **27**: 157–172; discussion 207-112.

45. Kennedy KF, Pencina MJ. A SAS® macro to compute added predictive ability of new markers predicting a dichotomous outcome. In, Institute for Advanced Analytics, 2010.

46. Kerr KF, Wang Z, Janes H *et al.* Net reclassification indices for evaluating risk prediction instruments: a critical review. *Epidemiology* 2014; **25**: 114–121.

APPENDIX