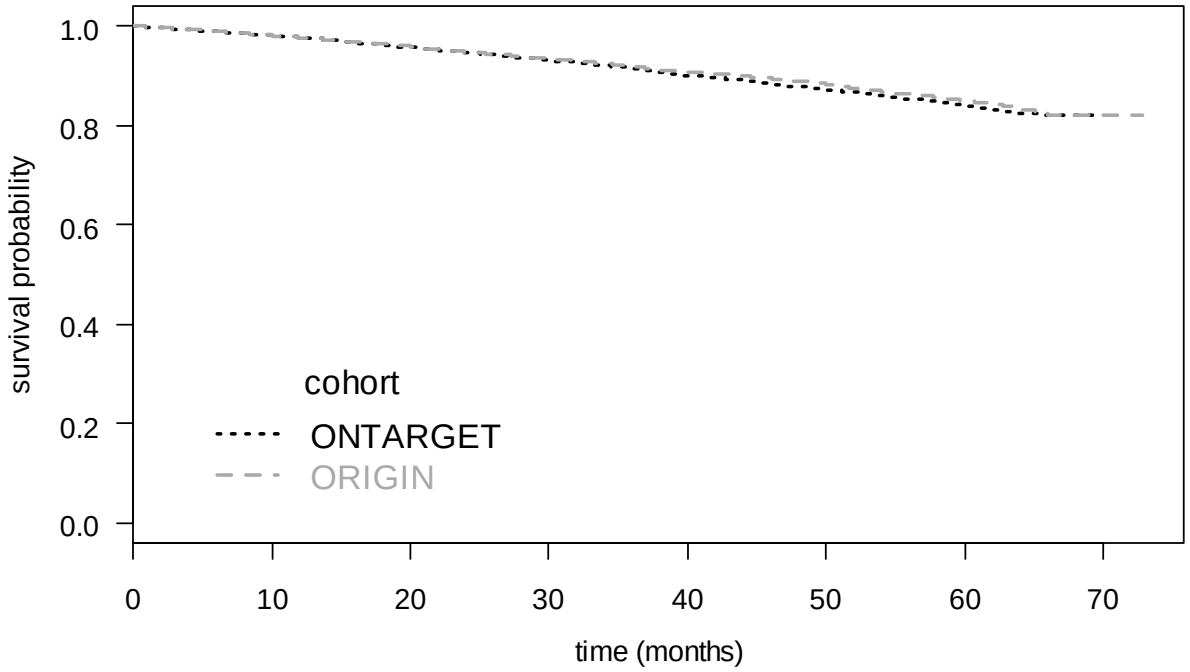
Supplemental Data for

Daniela Dunkler, Peggy Gao, Shun Fu Lee, Georg Heinze, Catherine M. Clase, Sheldon Tobe, Koon K. Teo, Hertzel Gerstein, Johannes F.E. Mann, Rainer Oberbauer on behalf of the ONTARGET & ORIGIN Investigators. Risk Prediction for Early Chronic Kidney Disease in Type 2 Diabetes.

Contents

This supplementary material has been provided by the authors to give readers additional information about their work.

Supplemental Figure 1. Kaplan-Meier Plot Comparing Follow-Up of ONTARGET and ORIGIN



Validation cohort: To receive a similar follow-up as in ONTARGET, the 2-year visit was used as baseline in the validation. In ORIGIN almost all necessary variables were determined at the 2-year visit, with the exception of diabetes and peripheral artery disease. Therefore, the status at baseline was used for these two variables.

Decision Curve Analysis

Methods

Using decision curve analysis (1), we quantified the clinical usefulness of our risk calculator for detecting individuals requiring referral to a renal specialist by computing its net benefit, which is a summary measure of the true and false detection rates.

For evaluating clinical usefulness, predicted probabilities near a pre-defined threshold, i.e., a probability above which individuals are classified as high-risk, are of greatest importance. We conducted a decision curve analysis to compute the 'net benefit' of using our prediction models to decide if an individual should be referred to a renal specialist (1). The net benefit is the true detection rate discounted by the rate of false detections, which are weighted by the assumed 'costs' of a false detection. This analysis makes the assumption, that if for a referral decision a threshold probability of, say, 10% is used, then implicitly the benefit of early detection of one individual who will later experience incidence or progression of CKD is rated to be equal to the costs of nine unnecessary referrals (=(100-10%)/10%). Thus, with a 10% threshold probability for a referral decision the net benefit is the proportion of true positives minus the proportion of false positives divided by nine. By varying the threshold, a decision-analysis curve is obtained which can be compared to the curves showing the net benefits of referring all or no patients.

Results

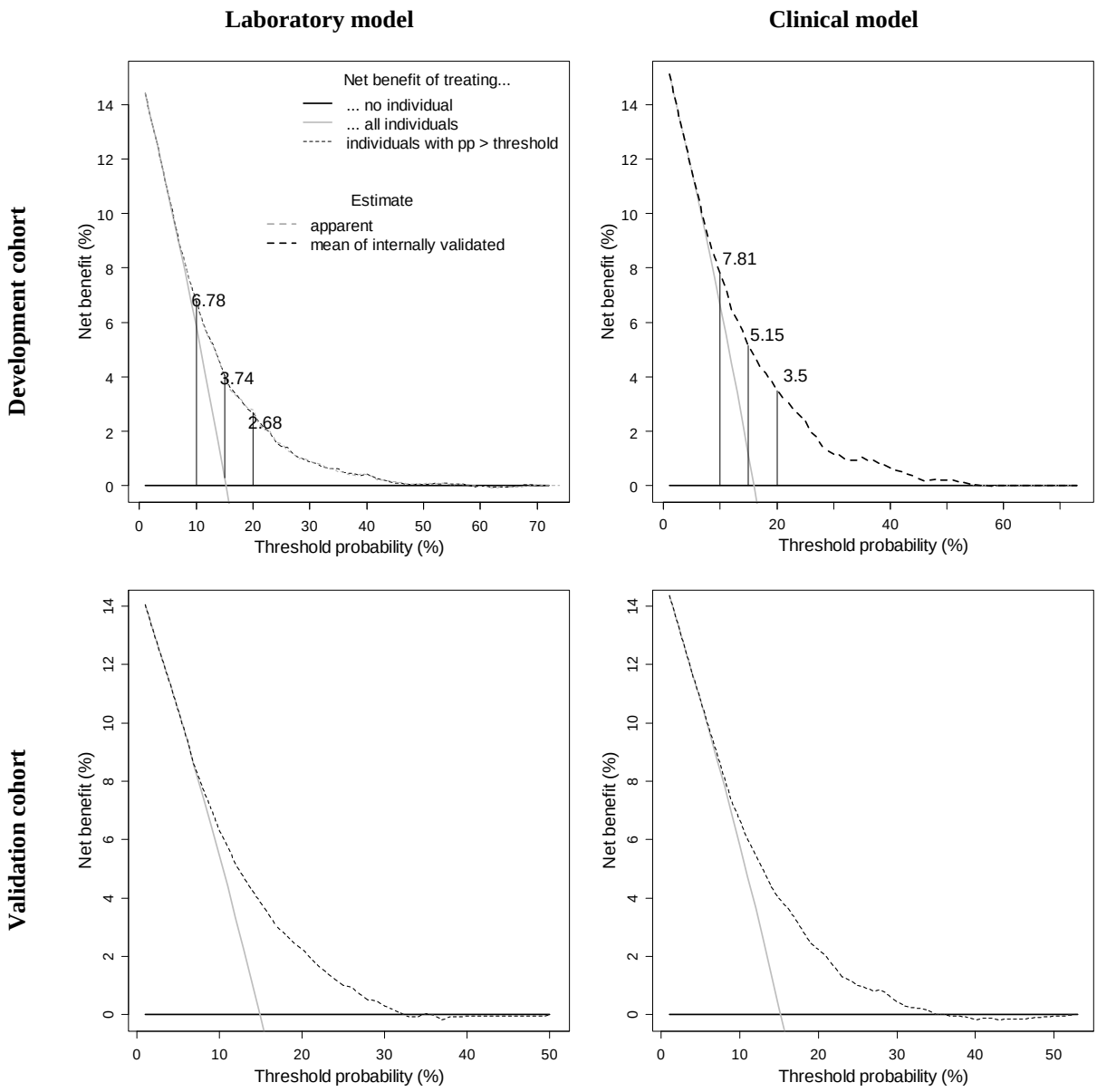
Laboratory model: Supplemental Figure 2 displays decision curves quantifying clinical usefulness. If a probability of incidence or progression of CKD of ≥ 10% was used to decide upon referral of an individual to a specialist, the net benefit of applying the laboratory model was 7.3%; equivalent to the benefit of 7 true-positive detections per 100 individuals without increasing the number of false detections.

Clinical model: If individuals with a probability of incidence or progression of CKD ≥ 10% are classified as high-risk, the net benefit of applying the clinical model was 7.8% (Supplemental Figure 2, right).

Supplemental Figure 2: Decision curve analysis for the laboratory (left panel) and the clinical model (right panel) for the outcome state 'alive with incidence and progression of CKD' based on the development cohort (first row) and the validation cohort (second row).

Decision curves show the net benefit of referring individuals to a specialist in kidney care if a certain threshold probability is exceeded and thus comparing the decision based on the prediction model to the strategies of referring all (dashed line) or no individuals (black horizontal line). A threshold probability of, e.g., 10% means that individuals with a probability of incidence and progression of CKD are classified as high-risk and should be referred to a renal specialist. At the threshold probability the cumulative 'costs' of false-negatives, i.e. high-risk

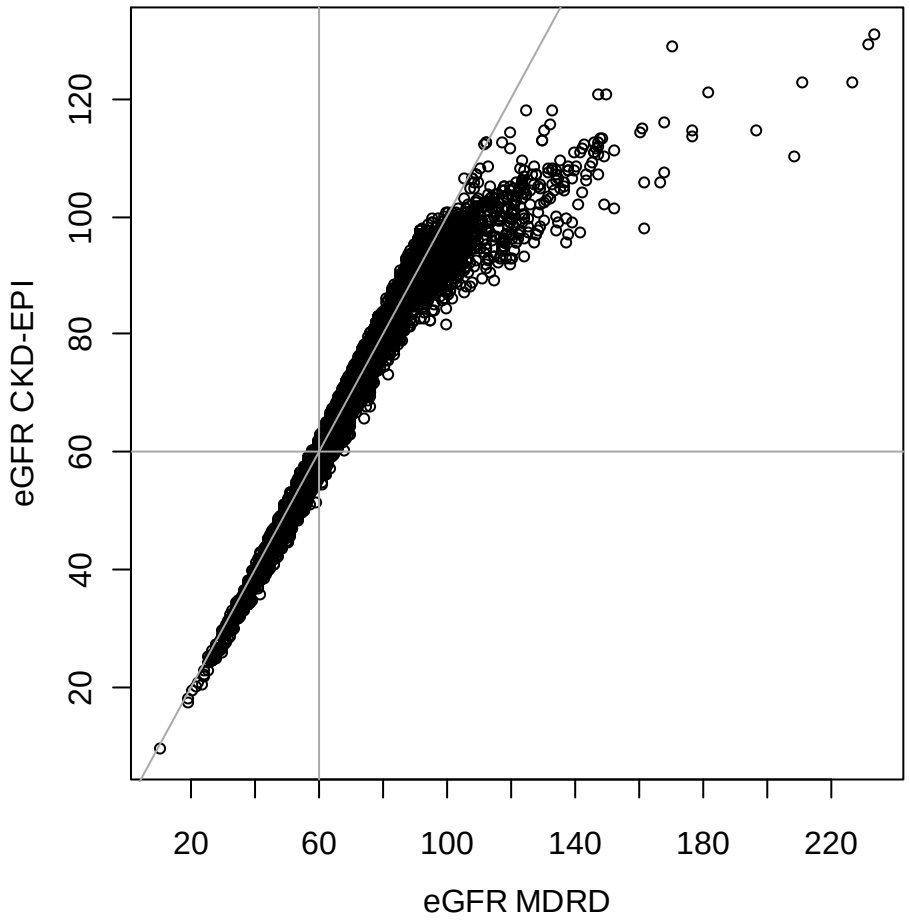
individuals who are not timely referred, and of false-positives, i.e. low-risk individuals who are unnecessarily referred, are assumed to be equal. Threshold probabilities of 10, 15 or 20% correspond to the assumptions that the missed referral of one individual who will develop CKD is 9, 5.7 or 4 times worse than one unnecessary referral, respectively. The numbers in the left panel (first row) indicate that the net benefit of applying the laboratory model: (a) 7.34% at a threshold probability of 10% (i.e. assuming one missed referral of an individual who will develop CKD is 9 times worse than one unnecessary referral), which mean that applying the model is equivalent to 7 additional true-detections per 100 individuals without increasing the number of false detections, (b) 4.53% at a threshold probability of 15%, and (c) 2.74% at a threshold probability of 20%. Note, that the net benefit of a perfect prediction model for CKD can maximally attain 16% (equaling the incidence of CKD).



Abbreviation: pp, predicted probability.

Supplemental Figure 3. Scatterplot Comparing eGFR CKD-EPI and eGFR MDRD (2,3).

Data stem from the development cohort.



Abbreviations: eGFR, estimated glomerular filtration rate.

Comments on the Statistical Analysis

Preliminary work. In ONTARGET the maximum absolute correlation among predictors for the laboratory and the clinical model were 0.25 and 0.39. Values of continuous predictors were truncated at their 0.5^{th} and 99.5^{th} percentiles.

Modeling algorithm. Multinomial logistic regression was applied to develop prediction models for the three outcome states (4). Using the polytomous outcome had the advantage that potential dependence between the different outcome states were incorporated compared to using multiple dichotomous models.

Fractional polynomials were applied to model nonlinear relationships of continuous predictors with the outcome (5). Optimal powers were selected by the RA2-algorithm with the significance level set to 0.157, which corresponds roughly to a selection according to the AIC criterion (6). Continuous predictors were roughly scaled to [0, 1] to ensure convergence of the modeling algorithm. For categorical predictors variable selection was conducted by likelihood-ratio-tests within each cycle of the RA2-algorithm.

After selection of predictors and fractional polynomials, interactions were tested by including all pair-wise product interaction terms into the model. If these interactions were (together) significant in the apparent model, all interactions were checked by graphical means. In neither the laboratory nor the clinical model interactions were significant. If these interactions were significant in a bootstrap resample, then all interactions were included into the respective model.

Model Development and Validation. First, the devised modeling algorithm was applied to the development cohort to construct 'apparent' prediction models. Second, it was applied to 500 bootstrap resamples to evaluate model stability and internal validation. Using 500 bootstrap resamples yields sufficiently small standard deviations for the optimism-corrected c-statistics in the clinical model for the outcome states alive without incidence or progression of chronic kidney disease (CKD), alive with incidence or progression of CKD, and death of 0.008, 0.011, and 0.010, respectively.

Applying the apparent model on the development cohort will typically give optimistic estimates of model performance. Therefore, the model performance was internally and externally evaluated. While internal validation focuses on the development process and the quality of the model in a similar population, external validation focuses on transportability to another population.

For internally validated or optimism-corrected estimates, for each bootstrap resample j the complete modeling algorithm was repeated and the optimism was estimated by $apparent - \left(bootstrap_{j,apparent} - bootstrap_{j,test}\right)$, where $apparent$ was the apparent estimate from the development data, $bootstrap_{j,apparent}$ was the estimate from the j^{th} bootstrap resample applied to the j^{th} bootstrap resample, and $bootstrap_{j,test}$ was the estimate from the j^{th} bootstrap resample applied to the development data (7).

Predictive accuracy is defined as the mean absolute difference between observed and expected outcome probabilities and can be computed for a model with and without predictors (8). The relative improvement in

predictive accuracy due to consideration of predictors is termed explained variation (9), ranging from 0% (no prognostic relevance) to 100% (perfect prediction). Relative importance of individual predictors can be quantified by the drop in explained variation if these predictors are removed from the model (10). The proportion of variation explained by individual predictors is given for each predictor (or groups of predictors) in Table 3 as the respective partial explained variation divided through the explained variation and multiplied with the Nagelkerke-R^2 in order to allow direct comparison between the Nagelkerke-R^2 of the model with the proportion of variation explained by individual predictors.

Performance Measures. If a performance measure has no direct extension to the polytomous outcome, the weighted mean of the three pairwise comparisons is given.

Shrinkage. Optimism-corrected calibration slopes were used as a global linear shrinkage factors (towards the overall mean) to correct the apparent laboratory and clinical models for optimism. The shrunken intercepts α_{1s} and α_{2s} are $\alpha_{1s} = \alpha_1 + (\bar{\eta}_1 - \alpha_1)(1 - s_1)$ and $\alpha_{2s} = \alpha_2 + (\bar{\eta}_2 - \alpha_2)(1 - s_2)$, where α_1 and α_2 are the apparent estimates of the intercepts, η_1 and η_2 are the linear predictors based on apparent regression coefficients, and s_1 and s_2 are global shrinkage factors for outcome 1 ('alive with incidence or progression of CKD') and 2 ('death'), respectively. The shrunken coefficients β_{1is} and β_{2is} for outcome 1 and 2 for $i = 1,...,k$ are simply $\beta_{i_1 s} = \beta_{i_1} s_1$ $\beta_{i_1 s} = \beta_{i_1} s_1$ $\beta_{i_1 s} = \beta_{i_1} s_1$ $\beta_{i_1 s} = \beta_{i_1} s_1$ $\beta_{i_1 s} = \beta_{i_1} s_1$, and $\beta_{i_2 s} = \beta_{i_2} s_2$.

Interpretation of Decision Curve Analysis (1). Individuals with type 2 diabetes mellitus who might be at risk for chronic kidney disease are currently not automatically transferred to a specialist. By using a prediction model, one can categorize individuals into low and high risk of developing chronic kidney disease. An individual with a very low predicted probability for incidence or progression of chronic kidney disease might not need to be transferred to a specialist, whereas a timely transferal may be necessary for an individual with a large predicted probability.

A *threshold probability* can be defined as the probability, starting from which an individual should be transferred to a specialist, because the potential benefits exceed its potential harms or efforts. Thus the choice of the threshold probability depends on the 'costs' of false-positives (i.e. low-risk individuals who are unnecessarily transferred) relative to the 'costs' of false-negatives (i.e. high-risk individuals who are not timely transferred). The net benefit of applying a prediction model for a specific threshold probability is defined as the true positive rate minus the false positive rate, the latter weighted by the relative costs of a false-positive prediction compared to a false negative one. The such-defined net benefit can range from minus infinity up to the prevalence of the outcome observed in a study population.

In order to decide if the *net benefit* of a prediction model for a specific threshold probability is of clinical importance, it can be compared to two hypothetical situations: while in the first, all individuals with type 2 diabetes mellitus are transferred to specialists to avoid false negatives, in the second situation no individual with type 2 diabetes mellitus is transferred to avoid any false positives. The second situation has a net benefit of 0, since there are neither true positives nor false positives. Therefore, compared to the second situation, a prediction model with a positive net benefit is beneficial. In the first situation the net benefit can be computed

by using the proportion of individuals who do and do not experience chronic kidney disease. If the net benefit of transferring all individuals to specialists is smaller than the net benefit of the prediction model, then the application of the prediction model is beneficial.

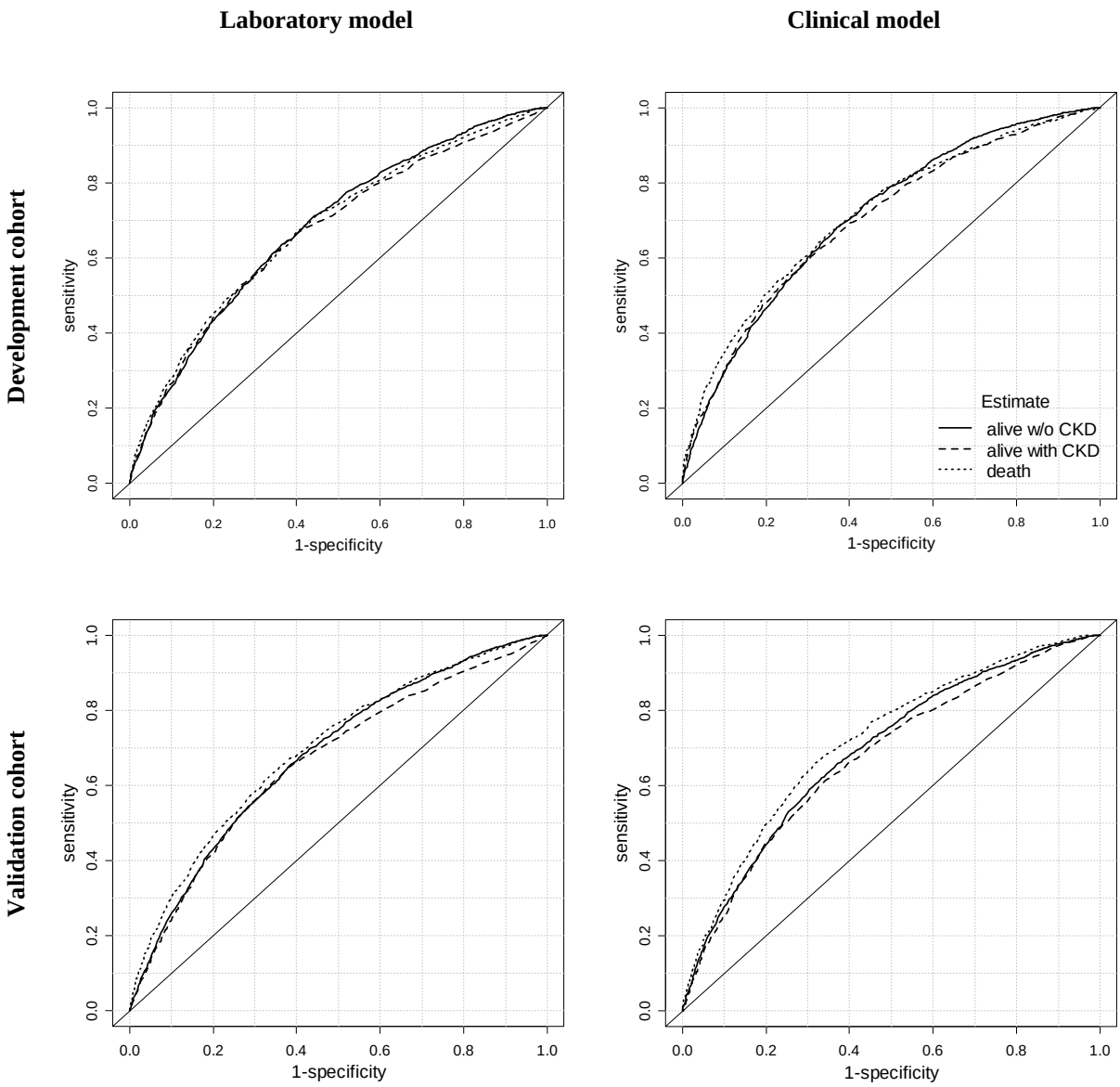
A *decision curve* gives the net benefit for various threshold probabilities for applying the prediction model, for assuming all individuals will experience chronic kidney disease or for assuming all individuals will not experience chronic kidney disease. The net benefit of the prediction model of very low or very high threshold probabilities is generally similar to the net benefit of assuming all individuals will experience chronic kidney disease or all individuals will not experience chronic kidney disease, respectively. Only between these two extremes the prediction model may be of value.

Supplemental Table 1. Performance of Prediction Models in the Development and Validation Cohorts for the Outcome 'Alive without incidence or progression of chronic kidney disease'.

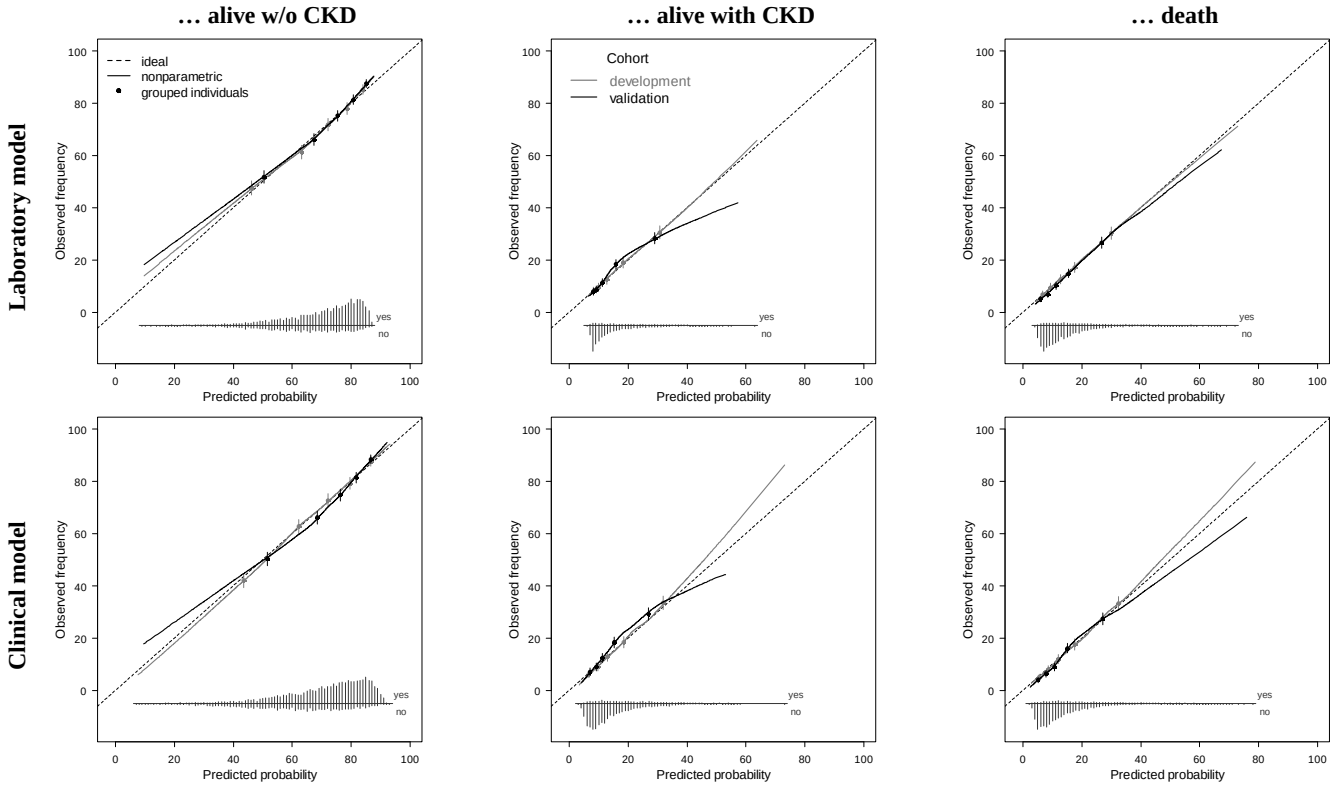
Performance Measures	Laboratory Model	Clinical Model
Explained Variation		
Nagelkerke-R^2		
optimism-corrected	11.70%	13.02%
externally-validated	11.34%	12.66%
Discrimination		
C-statistic		
optimism-corrected	0.68	0.69
externally-validated	0.68	0.69
Calibration [1]		
Calibration-in-the-large		
optimism-corrected	0	0
externally-validated	0.03	-0.04
Calibration slope		
optimism-corrected	0.98	0.91
externally-validated	1.01	1.03

[1] A multinomial logistic model with three outcomes has two estimates for calibration-in-the-large and the calibration slope. In an ideally calibrated prediction model calibration-in-the-large would be 0 and the calibration slope would be 1, indicating that predictions are not systematically biased. The optimism-corrected calibration slopes are used as shrinkage factors of the prediction models.

Supplemental Figure 4. Receiver-Operating Curves (ROC) for the Laboratory and the Clinical Model based on the Development and Validation Cohort.



Supplemental Figure 5. Calibration Plots



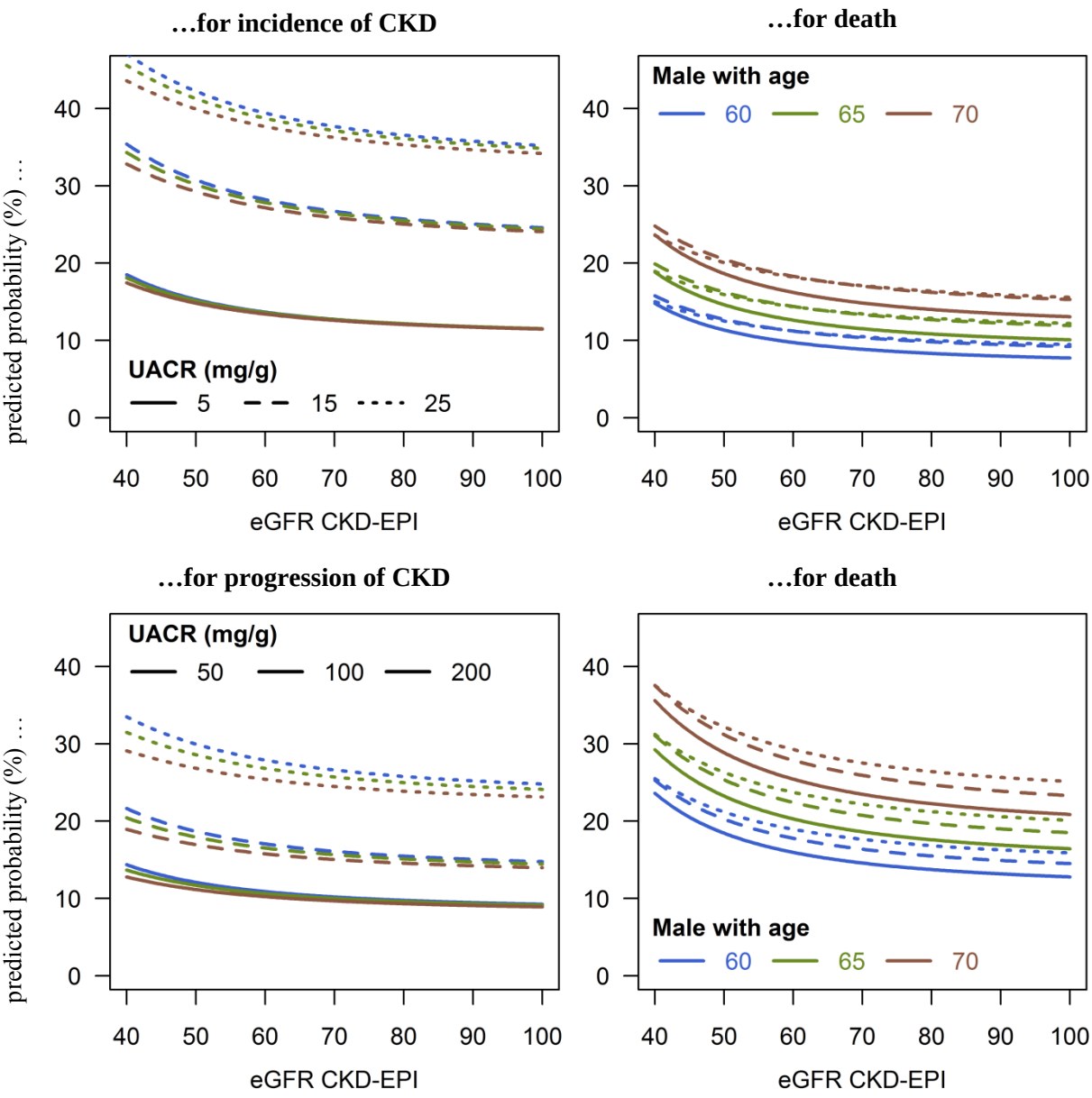
Calibration plots for the laboratory (first row) and the clinical model (second row) for each outcome state: alive without incidence or progression of CKD, alive with incidence or progression of CKD, and death based on internally validated estimates of the development cohort and on the validation cohort. Calibration plots depict the agreement of predicted probabilities and observed frequencies of the prediction model (continuous line) and for comparison a perfect prediction model (dashed line). Distribution of predicted probabilities is indicated by groups of participants (points). On the bottom the distribution of predicted probabilities for individual participants in the development cohort is depicted (vertical lines). Vertical lines upward represent participants with the outcome of the respective column; lines downwards represent participants from the other two outcome states.

Abbreviations: CKD, (incidence or progression of) chronic kidney disease.

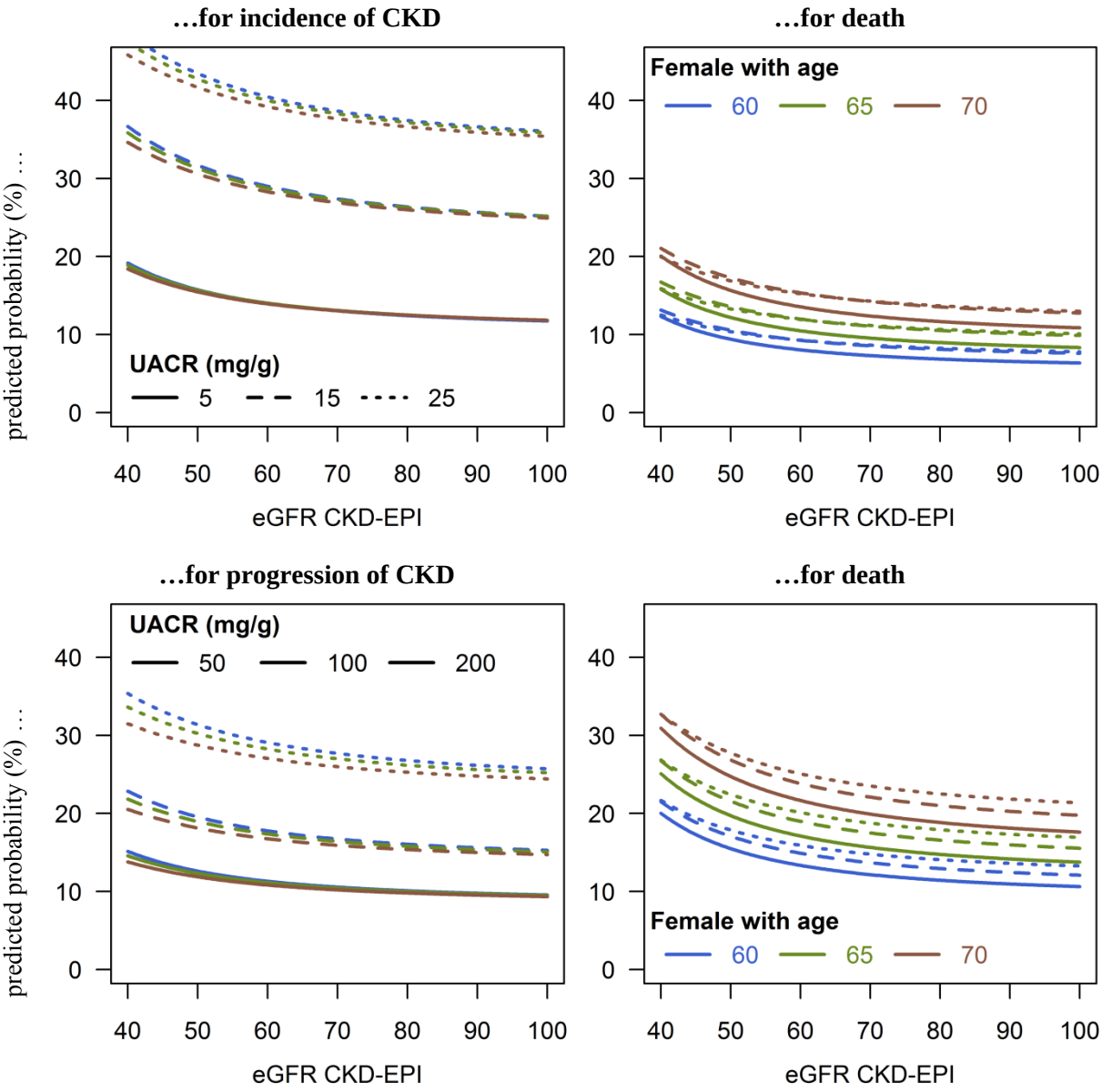
Supplemental Figure 6. Predicted Probabilities

for incidence or progression of chronic kidney disease after 5.5 years and death within 5.5 years computed by the laboratory model.

Supplemental Figure 6a. Hypothetical male individual aged 60, 65 or 70.



Supplemental Figure 6b. Hypothetical female individual aged 60, 65 or 70.



Supplemental Table 2. Prediction Equations

Supplemental Table 4a: Laboratory Prediction Equation

The laboratory prediction models includes the following predictors: UACR (mg/g), eGFR CKD-EPI (ml/min per 1.73m^2), gender and age.

Range of continuous predictors

Predictor	Minimum	Maximum
UACR (mg/g)	0.75	292.15
eGFR CKD-EPI (ml/min per 1.73m^2)	26.55	111.15
Age (years)	55	84

Prediction equation

Outcome state		Predictor	Transformation or coding	Shrunken coefficients
CKD	**Intercept**			-0.6727
	Predictors	**Albuminuria stage**[1]	microalbuminuria = 1	-0.1809
		d-UACR$_{tp}$	(d-UACR$_{tp}$+0.1)/5	-5.4497
			[(d-UACR$_{tp}$+0.1)/5]^3	4.7267
		eGFR	[(eGFR+0.1)/120]$^{-2}$	0.0890
		Gender	female = 1	0.0081
		Age	[age+0.1)/90]	-0.2088
Death	**Intercept**			-5.3823
	Predictors	**Albuminuria stage**[1]	microalbuminuria = 1	0.5402
		d-UACR$_{tp}$	(d-UACR$_{tp}$+0.1)/5	-2.0727
			[(d-UACR$_{tp}$ +0.1)/5]^3	1.5816
		eGFR	[(eGFR+0.1)/120]$^{-2}$	0.1120
		Gender	female = 1	0.5981
		Age	[age+0.1)/90]	5.3485

[1] versus normoalbuminuria = 0.

The predicted risk of an individual can be computed in the following manner:

1) Compute 'd-UACR to progression' (d-UACR$_{tp}$)[1] from UACR (mg/g) at baseline:

$$\text{d-UACR}_{tp} = \ln\left(\text{cutpoint}/\text{UACR (mg/g)}\right), \text{ with}$$

$$\text{cutpoint} = \begin{cases} 30 & \text{if } 0 \le \text{UACR (mg/g)} < 30 \\ 300 & \text{if } 30 \le \text{UACR (mg/g)} < 300 \end{cases}.$$

[1] d-UACR$_{tp}$ was defined as the difference between the participant-specific cutpoint of developing a new micro- or macroalbuminuria and UACR at baseline on the log-scale. A participant-specific cutpoint was required because new micro- or macro-albuminuria was defined by crossing the cutpoints of 30 and 300 mg/g (3.4 and 33.9 mg/mmol), respectively.

2) Compute the linear predictor for incidence or progression of CKD $lp_r(x)$:

$$
\begin{aligned}
lp_r(x) &= \beta_{r0} + \beta_{r1}x_1 + \ldots + \beta_{rj}x_j = \\
&= -0.76727\text{-}0.1809 * \text{albuminuria} - \\
&\quad 5.4497 * \left[\left(\text{d-UACR}_{\text{tp}} + 0.1\right)/5\right] + 4.7267 * \left[\left(\text{d-UACR}_{\text{tp}} + 0.1\right)/5\right]^3 + \\
&\quad 0.0890 * \left[\left(\text{eGFR} + 0.1\right)/120\right]^{-2} + 0.0081 * \text{female} - 0.2088 * \left[\left(\text{age} + 0.1\right)/90\right]
\end{aligned}
$$

3) Compute the linear predictor for death $lp_d(x)$:

$$
\begin{aligned}
lp_d(x) &= \beta_{d0} + \beta_{d1}x_1 + \ldots + \beta_{rj}x_j = \\
&= -5.3823 + 0.5402 * \text{albuminuria} - \\
&\quad 2.0727 * \left[\left(\text{d-UACR}_{\text{tp}} + 0.1\right)/5\right] + 1.5816 * \left[\left(\text{d-UACR}_{\text{tp}} + 0.1\right)/5\right]^3 + \\
&\quad 0.1120 * \left[\left(\text{eGFR} + 0.1\right)/120\right]^{-2} + 0.5981 * \text{female} - 5.3485 * \left[\left(\text{age} + 0.1\right)/90\right]
\end{aligned}
$$

4) Conditional probabilities for each outcome *y* given the predictor vector *x* are

$$
P\left(y = \text{'alive w/o renal endpoint'} \mid x\right) = \frac{1}{1 + e^{lp_r(x)} + e^{lp_d(x)}},
$$

$$
P\left(y = \text{'alive with renal outcome'} \mid x\right) = \frac{e^{lp_r(x)}}{1 + e^{lp_r(x)} + e^{lp_d(x)}} \text{ and}
$$

$$
P\left(y = \text{'death'} \mid x\right) = \frac{e^{lp_d(x)}}{1 + e^{lp_r(x)} + e^{lp_d(x)}}.
$$

Supplemental Table 2b: Clinical Prediction Equation

The clinical prediction models includes the following predictors: UACR (mg/mmol), eGFR CKD-EPI (ml/min per 1.73m^2), duration of diabetes (years), glucose (mmol/L), fasting LDL (mmol/L) , waist circumference (cm), number of antihypertensive drugs, age (years), gender, race (European, Asian, or Other), and the comorbidities peripheral artery disease (i.e. PTA , limb or foot amputation; PAD), stroke/TIA, laser therapy for diabetic retinopathy, and MACE. Comorbidity MACE, i.e. major atherosclerotic cardiac events, was defined as myocardial infarction, stable or unstable angina, CABG surgery, or PTCA/atherectomy/PCI. For number of antihypertensive drugs a score between 0 and 5 was devised, with one point for each group (RAS-blocker, calcium-channel-blocker, alpha-blocker, beta-blocker or diuretics) from which drugs were prescribed.

Range of continuous predictors

Predictor	Minimum	Maximum
UACR (mg/g)	0.75	292.15
eGFR CKD-EPI (ml/min per 1.73m^2)	26.55	111.15
Age	55	84
Glucose (mmol/L)	3.081	19.5
Fasting LDL (mmol/L)	0.8	6.097
Duration of diabetes (years)	0.0060	64.9778
Waist circumference (cm)	64.02	139

Prediction equation

Outcome state		Predictor	Transformation or coding	Shrunken coefficients
CKD	**Intercept**			-0.7382
	Predictors	**d-UACR$_{tp}$**	(d-UACR$_{tp}$+0.1)/5	-4.8303
			[(d-UACR$_{tp}$ +0.1)/5]^3	4.4693
		eGFR CKD-EPI	[(egfr+0.1)/120]$^{-2}$	0.0775
		Albuminuria stage[1]	microalbuminuria = 1	-0.2217
		Age	(age+0.1)/90	0.7529
		PAD	yes = 1	0.3621
		Glucose	[(glucose+0.1)/20]$^{-1}$	-1.1451
			ln([(glucose+0.1)/20])*[(glucose+0.1)/20]$^{-1}$	-0.5042
		Number of antihypertensive drugs	(score from 0 to 5) / 5	0.7667
		Ethnic group[3]	Asian = 1	0.3094
			Other = 1	0.2216
		Fasting LDL	[(ldl+0.1)/10]$^{-2}$	0.0069
			[(ldl+0.1)/10]^2	1.3196
		Duration of diabetes	[ln(diabduration+0.003)+6]/12	0.3271
		Stroke/TIA[2]	yes = 1	0.0865
		Gender[4]	female = 1	-0.0216
		Waist circumference	[(waist+0.1)/140]$^{-2}$	0.1080
			[(waist+0.1)/140]^3	0.9222
		MACE[2]	yes = 1	-0.1053
		Laser therapy for diabetic retinopathy[2]	yes = 1	-0.1006

[1] versus normoalbuminuria = 0.
[2] versus no = 0.
[3] versus European = 0.
[4] versus male = 0.

Outcome state		Predictor	Transformation or coding	Shrunken coefficients
Death	**Intercept**			-5.2880
	Predictors	**d-UACR$_{tp}$**	(d-UACR$_{tp}$+0.1)/5	-1.7979
			[(d-UACR$_{tp}$ +0.1)/5]^3	1.5969
		eGFR CKD-EPI	[(egfr+0.1)/120]$^{-2}$	0.0990
		Albuminuria stage[1]	microalbuminuria = 1	0.4115
		Age	(age+0.1)/90	5.1527
		PAD[2]	yes = 1	0.6782
		Glucose	[(glucose+0.1)/20]$^{-1}$	-1.4714
			ln([(glucose+0.1)/20])*[(glucose+0.1)/20]$^{-1}$	-0.6895
		Number of antihypertensive drugs	(score from 0 to 5) / 5	0.2970
		Ethnic group[3]	Asian = 1	-0.0353
			Other = 1	0.1458
		Fasting LDL	[(ldl+0.1)/10]$^{-2}$	0.0063
			[(ldl+0.1)/10]^2	2.5752
		Duration of diabetes	[log(diabduration+0.003)+6]/12	0.3132
		Stroke/TIA[2]	yes = 1	0.3867
		Gender[4]	female = 1	-0.1904
		Waist circumference	[(waist+0.1)/140]$^{-2}$	0.2402
			[(waist+0.1)/140]^3	1.3245
		MACE[2]	yes = 1	0.1769
		Laser therapy for diabetic retinopathy[2]	yes = 1	-0.0198

[1] versus normoalbuminuria = 0.
[2] versus no = 0.
[3] versus European = 0.
[4] versus male = 0.

1) and 4) as in the laboratory model.

2) Compute the linear predictor for incidence or progression of CKD $lp_r(x)$:

$$
\begin{aligned}
lp_r(x) = {} & -0.7382 - 0.2217 * \text{albuminuria} - \\
& 4.8303 * \left[\left(\text{d-UACR}_{\text{tp}} + 0.1 \right) / 5 \right] + 4.4693 * \left[\left(\text{d-UACR}_{\text{tp}} + 0.1 \right) / 5 \right]^3 + \\
& 0.0775 * \left[\left(\text{eGFR} + 0.1 \right) / 120 \right]^{-2} + 0.7529 * \left[\left(\text{age} + 0.1 \right) / 90 \right] + \\
& 0.3621 * \text{PAD} - 1.14513015 * \left[\left(\text{glucose} + 0.1 \right) / 20 \right]^{-1} - \\
& 0.5042 * \left[\left(\text{glucose} + 0.1 \right) / 20 \right]^{-1} * \ln \left[\left(\text{glucose} + 0.1 \right) / 20 \right] + 0.7667 * \text{nDrugs} / 5 + \\
& 0.3094 * \text{Asian} + 0.2216 * \text{Other} + 0.0069 * \left[\left(\text{ldl} + 0.1 \right) / 10 \right]^{-2} \\
& 1.3196 * \left[\left(\text{ldl} + 0.1 \right) / 10 \right]^2 + 0.3271 * \left[\ln \left(\text{diabduration} + 0.003 \right) + 6 \right] / 12 + \\
& 0.0865 * \text{stroke} - 0.0216 * \text{female} + 0.1080 * \left[\left(\text{waist} + 0.1 \right) / 140 \right]^{-2} + \\
& 0.9222 * \left[\left(\text{waist} + 0.1 \right) / 140 \right]^3 - 0.1053 * \text{MACE} - 0.1006 * \text{laser}
\end{aligned}
$$

3) Compute the linear predictor for death $lp_d(x)$:

$$
\begin{aligned}
lp_d(x) = {} & -5.2880 + 0.4115 * \text{albuminuria} - \\
& 1.7979 * \left[\left(\text{d-UACR}_{\text{tp}} + 0.1 \right) / 5 \right] + 1.5969 * \left[\left(\text{d-UACR}_{\text{tp}} + 0.1 \right) / 5 \right]^3 + \\
& 0.0990 * \left[\left(\text{eGFR} + 0.1 \right) / 120 \right]^{-2} + 5.1527 * \left[\left(\text{age} + 0.1 \right) / 90 \right] + \\
& 0.6782 * \text{PAD} - 1.4714 * \left[\left(\text{glucose} + 0.1 \right) / 20 \right]^{-1} - \\
& 0.6895 * \left[\left(\text{glucose} + 0.1 \right) / 20 \right]^{-1} * \ln \left[\left(\text{glucose} + 0.1 \right) / 20 \right] + 0.2970 * \text{nDrugs} / 5 - \\
& 0.0353 * \text{Asian} + 0.1458 * \text{Other} + 0.0064 * \left[\left(\text{ldl} + 0.1 \right) / 10 \right]^{-2} \\
& 2.5752 * \left[\left(\text{ldl} + 0.1 \right) / 10 \right]^2 + 0.3132 * \left[\ln \left(\text{diabduration} + 0.003 \right) + 6 \right] / 12 + \\
& 0.3867 * \text{stroke} - 0.1904 * \text{female} + 0.2402 * \left[\left(\text{waist} + 0.1 \right) / 140 \right]^{-2} + \\
& 1.3245 * \left[\left(\text{waist} + 0.1 \right) / 140 \right]^3 + 0.1769 * \text{MACE} - 0.0198 * \text{laser}
\end{aligned}
$$

References

1. Vickers AJ, Elkin EB: Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 26:565-574, 2006

2. Levey AS, Coresh J, Balk E, Kausz AT, Levin A, Steffes MW, Hogg RJ, Perrone RD, Lau J, Eknoyan G: National Kidney Foundation practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Ann Intern Med* 139:137-147, 2003

3. Levey AS, Stevens LA, Schmid CH, Zhang Y, Castro III, Feldman HI, Kusek JW, Eggers P, Van Lente F, Greene T, Coresh J: A new equation to estimate glomerular filtration rate. *Ann Intern Med* 150:604-612, 2009

4. Biesheuvel CJ, Vergouwe Y, Steyerberg EW, Grobbee DE, Moons KGM: Polytomous logistic regression analysis could be applied more often in diagnostic research. *J Clin Epidemiol* 61:125-134, 2008

5. Royston P, Sauerbrei W: Multivariable model-building. A pragmatic approach to regression analysis based on fractional polynomials for modelling continuous variables, Chichester, John Wiley & Sons, Ltd., 2008

6. Akaike H: A new look at the statistical model identification. *IEEE T Automat Contr* 19:716-723, 1974

7. Steyerberg EW, Harrell FE, Borsboom GJJM, Eijkemans MJC, Vergouwe Y, Habbema JD: Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol* 54:774-781, 2001

8. Schemper M: Predictive accuracy and explained variation. *Stat Med* 22:2299-2308, 2003

9. Heinze G, Schemper M: Comparing the importance of prognostic factors in Cox and logistic regression using SAS. *Comput Metho Prog Bio* 71:155-163, 2003

10. Schemper M: The relative importance of prognostic factors in studies of survival. *Stat Med* 12:2377-2382, 1993