Supplementary Material: Development of a web-based, adaptive clinical prediction tool for kidney replacement therapy in children with chronic kidney disease

Table of contents

Cross-validation and external validation methods

Cross-validation

We conducted 10-fold cross-validation of the entire model-building process[1,2]. In this procedure, 10 random samples comprising 90% of the data were used to develop the elementary and final models with the concatenation of the remaining 10% per fold used to evaluate model fit. In the model development stage of the cross-validation, we executed the same three-step process described above so the results from the functional forms of elementary model predictors, the random survival forest (RSF), and best subset selection were allowed to vary.

For the 10 folds, model parameters were estimated in each set of 90% of the data and yielded $\hat{S}_k$ as the estimate of the survival function for the *k*th fold (for *k*= 1 to 10). For the 10% excluded in that fold, standardized times were calculated as $w_{ki} = -\log(\hat{S}_k(t_{ki}))$ for the *i*th individual in the *k*th fold. If the prediction is valid, the full concatenation of the excluded data are expected to correspond to a sample subjected to censoring from the standard exponential distribution (e^{-t}). To explore this, we graphically depicted the survival function of w_{ki} overlayed with the standard exponential survival function.[3]

External validation

To externally validate our model, we assessed the calibration and discrimination in previously published data from the European Study Consortium for Chronic Kidney Disorders Affecting Pediatric Patients (ESCAPE) study[4]. Among 270 ESCAPE participants, we used the first available visit with complete data on U25 eGFR, proteinuria and diagnosis to evaluate the elementary model. To account for regional differences in KRT initiation, the outcome was time to KRT or U25eGFR of 20 ml/min|1.73m^2 which was interpolated between two visits when the second visit had U25eGFR < 20 ml/min|1.73m^2.

We conducted the Greenwood-Nam-D'Agostino goodness-of-fit test (calibration) for 5-year risk, calculated the c-statistic (discrimination), and evaluated the residual standardized times relative to the standard exponential (a measure of calibration) as described in the cross-validation methods.

References

1. Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med*. 2000;19(4):453-473.

2. Steyerberg EW, Harrell FE, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol*. 2001;54(8):774-781.

3. Ng DK, Matheson MB, Warady BA, Mendley SR, Furth SL, Muñoz A. Incidence of Initial Renal Replacement Therapy Over the Course of Kidney Disease in Children. *Am J Epidemiol*. 2019;188(12):2156-2164. doi:10.1093/aje/kwz220

4. Furth SL, Pierce C, Hui WF, et al. Estimating Time to ESRD in Children With CKD. *Am J Kidney Dis*. 2018;71(6):783-792. doi:10.1053/j.ajkd.2017.12.011

Supplementary Table S1. List of variables evaluated in random survival forest analysis to identify potential predictive candidates in subsequent parametric survival models to predict time to kidney replacement therapy. Continuous time-varying variables denoted by * were included as level at study origin (visit 2) and annualized change from the previous year (visit 1 to visit 2). Categorical time-varying variables denoted by † are included as presence at study origin (visit 2) and incident/initiated, persistent/continued, resolved/discontinued from the previous year (visit 1 to visit 2). Total time-fixed and time-varying variables included 172 predictors.

Sociodemographic	***Laboratory markers***	***Medications***
Sex	Sodium*	Any antihypertensive†
Age	Potassium*	Diuretic†
Maternal education	Chloride*	ACE/ARB†
Household income	CO2*	Active vitamin D†
Family history of kidney disease (first, second or third degree)	Acidosis†	Inactive vitamin D†
CKD severity	Glucose*	Phosphate binder†
U25eGFR*	Albumin*	Alkali therapy†
UPCR*	Hypoalbuminemia†	Growth hormone†
Years with CKD	Calcium*	Erythropoietin stimulating agent†
Anemia†	Phosphate*	Iron supplement†
change in hemoglobin z-score	Calcium-phosphate product*	Lipid lowering†
BP stage*	Elevated Ca×P†	Potassium binder†
Growth and development	Leukocyte count*	Corticosteroid†
Height z-score*	Erythrocyte count*	Immunosuppressant†
Short stature†	Platelets*	Bladder med†
BMI z-score*	Hematocrit*	Laxative†
Mid-upper arm circumference z-score*	Mean corpuscular hemoglobin concentration*	Antacid†
	Mean corpuscular hemoglobin*	Antibiotic†
Birth history	Mean corpuscular volume*	Asthma/allergy†
Low birth weight	Red cell distribution width*	CNS stimulant†
Premature		Nutrition supplement†
Small for gestational age	Triglycerides	Vitamin supplement†
	High density lipoprotein (HDL)	
	Low density lipoprotein	
	Total cholesterol	
	Non-HDL	
	Dyslipidemia	

* For continuous variables, the annualized change of the variable between visit 1 and visit 2 (study origin) was calculated as calculated as:

$$(Variable_{v2} - Variable_{v1}) \,/\, (Years\ from\ V1\ to\ V2)$$

† For binary variables, including comorbidities (e.g., anemia, metabolic acidosis) and therapies (e.g., ACE/ARB, alkali therapy), the time-varying variable was classified by four categories defined as:

Visit 1	Visit 2 (study origin)	Comorbidity presence	Therapy use
No	No	Absent	None
No	Yes	Incident	Initiated
Yes	No	Resolved	Discontinued
Yes	Yes	Persistent	Continued

Supplementary Table S2. Distribution of specific chronic kidney disease diagnoses represented in the study population (n= 890).

Non-glomerular/HUS kidney disease diagnosis	n= 704	Glomerular kidney disease diagnosis	n= 186
Aplastic/hypoplastic/dysplastic kidneys	23.4% (165)	Focal segmental glomerulosclerosis	35.0% (65)
Obstructive uropathy	22.4% (158)	Systemic immunological disease (including systemic Lupus erythematosus)	17.7% (33)
Reflux nephropathy	16.6% (117)	Chronic glomerulonephritis	10.8% (20)
Hemolytic uremic syndrome (HUS)	6.8% (48)	IgA Nephropathy (Berger's)	7.5% (14)
Congenital Urologic Disease (Bilateral Hydronephrosis)	6.0% (42)	Familial nephritis (Alport's)	6.5% (12)
Non-glomerular Other	5.5% (39)	Membranoproliferative glomerulonephritis type I	5.4% (10)
Polycystic kidney disease (Autosomal recessive)	4.0% (28)	Glomerular other	4.8% (9)
Renal infarct	3.3% (23)	Henoch schonlein nephritis	3.8% (7)
Cystinosis	1.7% (12)	Idiopathic crescentic glomerulonephritis	3.2% (6)
Pyelonephritis/Interstitial nephritis	1.6% (11)	Membranous nephropathy	2.2% (4)
Medullary cystic disease/Juvenile nephronophthisis	1.4% (10)	Congenital nephrotic syndrome	1.6% (3)
Perinatal Asphyxia	1.4% (10)	Membranoproliferative glomerulonephritis type II	1.1% (2)
Syndrome of agenesis of abdominal musculature	1.1% (8)	Sickle cell nephropathy	0.5% (1)
VACTERL (VATER) Syndrome	1.0% (7)		
Branchio-oto-Renal Disease/Syndrome	1.0% (7)		
Wilms' tumor	0.9% (6)		
Methylmalonic Acidemia	0.9% (6)		
Polycystic kidney disease (Autosomal dominant)	0.7% (5)		
Oxalosis	0.3% (2)		

Supplementary Table S3. AIC values as a measure of test error from different functional forms of GFR, proteinuria (UPCR) and interactions with diagnosis (glomerular with no HUS vs. non-glomerular with HUS) as modifiers of the location β parameter in generalized gamma parametric survival model. **Bold** indicated lowest AIC. The AIC of the null model with no covariates (i.e., predictors) was 2021.164.

Functional form of GFR	Functional form of UPCR	Modifiers			
		GFR UPCR	GFR×Diagnosis UPCR	GFR UPCRxDiagnosis	GFRxDiagnosis UPCRxDiagnosis
Continuous	Continuous	1492.528	1490.358	1487.446	1488.356
Natural spline (knot at 45)	Continuous	1486.931	1480.883	1481.642	1480.863
Linear spline (knot at 45)	Continuous	1486.270	1480.123	1480.768	1480.225
Continuous	Natural spline (knot at 0.5)	1483.722	1483.241	1483.327	1484.559
Natural spline (knot at 45)	Natural spline (knot at 0.5)	1479.916	1476.965	1479.498	1479.106
Linear spline (knot at 45)	Natural spline (knot at 0.5)	1479.402	1476.493	1478.753	1478.872
Continuous	Linear spline (knot at 0.5)	1484.409	1483.604	1483.328	1484.608
Natural spline (knot at 45)	Linear spline (knot at 0.5)	1480.263	1476.778	1479.309	1479.286
Linear spline (knot at 45)	Linear spline (knot at 0.5)	1479.907	**1476.477**	1478.661	1479.052

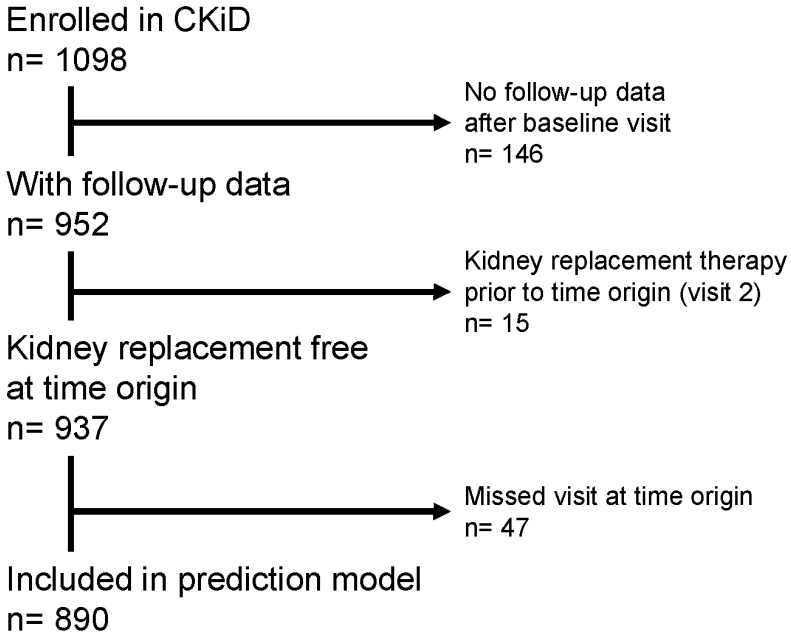
Supplementary Table S4. Ranking of variables from random survival forest according to top 20 minimal depth of maximal subtree classification and top 20 variable importance.

Variable	Min Depth of Max Subtree	Min Depth Rank	Importance	Importance Rank
U25eGFR	5.503	#1	0.0532	#1
UPCR	5.914	#2	0.0273	#2
Albumin	8.357	#3	0.0041	#6
Ann change in UPCR	8.465	#4	0.0046	#5
Persistent anemia	10.582	#12	0.0080	#3
Potassium	9.095	#6	0.0035	#9
Anemia	11.299	#14	0.0058	#4
Ann change in U25eGFR	10.117	#8	0.0037	#7
Chloride	9.145	#7	0.0036	#8
Phosphate	8.780	#5	0.0026	#15
Hematocrit	10.482	#10	0.0033	#10
Red blood cell count	10.525	#11	0.0027	#13
CO_2	10.325	#9	0.0007	#21
ESA initiation	11.393	#15	0.0027	#14
ESA use	12.024	#21	0.0032	#11
Calcium × Phosphate	11.659	#17	0.0011	#18
Active Vitamin D	12.717	#28	0.0028	#12
Hypoalbuminemia	12.427	#23	0.0016	#16
Non-HDL	11.654	#16	0.0006	#26
ACE/ARB discontinuation	12.468	#25	0.0014	#17
Blood pressure stage	12.013	#20	0.0006	#25
Calcium	10.720	#13	0.0003	#43
Red cell distribution width	11.680	#18	0.0004	#34
LDL	12.001	#19	0.0003	#36
Persistent active vitamin D use	13.238	#41	0.0010	#19
Persistent hypoalbuminemia	13.356	#48	0.0007	#20

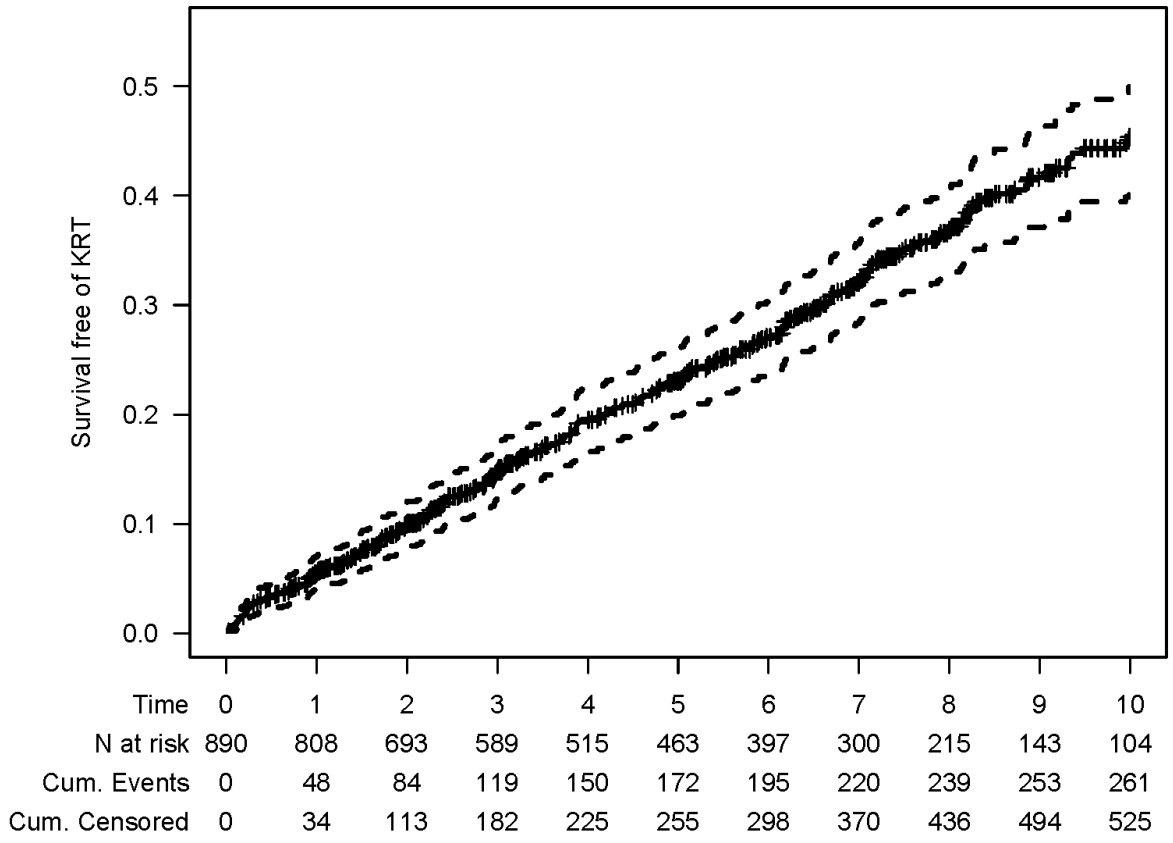
Supplementary Table S5. Baseline descriptive characteristics of the ESCAPE cohort (n= 270) used for external validation.

Characteristic	**Median [IQR] or n (%)**
Demographic and height	
Age, years	11.9 [8.6, 14.9]
Male sex	158 (59%)
Height, m	1.43 [1.25, 1.61]
Kidney disease characteristics	
Glomerular non-HUS diagnosis	16 (6%)
Serum creatinine, mg/dL	1.5 [1.2, 2.0]
U25eGFR, ml/min|1.73 m²	36 [27, 44]
Urine protein/creatinine ratio, mg/mg	0.33 [0.10, 0.89]
Urine protein/creatinine > 2	32 (12%)

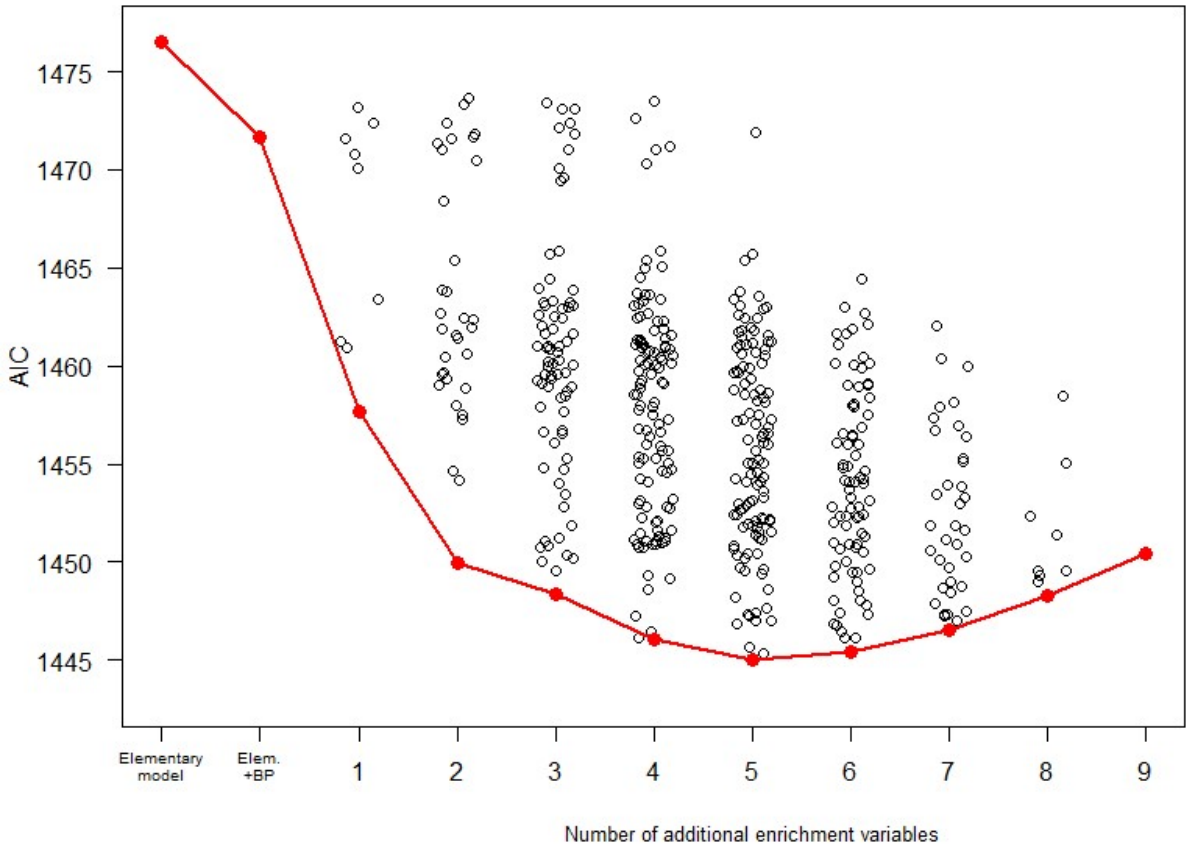
Supplementary Figure S1. Cohort flow diagram of those with no follow-up time, prevalent kidney replacement therapy event prior to time origin, and those with missed visits at the time origin.



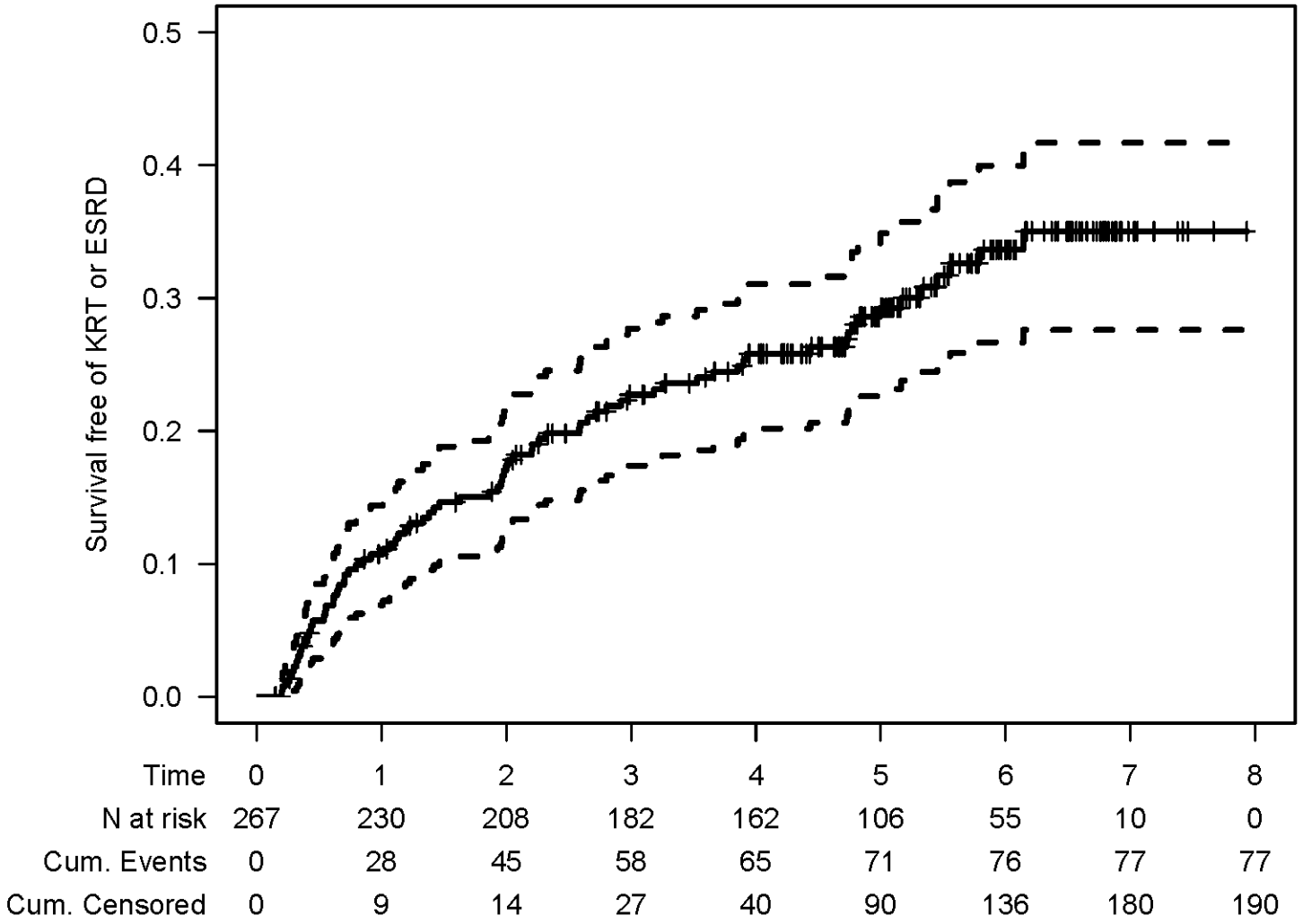
Supplementary Figure S2. Kaplan-Meier estimates of cumulative incidence of kidney replacement therapy in the CKiD cohort with 95% confidence intervals (discontinuous lines) and censored observations indicated by vertical ticks.



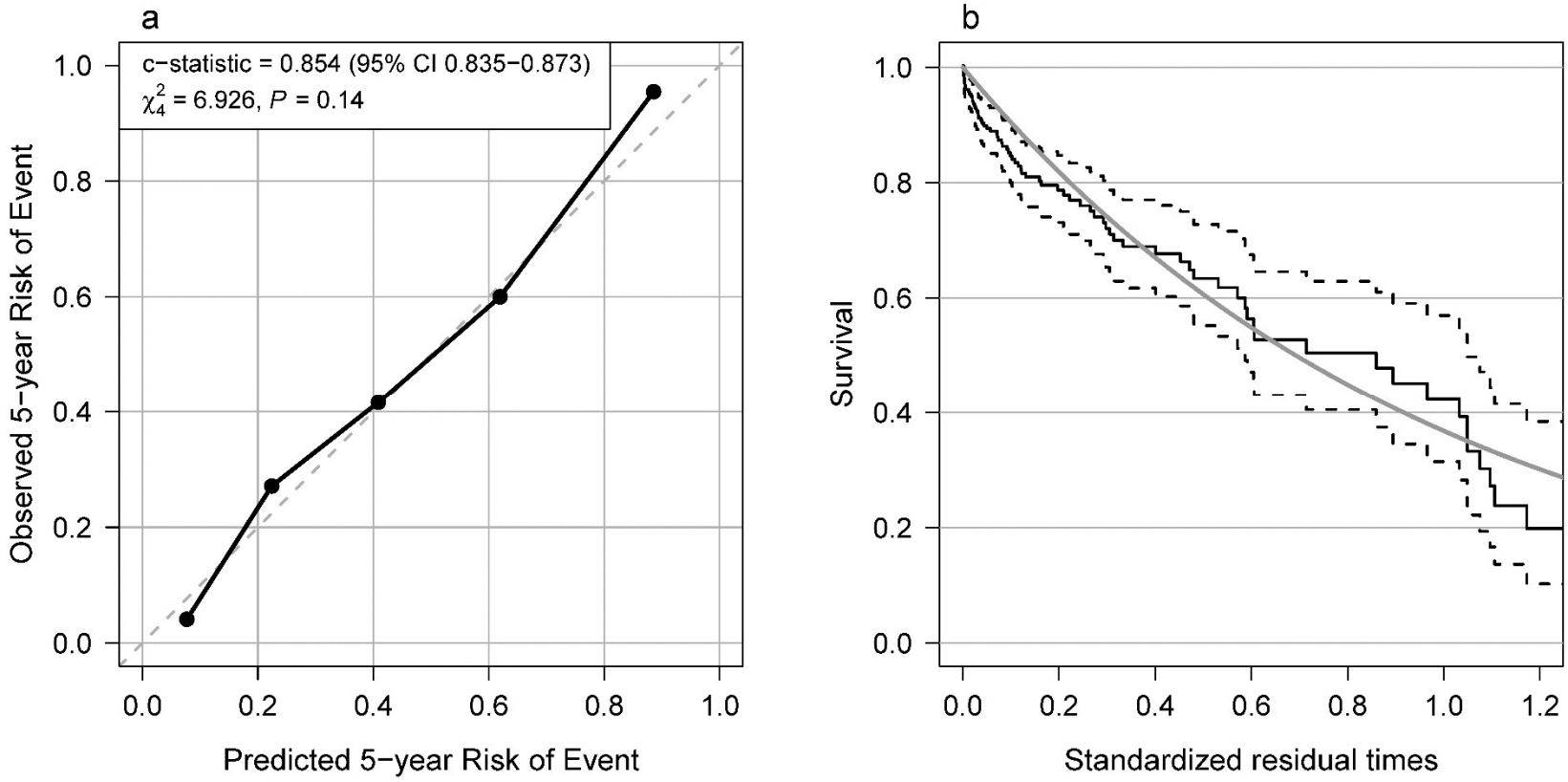
Supplementary Figure S3. Results from best subset selection of top 9 variables identified from random survival forest analysis presenting Akaike information criterion (AIC) on best subset models with additional variables (2^9= 512, including the elementary model as the null model). Red dots indicate the lowest AIC within each set of additional enrichment variables. The lowest AIC of all models was identified in the model with 5 additional enrichment variables and this model corresponds to the enriched model in Table 2.



Supplementary Figure S4. Kaplan-Meier estimates of cumulative incidence of ESKD in the ESCAPE cohort with 95% confidence intervals (discontinuous lines) and censored observations indicated by vertical ticks.



Supplementary Figure S5. Results and interpretation from external validation of elementary model using data from the ESCAPE study including calibration plot depicting observed risk on predicted risk from at 5-year risk of kidney replacement therapy or eGFR < 20 ml/min|1.73m^2 and survival function of standardized residual times for participants. The calibration plot demonstrates close correspondence between observed and predicted 5-year risk and the survival function aligns closely with the expected standard exponential for strong model fit.



Supplementary Figure S5 presents data from the external validation of the elementary model using European ESCAPE data. The left panel presents the Greenwood-Nam-D'Agostino test for 5-year risk comparing the observed 5-year risk (y-axis) to the predicted 5-year risk (x-axis) using bins of participants of similar predicted 5-year risk with at least 5 observed events per bin. There were no significant differences between the observed risk and predicted risk (χ_4^2 = 6.926; p= 0.140), the calibration slope was 0.996 (95%CI: 0.903, 1.088, p= 0.90), and the c-statistic was 0.854 (95%CI: 0.835, 0.873) indicating strong discrimination. The right panel presents the standardized times of residuals which were congruent with the standard exponential function depicted in green.

Supplementary Appendix S1. R code to generate 10th, 25th and 50th percentile of kidney replacement therapy risk from elementary model, partially enriched models 1 to 4, and the fully enriched model using example profile provided in the Results section of the paper.

```
# Demonstration of CKiD risk prediction calculator
# install.packages("flexsurv") ## Install flexsurv package if not in library
already
library(flexsurv)

### Step 1: Input the profile of a hypothetical patient
## Necessary components for the Elementary model described in Table 2 of
paper
kid.gfr <- 60 # GFR in ml/min|1.73m²
kid.upc <- 0.8 # Urine protein-creatinine ratio in mg/mgCr; check units
carefully
kid.glomdx <- 1 # Glomerular, non-HUS diagnosis = 1; Non-glomerular or HUS
diagnosis = 0

# Components for the Partially Enriched and Enriched models described in
Table 2 of paper
kid.highbp <- 1 # High blood pressure = 1; normal blood pressure = 0
kid.anemia <- 1 # Anemia as low hemoglobin = 1; normal hemoglobin = 0
kid.albumin <- 4.5 # Serum albumin in g/dL; check units carefully
kid.chloride <- 105 # Chloride in mmol/dL; check units carefully
kid.co2 <- 22 # Bicarbonate in mmol/L; check units carefully
kid.prevgfr <- 67 # Previous GFR from about one year ago, ml/min|1.73m²
kid.prevgfryrs <- 1 #GFR was 48 when checked 1 years ago, can be in decimals;
e.g., a value of 0.95 is 0.95 years earlier than the current GFR input date

### Step 2: Use models described in Table 2 to output the times by which 10%,
25% and 50% will experience the outcome (KRT) based on the profile in Step 1.
## Times are expressed as years (note: 1 month= 0.0833 years)
## Uses the function qgengamma in the flexsurv package which provides values
of quantiles for the specified generalized gamma distribution
  # the first argument specifies the desired quantiles (10th, 25th and 50th
or median) which are measures of predicted risk from the prediction tool
  # The next three arguments correspond to coefficients reported in the paper
that modify Beta (mu), sigma (sigma) and kappa (Q)
percentile.header <- c("10th pctile", "25th pctile", "50th pctile") # To
improve readability of output, label header of the first three times with
corresponding percentiles

# Elementary model
times.elem <- round(qgengamma(c(0.1,0.25,0.5), mu = 2.8624 +
2.0868*log(kid.gfr/45) - 1.0758*max(c(0,log(kid.gfr/45))) - 0.1532*kid.glomdx
+ 1.0386*kid.glomdx*log(kid.gfr/45) -
1.2154*kid.glomdx*max(c(0,log(kid.gfr/45))) - 0.1959*log(kid.upc/0.5) -
0.3146*max(c(0,log(kid.upc/0.5))), sigma = 0.7440 + 0.2627*kid.glomdx, Q =
0.4253 - 0.0005*kid.glomdx),3)
print(as.data.frame(times.elem,row.names=percentile.header))

# Partially enriched model 1: elementary + hypertension + deltaGFR only
times.pe1 <- round(qgengamma(c(0.1,0.25,0.5), mu = 2.8993 +
1.9511*log(kid.gfr/45) - 1.0091*max(c(0,log(kid.gfr/45))) - 0.0718*kid.glomdx
+ 0.8132*kid.glomdx*log(kid.gfr/45) -
```

```
1.0010*kid.glomdx*max(c(0,log(kid.gfr/45))) - 0.2123*log(kid.upc/0.5) -
0.2721*max(c(0,log(kid.upc/0.5))) - 0.2229*kid.highbp +
0.4349*(log(kid.gfr/kid.prevgfr)/kid.prevgfryrs), sigma = 0.7454 +
0.1744*kid.glomdx, Q = 0.3888 + 0.2204*kid.glomdx),3)
print(as.data.frame(times.pe1,row.names=percentile.header))

# Partially enriched model 2: elementary + hypertension + anemia only
times.pe2 <- round(qgengamma(c(0.1,0.25,0.5), mu = 2.9967 +
1.9769*log(kid.gfr/45) - 1.1370*max(c(0,log(kid.gfr/45))) - 0.0511*kid.glomdx
+ 0.9807*kid.glomdx*log(kid.gfr/45) -
1.0714*kid.glomdx*max(c(0,log(kid.gfr/45))) - 0.2127*log(kid.upc/0.5) -
0.2566*max(c(0,log(kid.upc/0.5))) - 0.2337*kid.highbp - 0.3617*kid.anemia,
sigma = 0.7441 + 0.2521*kid.glomdx, Q = 0.3513 + 0.1023*kid.glomdx),3)
print(as.data.frame(times.pe2,row.names=percentile.header))

# Partially enriched model 3: elementary + hypertension + albumin + co2 +
chloride only
times.pe3 <- round(qgengamma(c(0.1,0.25,0.5), mu = 3.7450 +
2.1789*log(kid.gfr/45) - 1.3874*max(c(0,log(kid.gfr/45))) - 0.1071*kid.glomdx
+ 0.6812*kid.glomdx*log(kid.gfr/45) -
0.5795*kid.glomdx*max(c(0,log(kid.gfr/45))) - 0.1720*log(kid.upc/0.5) -
0.2102*max(c(0,log(kid.upc/0.5))) - 0.2242*kid.highbp + 0.3802*kid.albumin +
0.0005*kid.co2 - 0.0235*kid.chloride, sigma = 0.7204 + 0.2426*kid.glomdx, Q =
0.4467 + 0.0768*kid.glomdx),3)
print(as.data.frame(times.pe3,row.names=percentile.header))

# Partially enriched model 4: elementary + hypertension + anemia + albumin +
co2 + chloride corresponds to an enriched "Current day" model
times.pe4 <- round(qgengamma(c(0.1,0.25,0.5), mu = 4.4095 +
2.0944*log(kid.gfr/45) - 1.4511*max(c(0,log(kid.gfr/45))) - 0.0231*kid.glomdx
+ 0.7360*kid.glomdx*log(kid.gfr/45) -
0.6282*kid.glomdx*max(c(0,log(kid.gfr/45))) - 0.1984*log(kid.upc/0.5) -
0.1803*max(c(0,log(kid.upc/0.5))) - 0.2339*kid.highbp - 0.3162*kid.anemia +
0.3168*kid.albumin - 0.0074*kid.co2 - 0.0248*kid.chloride, sigma = 0.7242 +
0.2172*kid.glomdx, Q = 0.4163 + 0.1367*kid.glomdx),3)
print(as.data.frame(times.pe4,row.names=percentile.header))

# Fully enriched model: elementary + hypertension + anemia + albumin + co2 +
chloride + deltaGFR corresponds to "Current day" + deltaGFR (longitudinal)
model which is fully enriched
times.full <- round(qgengamma(c(0.1,0.25,0.5), mu = 4.4493 +
1.9837*log(kid.gfr/45) - 1.4158*max(c(0,log(kid.gfr/45))) + 0.0268*kid.glomdx
+ 0.6275*kid.glomdx*log(kid.gfr/45) -
0.5253*kid.glomdx*max(c(0,log(kid.gfr/45))) - 0.1968*log(kid.upc/0.5) -
0.1759*max(c(0,log(kid.upc/0.5))) - 0.2282*kid.highbp - 0.3033*kid.anemia +
0.3172*kid.albumin - 0.0071*kid.co2 - 0.0255*kid.chloride +
0.3934*(log(kid.gfr/kid.prevgfr)/kid.prevgfryrs), sigma = 0.7262 +
0.1390*kid.glomdx, Q = 0.3914 + 0.2787*kid.glomdx),3)
print(as.data.frame(times.full,row.names=percentile.header))
```

Supplementary Appendix S2. Content from online clinical calculator describing model purpose, use, strengths and limitations designed for clinicians.

Website: https://ckid-gfrcalculator.shinyapps.io/CKiD_KRT_Risk/

Pediatric CKiD Kidney Replacement Therapy Risk Calculator

This calculator provides estimated times to when a patient diagnosed with CKD may require kidney replacement therapy (KRT). These estimates are based on a set of variables that are routinely measured on patients. It is intended to aid clinicians in managing kidney disease for pediatric patients below the age of 18 years.

The tool requires information on current markers of kidney health: CKD diagnosis, GFR and proteinuria. Additional variables include blood pressure category, anemia, serum albumin, serum chloride, serum bicarbonate and GFR from approximately one year ago. Not all additional variables are required for a valid estimate of estimated time to kidney replacement therapy, but additional variables will yield improved prediction.

Please note these are estimated times to needing KRT according to information provided. Please interpret the estimated times within the context of the patient's full clinical profile.

Frequently Asked Questions

About the calculator

What is the purpose of this calculator?

The calculator is designed to help clinicians estimate the time to initiation of kidney replacement therapy (transplant or dialysis) for a pediatric patient with diagnosed kidney disease. The calculation is based on the patient's clinical variables.

How was this calculator developed?

This calculator is based on data from a study of North American children and adolescents with a pediatric diagnosis of kidney disease (specifically, the Chronic Kidney Disease in Children study, called CKiD). Statistical learning methods (random survival forests, best subset regression) were used to identify candidate predictors of time to kidney replacement therapy and determine the most predictive survival models using the generalized gamma distribution. The model was internally validated using cross-validation.

How does this calculator compare with other online risk calculators?

Other risk calculators may be based on data from adults and have not been well-validated in children. Some risk calculators are based on GFR and urine protein levels only. This calculator uses multiple variables that are commonly collected at clinical visits. In addition, other risk calculators predict time to kidney replacement therapy or a 50% decline in GFR. This calculator is based on the occurrence of kidney replacement therapy alone. Lastly, this calculator uses patient history of GFR (change in GFR over the course of a year) to better estimate when kidney replacement therapy is expected to occur.

Does this calculator collect personal or private information?

No, this calculator does not collect or store any personal information or clinical data. The calculator does not ask for name, date of birth or any information of that sort.

Using the calculator

Is this calculator applicable to any patient?

This calculator is designed to be used by healthcare providers for pediatric patients (17 years of age or younger) with a diagnosis of chronic kidney disease (non-glomerular or glomerular etiology) and GFR < 90 ml/min|1.73m2 and is expected to be generalizable to North American patients who fit this profile.

What if the patient does not have data (e.g., not measured, or not asked) for one or more questions?

This calculator is unique in that the only requirements for KRT prediction are diagnosis, GFR and proteinuria level. Additional data including BP category, anemia, serum biomarkers and change in GFR over 1 year will improve the prediction, but are not necessary for a valid estimate.

Does previous GFR have to be from exactly one year ago?

No, the variables do not have to be from a visit exactly one year ago, but approximately one year ago. Clinical information from one year ago is used to indicate patient history. The calculator is based on models using data from clinical visits in the CKiD study approximately one year apart, so it is recommended to use that same time frame (plus or minus a few months). If you have information available from a patient's clinical visit approximately one year ago, you can use that information in the calculator to estimate time to kidney replacement therapy, but it is not necessary.

Interpreting the results

Do the results identify when exactly the patient will have renal replacement therapy?

No, the calculator cannot predict exactly when or if a particular patient will require renal replacement therapy. There are many factors that contribute to disease progression and the calculator does not capture all of them. These are estimates based on a large, representative cohort of children and adolescents with chronic kidney disease. Since the calculator provides estimates of time, variability of outcomes is shown for when 50%, 25% and 10% of patients with a similar profile will have KRT.

How should I interpret the results?

The calculator will estimate the time when 50%, 25% and 10% of patients with the same clinical profile as your patient will have renal replacement therapy. For example, if the results state 'Among patients with the same profile, 25% will have KRT by 6 years', this means that among patients with similar clinical profiles to your patient, 25% will be expected to have KRT within 6 years and 75% will be expected to have KRT after 6 years.

Supplementary Table S6. List of principal site investigators of the Chronic Kidney Disease in Children (CKiD) cohort study.

Study Investigator(s)	Institution	City	State/Province
Sahar Fathallah-Shaykh, MD	University of Alabama at Birmingham (Children's of Alabama)	Birmingham	AL
Anjali Nayak, MD; Martin Turman, MD	Phoenix Children's Hospital	Phoenix	AZ
Tom Blydt-Hansen, MD, FRCPC	British Columbia Children's Hospital	Vancouver	British Columbia, Canada
Cynthia Wong, MD; Steve Alexander, MD	Stanford University	Palo Alto	CA
Ora Yadin, MD	University of California – Los Angeles (UCLA)	Los Angeles	CA
Elizabeth Ingulli, MD; Robert Mak, MD, PhD	University of California – San Diego (UCSD)	San Diego	CA
Cheryl Sanchez-Kazi, MD	Loma Linda University	Loma Linda	CA
Asha Moudgil, MD	Children's National Medical Center	Washington	DC
Samina Muneeruddin, MD	Nemours Hospital for Children- Delaware Valley	Wilmington	DE
Carolyn Abitbol, MD; Marissa DeFrietas, MD; Chryso Katsoufis, MD; Wacharee Seeherunvong, MD	University of Miami	Miami	FL
Larry Greenbaum, MD, PhD	Children's Healthcare of Atlanta / Emory University	Atlanta	GA
Lyndsay Harshman, MD	University of Iowa	Iowa City	IA
Priya Verghese, MD	Ann & Robert H. Lurie Children's Hospital of Chicago	Chicago	IL
Sonia Krishnan, MD	University of Illinois at Chicago	Chicago	IL
Amy Wilson, MD	Riley Hospital for Children at Indiana University Health	Indianapolis	IN
Stefan Kiessling, MD; Margaret Murphy, PhD	University of Kentucky	Lexington	KY
Siddharth Shah, MD, Janice Sullivan, MD; Sushil Gupta, MD	University of Louisville (Novak Center for Children's Health)	Louisville	KY
Samir El-Dahr, MD; Stacy Drury, MD	Tulane University	New Orleans	LA
Nancy Rodig, MD	Boston Children's Hospital	Boston	MA
Allison Dart, MD MSc, FRCPC	University of Manitoba (The Children's Hospital	Winnipeg	Manitoba, Canada

	Research Institute of Manitoba)		
Meredith Atkinson, MD	Johns Hopkins University (Johns Hopkins Children's Center)	Baltimore	MD
Arlene Gerson, PhD		Baltimore	MD
Tej Matoo, MD	Children's Hospital of Michigan / Wayne State University	Detroit	MI
Zubin Modi, MD	University of Michigan	Ann Arbor	MI
Jason Thomas, MD	Spectrum Health Hospitals / Helen DeVos Children's Hospital	Grand Rapids	MI
Bradley Warady, MD; Rebecca Johnson, PhD	Children's Mercy Hospital	Kansas City	MO
Vikas Dharnidharka, MD	Washington University in St. Louis (St. Louis Children's Hospital)	St. Louis	MO
Stephen Hooper, PhD	University of North Carolina	Chapel Hill	NC
Susan Massengill, MD	Levine Children's Hospital	Charlottesville	NC
Liliana Gomez-Mendez, MD	East Carolina University	Greenville	NC
Matthew Hand, DO	Dartmouth-Hitchcock Medical Center	Lebanon	NH
Joann Carlson, MD	Rutgers-Robert Wood Johnson Medical School	New Brunswick	NJ
Craig Wong, MD, MPH	University of New Mexico Health Sciences Center	Albuquerque	NM
Frederick Kaskel, MD, PhD; Shlomo Shinnar, MD, PhD	Albert Einstein College of Medicine/Montefiore Medical Center	Bronx	NY
Jeffrey Saland, MD	Icahn School of Medicine at Mount Sinai	New York	NY
Marc Lande, MD; George Schwartz, MD	University of Rochester Medical Center	Rochester	NY
Anil Mongia, MD	State University of New York, Downstate Medical Center	Brooklyn	NY
Donna Claes, MD; Mark Mitsnefes, MD	Cincinnati Children's Hospital	Cincinnati	OH
Katherine Dell, MD	Case Western Reserve University/Cleveland Clinic Children's	Cleveland	OH
Hiren Patel, MD	Nationwide Children's Hospital	Columbus	OH
Pascale Lane, MD	University of Oklahoma Health Sciences Center	Oklahoma City	OK
Rulan Parekh, MD; Lisa Robinson, MD	Hospital for Sick Children (Sick Kids)	Toronto	Ontario, Canada

Amira Al-Uzri, MD, MCR; Kelsey Richardson, MD	Oregon Health and Science University	Portland	OR
Susan Furth, MD, PhD; Larry Copelovitch, MD	Children's Hospital of Philadelphia	Philadelphia	PA
Elaine Ku, MD, MAS	University of California – San Francisco (UCSF)	San Francisco	SF
Joshua Samuels, MD	University of Texas Health Science Center at Houston	Houston	TX
Poyyapakkam Srivaths, MD	Baylor College of Medicine (Texas Children's Hospital)	Houston	TX
Samhar Al-Akash, MD	Driscoll Children's Hospital	Corpus Christi	TX
Davoud Mohtat, MD	INOVA Children's Hospital / Pediatric Specialists of Virginia	Fairfax	VA
Victoria Norwood, MD	University of Virginia	Charlottesville	VA
Joseph Flynn, MD	Seattle Children's Hospital	Seattle	WA
Cynthia Pan, MD	Medical College of Wisconsin	Milwaukee	WI
Sharon Bartosh, MD	University of Wisconsin	Madison	WI