Towards the best kidney failure prediction tool: a systematic review and selection aid

Chava L. Ramspek[1], Ype de Jong[1,2], Friedo W. Dekker[1] and Merel van Diepen[1]

[1]Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands and [2]Department of Internal Medicine, Leiden University Medical Center, Leiden, The Netherlands

Correspondence and offprint requests to: Chava L. Ramspek; E-mail: c.l.ramspek@lumc.nl

ABSTRACT

Background. Prediction tools that identify chronic kidney disease (CKD) patients at a high risk of developing kidney failure have the potential for great clinical value, but limited uptake. The aim of the current study is to systematically review all available models predicting kidney failure in CKD patients, organize empirical evidence on their validity and ultimately provide guidance in the interpretation and uptake of these tools.

Methods. PubMed and EMBASE were searched for relevant articles. Titles, abstracts and full-text articles were sequentially screened for inclusion by two independent researchers. Data on study design, model development and performance were extracted. The risk of bias and clinical usefulness were assessed and combined in order to provide recommendations on which models to use.

Results. Of 2183 screened studies, a total of 42 studies were included in the current review. Most studies showed high discriminatory capacity and the included predictors had large overlap. Overall, the risk of bias was high. Slightly less than half the studies (48%) presented enough detail for the use of their prediction tool in practice and few models were externally validated.

Conclusions. The current systematic review may be used as a tool to select the most appropriate and robust prognostic model for various settings. Although some models showed great potential, many lacked clinical relevance due to being developed in a prevalent patient population with a wide range of disease severity. Future research efforts should focus on external validation and impact assessment in clinically relevant patient populations.

Keywords: kidney failure, prediction model, prognostic, systematic review

BACKGROUND

Chronic kidney disease (CKD) may lead to kidney failure, although the rates of progression vary substantially between individuals [1]. Prediction tools that can identify patients at high risk of developing kidney failure could have great clinical value. They could be used to inform individualized decision making, employed in determining the appropriate time for referral to nephrologists and used in the planning and preparation of renal replacement therapy (RRT). Prediction tools might also offer opportunities for risk stratification in research and improvement of health policies [2].

Multiple prediction models have been developed to identify individuals at high risk of kidney failure and have been previously described in two systematic reviews [3, 4]. Many of these models showed good predictive abilities in development. However, despite nephrologists and patients acknowledging a lack of prognosis discussions in practice, clinical uptake of these tools is still limited [5]. Policymakers also seem hesitant in endorsing prediction tools. The most recent Kidney Disease: Improving Global Outcomes guideline recommends the use of prediction models for timely referral for planning RRT [6]. However, the guideline, fails to provide guidance on which risk prediction tool should be used to do so. The lack of uptake by clinicians and policymakers has been partly attributed to substandard methodology, a lack of external validation and a shortage of easy calculation options [7].

The last two published reviews in 2012 and 2013 included eight studies each on prediction of kidney failure in CKD patients [3, 4]. Since then the number of available models has greatly increased. A new systematic review of the available models is the first step towards the use and recommendation of robust prognostic tools. The aim of the current study is therefore to systematically review all available models predicting kidney failure in CKD patients, organize empirical evidence on their validity and ultimately provide guidance in the selection of the best prediction tool for various settings.

MATERIALS AND METHODS

Data sources and searches

The current review was framed by the search for prognostic prediction models for CKD patients, predicting the future event

of kidney failure. To ensure transparent reporting and accurate study appraisal, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) and Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies (CHARMS) guidelines were followed where applicable [8–10]. The completed PRISMA checklist is provided as supplementary material. We searched the PubMed and EMBASE databases on 31 December 2017 for English-language studies regarding risk prediction in CKD patients. The search strategies were designed to include relevant development, validation and implementation studies and are provided in Appendix A1.

Study selection

Titles, abstracts and full-text articles were sequentially screened for inclusion by two independent researchers (C.L.R. and Y.J.). Discrepancies on inclusion of full-text articles were resolved by consulting a third co-author (M.D.). Articles were included if they met the following predefined selection criteria: (i) the study must develop, validate, update or implement a multivariate prognostic prediction model, with a prediction research question as the aim, as opposed to an aetiological or methodological goal; (ii) the study must present at least one measure to assess model performance; (iii) the study population must consist of adult CKD patients and (iv) the study outcome must include kidney failure or end-stage renal disease. The references of included studies and related reviews were manually screened in order to identify additional relevant studies.

Data extraction and quality assessment

Following selection, two reviewers (C.L.R. and Y.J.) independently conducted the data extraction and quality assessment. Discrepancies were discussed with input from an additional co-author (M.D.) where necessary. Conforming with CHARMS recommendations, information on the source of the data, population, outcome, sample size, missing data, model development and model performance were extracted and summarized. Additionally, data on external validations of models were extracted. Furthermore, the risk of bias and clinical usefulness were judged by both reviewers independently. In order to facilitate further comparison, studies were grouped by study population, which ranged from very broad (general CKD) to specific CKD subgroups such as immunoglobulin A (IgA) nephropathy or diabetic nephropathy. Quality and risk of bias were assessed in both development and validation studies by making use of a novel tool, the Prediction Study Risk of Bias Assessment Tool (PROBAST). Although this tool has yet to be published in its complete form, there is no other formal risk of bias assessment available that is applicable to prediction studies. The PROBAST is specifically designed for systematic reviews of prediction studies and is used as a domain-based approach with 23 signalling questions that categorize the risk of bias into high, low or unclear for five separate domains: participant selection, predictors, outcome, sample size and missing data, and analysis. It also assesses the usability of a model. It has been used in multiple reviews in the past year and was presented in part at the

2016 Cochrane Colloquia [11]. The final test version of PROBAST was obtained through personal e-mail contact with Dr R.C.G. Wolff.

Data synthesis

Given the multitude of different models and heterogeneity in study characteristics, we opted for a narrative synthesis of results supported by extensive tables and figures with study characteristics listed per article. Model performance was evaluated by examining the discrimination and calibration of included prediction tools. Discrimination is most often described by the C-statistic and indicates how well the model discriminates between patients with and without the event of interest. It lies between 0.5 and 1, where 0.5 is similar to tossing a coin and 1 indicates perfect discrimination [12]. Important to take into account is that the C-statistic of the same model can vary greatly, dependent on the population on which the model is tested. When a population is heterogeneous in the predictors that make up the prediction tool, the C-statistic may increase substantially [13]. On the other hand, calibration describes the agreement between the absolute number of predicted events and observed events population wide. It is best represented in a plot, wherein the predicted probability of kidney failure is plotted against the observed rate of kidney failure [12]. To evaluate the sample size and risk of overfitting in development studies, the events per candidate predictor (EPV) were extracted. A minimum of 10 EPV has been suggested as rule of thumb for an acceptable sample size in model development studies [14]. For external validation studies, it has been recommended to include a minimum of 100 events in total to obtain a precise estimate of performance [15].

RESULTS

Study selection

The study selection process is described in a flowchart (Figure 1). Overall, 2183 titles were identified, of which 431 abstracts were assessed, and 90 full-text publications were evaluated in depth. From these articles, a final 42 studies met all inclusion criteria and were included in the current review. Most full-text exclusions were due to the predicted outcome not including kidney failure or the lack of a multivariate model. Although prediction research has seen a great surge in nephrology over the last few years, the first included predictive model was published in 1986 for IgA nephropathy patients. Since the beginning of the 2000s, a substantial increase of published models is apparent, as can be seen in Figure 2. Although the number of developed models has increased almost every year, the number of validation studies has remained small. Of the 42 included studies, 7 exclusively externally validated already existing models [16–22]. Besides development, 10 studies also externally validated their own or previously published models. Disconcertingly, no study assessing the impact of using such a prediction tool was found, which ultimately is the only way of assessing whether the model can improve patient care.

Characteristics of development studies

A total of 35 studies were published on the development of novel tools to predict kidney failure in CKD patients. Generally,
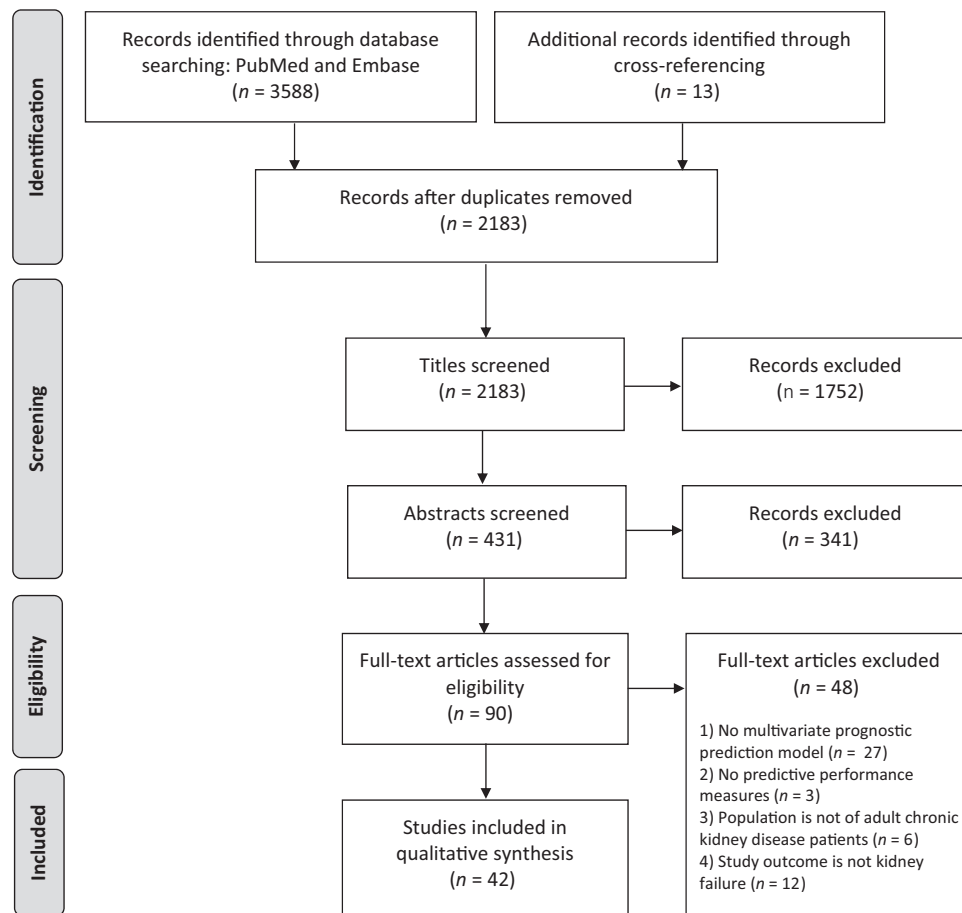
C. L. Ramspek *et al.*

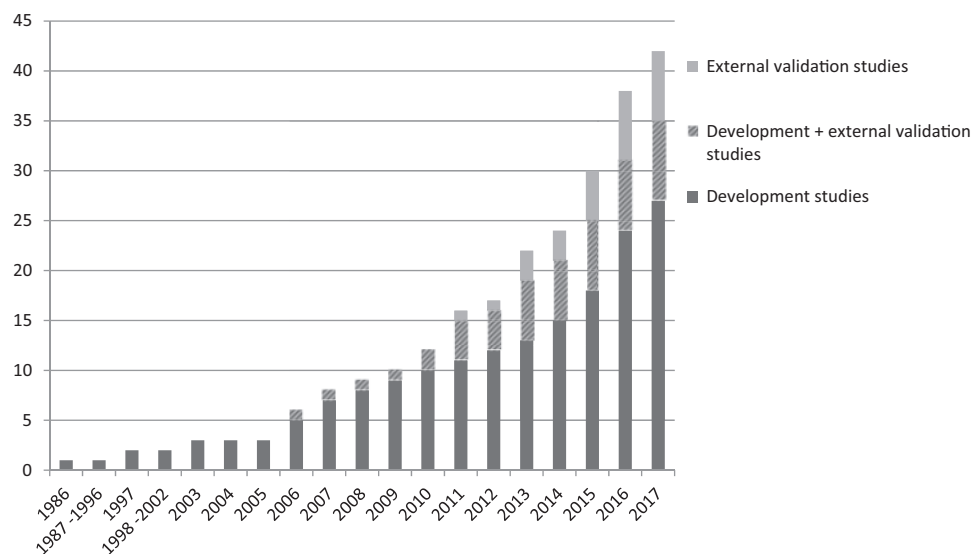FIGURE 1: PRISMA flow diagram of study inclusion.

FIGURE 2: Cumulative number of published development and validation studies for models that predict kidney failure in CKD patients ($N = 42$).

a distinction can be made between models developed for a general CKD patient population ($n = 16$) and models developed for a population with a specified primary renal disease ($n = 19$), mainly IgA nephropathy or diabetic nephropathy. The characteristics of all included development studies are described in Table 1. Since each study developed between 1 and 12 prediction models, the results presented in Table 1 concern the final model(s) as selected by the authors or the model with the best performance if no final model was suggested. The population size differed greatly between studies and ranged from 75 to

Table 1. Baseline characteristics of model development studies ($N = 35$)

Study	Country	Design	Population	Mean GFR	N total, No. of events	Outcome	Time frame (years)	EPV	Model type	N predictors	Internal validation	C-statistic	Calibration	Presented model
General CKD														
Cheng et al. [23]	Taiwan	Single-centre cohort	General CKD Stage 4	–	463, 132	GFR <15	0.5	3	CART	11	Cross-validation	D: 0.72	–	Decision rules
Schroeder et al. [24]	USA	Multicentre cohort	General CKD Stages 3 and 4[a]	47	22 460, 737	RRT	5	74	Cox	8	Bootstrap (+external)	IV: 0.96	D: plot	Formula[b] and score
Hsu et al. [25]	USA	Cohort	General CKD GFR 20–70	44	2466, 581	RRT, 50% GFR↓	–	36	Cox	12	–	D: 0.89	–	HRs
Tangri et al. [26] AJKD	Canada	Single-centre cohort	General CKD Stages 3–5	36	3004, 344	RRT	Dynamic	43	Cox	8	Bootstrap, cross-validation	IV: 0.91	D: plot, test	Formula
Xie et al. [27]	USA	Multicentre cohort	General CKD Stages 3–5[c]	49	28 779, 1730	RRT	1, 3, **5**, 7	115	Cox	5	Cross-validation	IV: 0.92	–	HRs
Marks et al. [28]	Scotland	Multicentre cohort	General CKD Stages 3–5	33	3396, 142	RRT	5	24	Logistic	5	– (external)	D: 0.94	D: test	Formula
Maziarz et al. [29]	USA	Multicentre cohort	General CKD Stages 3–5[c]	–	28 779, 1730	RRT	1, 3, **5**	115	Cox	5	Cross-validation	IV: 0.92	–	HRs
Levin et al. [30]	Canada	Multicentre cohort	General CKD Stages 3–4	28	2402, 142	RRT	1	9	Cox	7	Bootstrap	D: 0.87	D: test	HRs
Maziarz et al. [31]	USA	Multicentre cohort	General CKD Stages 3–5[c]	–	16 656, 959	RRT	1, 3, **5**	63	Cox	5	Cross-validation	IV: 0.90	–	–
Drawz et al. [32]	USA	Single-centre cohort	General CKD Stages 4 and 5[d]	25	1866, 77	RRT	1	4	Cox	6	Bootstrap (+external)	IV: 0.86	–	Formula
Smith et al. [33]	UK	Multicentre cohort	General CKD Stages 3 and 4	32	158, 40	Death, RRT	2	4	Cox	10	–	D: 0.81	–	HRs
Tangri et al. [34]	Canada	Single-centre cohort	General CKD Stages 3–5	36	3449, 386	RRT	1, 3, **5**	16	Cox	4, 8	– (external)	4v: 0.91 8v: 0.92	–	Formula and web calculator
Landray et al. [35]	UK	Single-centre cohort	General CKD Stages 3–5	22	382, 190	RRT	–	4	Cox	4	– (external)	D: 0.87	D: plot	HRs
Johnson et al. [36]	USA	Multicentre cohort	General CKD Stages 3 and 4[a]	–	9782, 323	RRT	5	54	Cox	6	Bootstrap	IV: 0.89	D: plot	Score
Johnson et al. [37]	USA	Multicentre cohort	General CKD Stages 3–5[a]	–	6541, 369	RRT	5	41	Cox	6	–	D: 0.91	–	HRs
Dimitrov et al. [38]	Italy	RCT	General CKD GFR 20–70	43	344, 80	ESRD	–	7	ANN	4	–	D: 0.80	–	Decision tree
Specified renal disease														
Bidadkosh et al. [39]	Multinational	RCT	Diabetic nephropathy	33	861, 60	ESRD, 40% GFR↓	–	6	Cox	8	–	D: 0.79	–	–
Tang et al. [40]	China	Single-centre cohort	Lupus nephritis	78	599, 145	RRT, 50% GFR↓, GFR <15	–	4	Cox	8	Split sample	–	–	HRs and score
Barbour et al. [41]	Multinational	Multicentre cohort	IgA nephropathy	68	901, 162	GFR <15, 50% GFR↓	5	21	Cox	8	Bootstrap	D: 0.80	D: plot	Formula
Li et al. [42]	Taiwan	Single-centre cohort	Diabetic nephropathy	–	131, 22	RRT	–	2	Cox	4	Cross-validation	D: 0.90	–	HRs and score
Pesce et al. [43]	Multinational	Multicentre cohort	IgA nephropathy	87	1040, 241	Time to ESRD	3–8	24	ANN	6	Split sample + cross-validation	IV: 0.90	–	Web calculator (out of service)
Diciolla et al. [44]	Multinational	Multicentre cohort	IgA nephropathy	–	1040, 241	RRT	5	40	ANN	6	Cross-validation	–	–	Web calculator (out of service)
Hoshino et al. [45]	Japan	Single-centre cohort	Diabetic nephropathy	44	205, –	RRT	10	–	Cox	4	Cross-validation	IV: 0.93	–	HRs
Tanaka et al. [46]	Japan	Multicentre cohort	IgA nephropathy	–	698, 73	RRT	5	7	Cox	5	– (external)	D: 0.87	D: plot, test	HRs and score
Xie et al. [47]	China	Single-centre cohort	IgA nephropathy	88	619, 67	ESRD	2, 5, 10	2	Cox	4	–	D: 0.85	–	HRs
Berthoux et al. [48]	France	Single-centre cohort	IgA nephropathy	75	332, 45	Death, RRT	10, 20	8	Score	3	– (external)	–	–	HRs and score
Desai et al. [49]	Multinational	Multicentre RCT	Diabetic nephropathy	35	995, 222	RRT	–	6	Cox	19	Bootstrap	D: 0.85	–	HRs
Day et al. [50]	UK	Single-centre cohort	Pauci-immune GN	–	390, 54	RRT	1	9	Cox	2	–	D: 0.83	–	HRs
Goto et al. [51]	Japan	Multicentre cohort	IgA nephropathy[e]	–	2283, 252	RRT	10	18	Cox	8	Bootstrap + split-sample	D: 0.94 IV: 0.94	–	Score
Kent et al. [52]	Multinational	Multiple RCT's	Non-diabetic CKD	–	1860, 311	RRT/100% SCr↑	–	62	Cox	5	–	D: 0.83	D: plot, test	HRs
Keane et al. [53]	Multinational	RCT	Diabetic nephropathy	–	1513, 341	RRT	–	12	Cox	4	Jackknife	–	D: plot	HRs
Magistroni et al. [54]	Italy	Single-centre cohort	IgA nephropathy	83	237, 40	RRT	10	2	Cox	4	– (external)	–	–	Score
Wakai et al. [55]	Japan	Multicentre cohort	IgA nephropathy[e]	–	2269, 207	ESRD	7	19	Cox	8	Bootstrap + split-sample	D: 0.94 IV: 0.93	–	Score
Frimat et al. [56]	France	Single-centre cohort	IgA nephropathy	–	210, 33	RRT	7	2	Cox	7	–	–	D: plot	Score
Beukhof et al. [57]	The Netherlands	Single-centre cohort	IgA nephropathy	94	75, 14	RRT	10	1	Cox	5	–	–	–	Nomogram

[a]Both studies by Johnson et al. [36, 37] overlap in patient population and include the same predictors. The study by Schroeder et al. [24] updates this same model [36] (the KPNW) by including additional predictors and excluding some original predictors. [b]The formula as provided in the supplement of Schroeder et al.'s article [24] does not provide the knot locations for spline terms. These are available from the authors upon request. [c]The study by Xie [27] and Maziarz [29] includes the same patient population. Part of this population is included in Maziarz et al. [31]. All three studies include the same predictors in the same four models but re-estimate β-coefficients for different subsets. The population is of underserved/uninsured patients. [d]Population of veterans ≥65 years old. [e]Overlap in patient population. The study by Goto et al. [51] has an extended follow-up of 3 years in the same cohort as the study by Wakai et al. [55]. "–" not reported. (e)GFR, (estimated) glomerular filtration rate in mL/min/1.73 m²; EPV, events per variable/candidate predictor; D, development; IV, internal validation; CART, classification and regression tree; ANN, artificial neural networks; RCT, randomized control trial; SCr, serum creatinine.

28 779 patients. A small sample size was a problem in 17/35 studies, as they had <10 EPV, thus running the substantial risk of overfitting their model [14]. To assess to what extent these models are overfit, external validation is of key importance. Before the validity of these models is tested, they should not be used in practice.

For specific renal diseases, the baseline was almost always the first biopsy (and disease confirmation), providing a clear moment in time for when to use the prognostic model or score. Models developed in general CKD, however, rarely defined the moment in time when their prediction tool should be used, as most of these studies enrolled prevalent CKD patients with a wide range of disease severity. Only two models were developed on incident patients, who were included at the first referral to a nephrologist [26, 34]. There was some variation in outcome definitions, but for most studies, renal failure was defined as the need for RRT (dialysis start or kidney transplantation). Five studies used estimated glomerular filtration rate (eGFR) or creatinine as a proxy for kidney failure. Two development studies used RRT start or death as a composite outcome measure. A total of four studies did not report their definition of ESRD. The time frame over which the models predict kidney failure ranged from 6 months to 20 years and nine studies failed to define a prediction time frame, presumably using the maximum study follow-up. The specific predictors included per development study are presented in Figure 3. There is a large amount of overlap in final predictors with almost all studies including age, sex, eGFR (or serum creatinine), proteinuria and histological features for IgA nephropathy tools.

Study	Age	Sex	Ethnicity	Blood pressure	Other	Diabetes	Hypertension	Other	Histology (biopsy)	eGFR	Serum Creatinine	Proteinuria / albuminuria	Serum Albumin	Phosphate	Calcium	Hemoglobin	Other
	patiënt characteristics					comorbidities				laboratory variables							
General CKD																	
Cheng, 2017	X				X_4	X	X	X				X					X_2
Schroeder, 2017	X	X		X	X	X		X		X		X				X	
Hsu, 2016	X	X	X	X	X_4	X		X		X		X					
Tangri, 2016 *AJKD*	X	X								X		X	X	X	X		X
Xie, 2016	X	X	X							X		X					
Marks, 2015	X	X			X							X					
Maziarz, 2015	X	X	X							X		X					
Levin, 2014	X	X								X		X	X			X	X_3
Maziarz, 2014	X	X	X							X		X					
Drawz, 2013	X			X				X		X			X				X
Smith, 2013								X		X		X					X_2
Tangri, 2011 4v model	X	X								X		X					
Tangri, 2011 8v model	X	X								X		X	X	X	X		X
Landray, 2010		X									X	X		X			
Johnson, 2008	X	X				X	X	X		X							
Johnson, 2007	X	X				X	X	X		X							
Dimitrov, 2003											X	X		X	X		
Specified Renal Disease																	
Bidadkosh, 2017	X	X			X						X	X				X	X_2
Tang, 2017								X	X_6	X							
Barbour, 2016			X						X_4	X		X					
Li, 2016										X		X					X_2
Pesce, 2016	X	X					X		X		X	X					
Diciolla, 2015	X	X					X		X		X	X					
Hoshino, 2015	X								X	X		X					
Tanaka, 2013									X_3	X		X					
Xie, 2012				X					X			X				X	
Berthoux, 2011							X		X			X					
Desai, 2011	X	X	X		X_2			X_5	X			X	X			X	X_4
Day, 2010									X		X						
Goto, 2009	X	X		X					X		X	X	X				X
Kent, 2007	X	X		X							X	X					
Keane, 2006											X	X	X			X	
Magistroni, 2006	X			X							X	X					
Wakai, 2006	X	X		X					X		X	X	X				X
Frimat, 1997		X									X	X					
Beukhof 1986	X								X			X					X_2

FIGURE 3: Predictors included in development studies ($N = 35$). The inclusion of a predictor is shown as 'X'. The subscript under X (e.g. 'X_2') indicates the number of predictors included from that category.

Concerning the reporting of performance measures, discrimination measures were reported far more often than calibration measures. Discrimination in the form of a C-statistic was reported in 28/35 studies. The C-statistic ranged from 0.72 to 0.96 and was generally high, indicating good to excellent discrimination in most studies. Calibration was presented far less frequently, with only 11 studies presenting a calibration plot, bar chart or test.

In order to calculate an individual's risk, the model constant and hazard ratios (HRs)/regression coefficients per predictor are needed. Many studies only presented HRs per predictor without the constant (intercept or baseline hazard value), and some gave no data on the model equation at all. The full formula for the developed model was presented in only 6/35 studies. Just three studies provided a web calculator for easy use, of which two web calculators are no longer working. A total of 13 studies provided a simplified scoring system. In total, 25 final models were validated in some form, either internally and/or externally. Cross-validation, bootstrapping and random split sample were the most used forms of internal validation.

Characteristics of external validation studies

A total of 17 studies externally validated one or more of the developed prediction tools. The characteristics of these models and validations can be found in Table 2. Most validation studies were performed by the same group of researchers who developed the models and were often presented in the same publication as the development. Compared with the development performance, the C-statistic was lower in 68% of the validations. Two studies updated the validated model by recalibrating the baseline hazard and two studies added predictors to the existing model. In total, five risk scores predicting prognosis in IgA nephropathy patients and seven prognostic tools for general CKD patients were externally validated. Only the Absolute Renal Risk (ARR) score, Goto score and Kidney Failure Risk Equation (KFRE) (three, four and eight variables) were validated multiple times. The largest validation study of the KFRE was performed by Tangri et al. [18] and summarized the validation of the KFRE in >30 countries, including more than half a million patients.

Risk of bias

Risk of bias was assessed in all 42 included studies, using signalling questions from the PROBAST specified for detecting methodological flaws in both development and validation prediction studies. Overall, the risk of bias was high, as can be seen in Figure 4A and B. Forty-one of 42 studies received a high risk of bias in at least one of the five domains; the only study with an overall low risk of bias was by Schroeder et al. [24]. The majority of studies had a high risk of bias in the domain sample size and missing data. This was often due to the use of complete case analysis, which is generally an inappropriate method of handling missing data. A small sample size was a frequent problem limiting model usage, as a small sample often results in an overfit model and thereby biased results. In the domain statistical analysis, 83% of studies had a high risk of bias. The most

common reason was incomplete reporting of performance measures, as few studies reported sufficient calibration results. Also, many studies did not correct their model for overfitting through internal validation. The usability of the model was assessed in a separate domain. If in the publication the full model formula, a calculator or a risk score with absolute risk table was available, then the tool was considered usable. Less than half the studies (48%) presented enough detail for the use of their prediction tool in practice. The usable models that specified a prediction time frame are presented in Figure 5, categorized by the type of patient population and outcome. This figure may be employed as a selection guide when wanting to calculate an individuals' prognosis, taking into account that many of the models have significant shortcomings and may not be ready for clinical use.

DISCUSSION

This systematic review provides an overview of all development and validation studies of predictive models for progression of CKD to kidney failure. Since the last reviews on this topic, the number of publications has more than doubled [3]. Most included studies report high model performance measures, implying that calculating an individual's risk of renal failure with high accuracy is attainable. This is further emphasized by the similar predictors included in various models. There were, however, substantial shortcomings in many publications. As in many medical prediction studies, aetiological and prediction goals were often confused, limiting interpretability and applicability [7, 58]. First, more than half the tools provided insufficient details to calculate an individual's prognosis of kidney failure, rendering it useless for its intended purpose. Second, the clinical relevance of many models is limited due to the selection of the derivation population. Third, a high risk of bias was observed across studies, mainly due to the high risk of overfitting, inadequate handling of missing data and incomplete reporting of performance measures. Fourth, sufficient validation was largely lacking, increasing research waste and limiting the reliability of models. And finally, not a single impact study on the effect of clinical uptake has been performed. It is therefore not surprising that clinical uptake of models remains sporadic and guidelines on which model to use are lacking.

Providing absolute evidence for the single 'best' prognostic tool to use is complicated by differences between studies, mainly concerning varying study populations, use of different prediction baselines, use of varying time frames and multiple outcome definitions. A selection guide including all usable models is presented that may assist clinicians and patients in choosing the tool appropriate to their setting (Figure 5). There are many factors to take into account when selecting the most appropriate model, depending on the user's wishes and specific clinical setting. Users should be wary of overfitting in models developed on a small sample size and we would advise against the use of these models unless validated in a sufficiently large sample. Based on our results, we would advise the use of a tool with an overall low risk of bias that has shown good performance in external validation in a similar population to the

C. L. Ramspek et al.

Table 2. Characteristics of external validation studies and model performance in validations (N = 17)

Study	Model validated	Independent	Validation type	Country	Population	GFR mean	N total, No. of events	Outcome	Time frame (years)	Model updated	C-statistic	Calibration	Updated model presented
General CKD													
Schroeder et al. [24]	KPNW model (Johnson)	No	External	USA	General CKD Stages 3 and 4	48	16 553, 360	RRT	5	Baseline hazard recalibrated to Colorado	0.95	Plot	No
Lennartz et al. [19]	KFRE 4v (Tangri)	No	External	Germany	General CKD Stages 2–4	46	565, 52	RRT	3	Baseline hazard, addition ultrasound parameters	4v, update: 0.91, 0.91	Plot	Formula
Tangri et al. [18]	KFRE 4v and 8v (Tangri)	No	External	>30 countries	General CKD Stages 3–5	46	721 357, 23 829	RRT	2, **5**	Baseline hazard recalibrated to Europe	4v, 8v: 0.88, 0.88	Plot	Formula and web calculator
Grams et al. [16]	KFRE 4v (Tangri)	Yes	External	USA	CKD, GFR 20–65[a]	–	1094, –	RRT	1, **5**	No	0.83	–	–
Marks et al. [28]	Model 7 (Marks)	No	Temporal	Scotland	General CKD Stages 3–5	47	18 687, 222	RRT	5	No	0.96	Plot, test	–
	KFRE 3v and 4v (Tangri)	Yes	External	Scotland	General CKD Stages 3–5	47	18 687, 222	RRT	5	No	3v, 4v: 0.94, 0.95	Plot	–
Levin et al. [30]	KFRE 8v (Tangri)	No	External	Canada	General CKD Stages 3b–4	28	2402, 142	RRT	1	Coefficients re-estimated, biomarkers added	8v, update: 0.86, 0.87	–	HRs
Drawz et al. [32]	VA risk score (Drawz)	No	Geographic	USA	General CKD Stage 4 and 5[b]	25	819, 33	GFR <15, RRT	1	No	0.82	–	–
	KFRE 8v (Tangri)	Yes	External	USA	General CKD Stage 4 and 5[b]	25	2684, 110	GFR <15, RRT	1	No	0.78	–	–
Peeters et al. [17]	KFRE 3v, 4v, and 8v (Tangri)	Yes	External	The Netherlands	General CKD Stages 3–5	33	595, 114	RRT	5	No	3v, 4v, 8v: 0.88, 0.88, 0.89	Plot, test	–
Tangri et al. [34]	KFRE 3v, 4v, and 8v (Tangri)	No	External	Canada	General CKD Stages 3–5	31	4942, 1177	RRT	1, 3, **5**	No	3v, 4v, 8v: 0.79, 0.83, 0.84	Plot, test	–
Landray et al. [35]	CRIB score (Landray)	No	External	UK	General CKD Stages 3–5	22	213, 66	RRT	–	No	0.91	Plot	–
Specified renal disease													
Knoop et al. [21]	ARR score (Berthoux)	Yes	External	Norway	IgA nephropathy	–	1134, 320	Death, RRT	5, **10**, 15	Coefficients re-estimated, age and GFR added	0.79 update: 0.89	–	Formula
Mohey et al. [22]	ARR score (Berthoux)	No	External	France	Secondary IgA nephropathy	82	74, 19	GFR <15, death	10, 20	No	–	–	–
Tanaka et al. [46]	Tanaka score	No	External	Japan	IgA nephropathy	–	702, 85	RRT	5	No	0.89	Plot, test	–
Xie et al [47]	Goto score	Yes	External	China	IgA nephropathy	88	619, 67	ESRD	2, 5, **10**	No	0.82	–	–
	RENAAL score (Keane)	Yes	External	China	IgA nephropathy	88	619, 67	ESRD	2, 5, **10**	No	0.79	–	–
	ARR score (Berthoux)	Yes	External	China	IgA nephropathy	88	619, 67	ESRD	2, 5, **10**	No	0.73	–	–
Berthoux et al. [48]	ARR score (Berthoux)	No	Temporal	France	IgA nephropathy	–	250, 38	Death, RRT	10, 20	No	–	–	–
Bjorneklett [20]	Goto score	Yes	External	Norway	IgA nephropathy	67	633, 146	RRT	10, 20	Coefficients re-estimated, classification simplified	–	–	No
Magistroni et al. [54]	Magistroni score	No	External	Italy	IgA nephropathy	–	73, 8	RRT	10	No	–	–	–

[a]Hypertensive CKD population. [b]Population of veterans ≥65 years old. v, variable; (e)GFR, (estimated) glomerular filtration rate in mL/min/1.73 m^2; EPV, events per variable/candidate predictor; '–' not reported. The time-frame for which the model performance and other model specifics are reported in bold in the table.
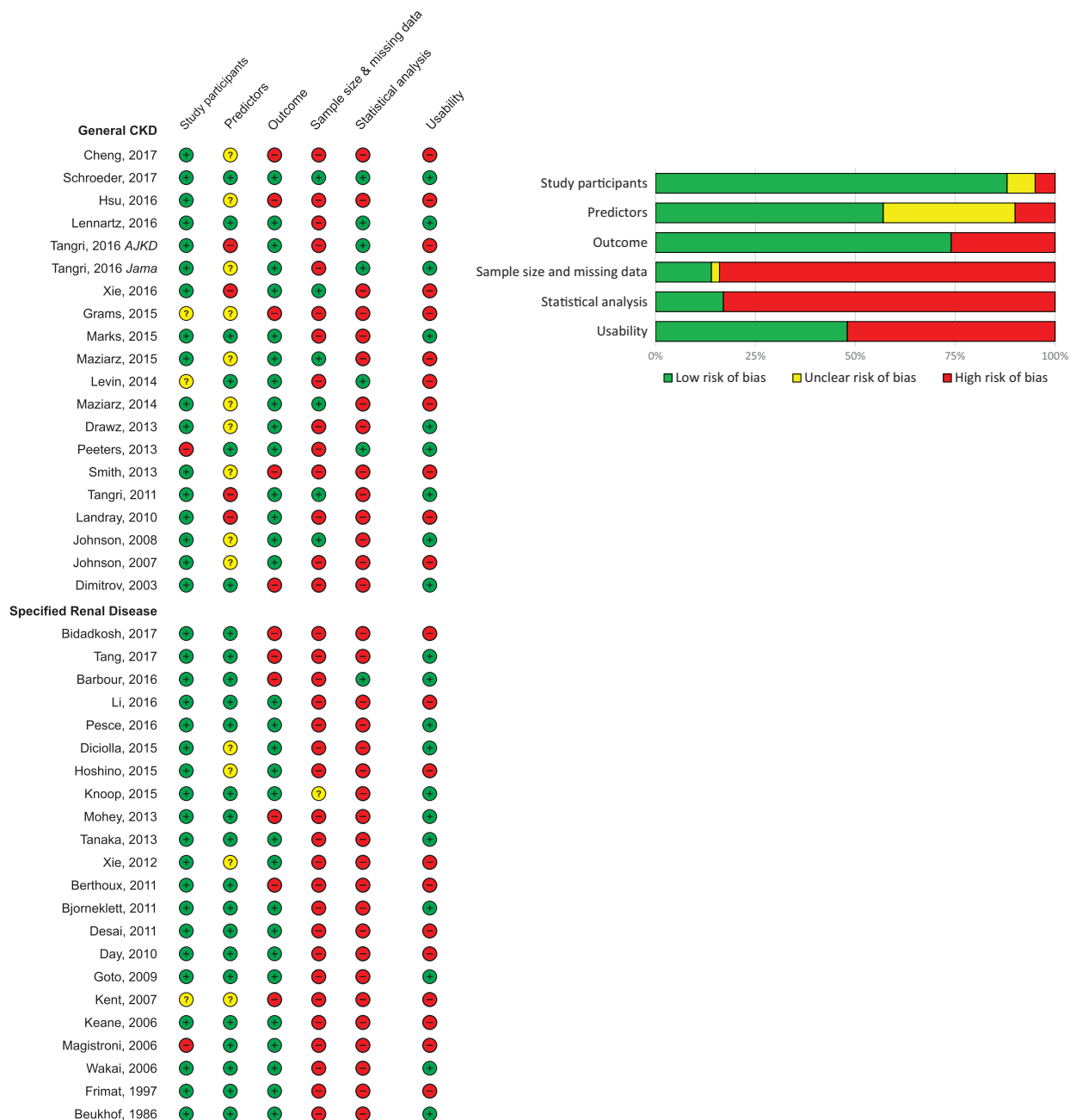
FIGURE 4: (**A**) Risk of bias and usability of prediction models ($N = 42$). Assessed using the PROBAST. The five risk of bias domains were evaluated as low risk (+), unclear risk (?) or high risk (−). Usability was evaluated as yes (+) or no (−). (**B**) PROBAST risk of bias summary for all studies ($N = 42$).

population in which the use is intended and has ideally been assessed in an impact study.

For kidney failure prediction in a general CKD cohort with Stages 3–5 patients, we would recommend the four- or eight-variable KFRE, as it has been externally validated extensively for a time frame of 2 and 5 years. Although the development study potentially introduced bias by selecting predictors that were recorded up to 365 days after prediction baseline and by using univariate analysis to select predictors, the model has shown consistently good performance in CKD Stages 3–5 patients from less-biased external validation studies [18, 34]. Alternatively, for 5-year predictions, the Kaiser Permanente Northwest (KPNW) model as updated and externally validated

by Schroeder et al. [24] also has great potential, mainly due to its methodological rigor and low risk of bias, although it is less easy to use than the KFRE. Various other general CKD models showed promising results in development but should be further externally validated to ensure consistency of performance before clinical use [26, 28, 32]. For prediction of disease progression in IgA nephropathy patients, a large number of models are available. However, these models, were generally developed on a small sample size and often had a high risk of bias. The most evidence on validity was found for the risk scores developed by Goto et al. [51] and the ARR (by Berthoux et al. [48]). The Goto score contains some risk of bias due to a complete case analysis and univariate selection of predictors, but was developed on a
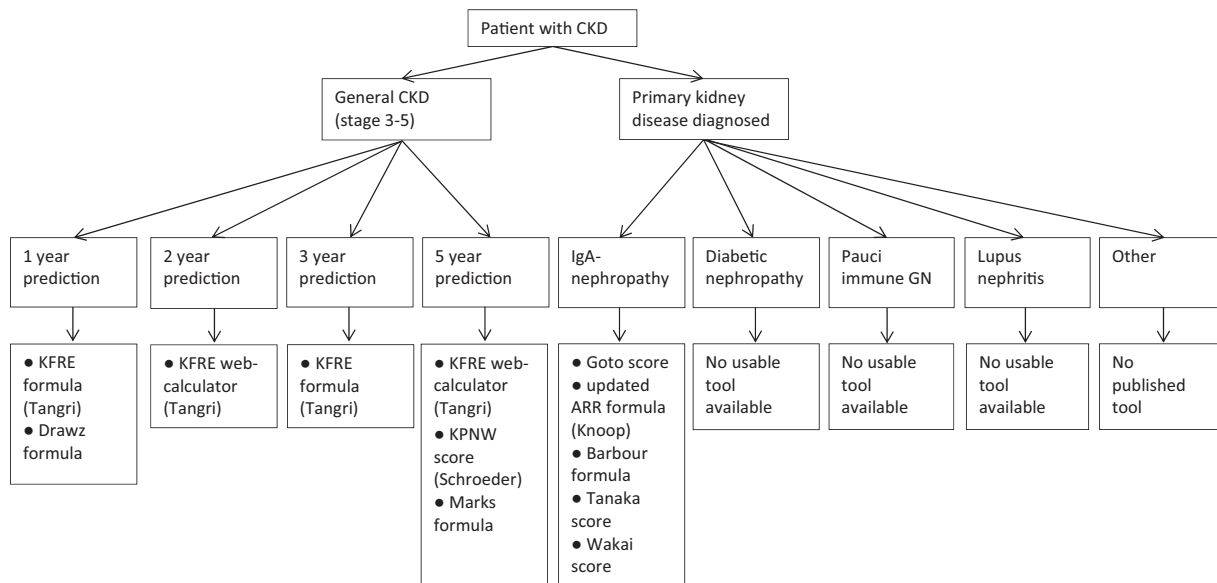
FIGURE 5: Model selection guide for CKD patients. In this graph, only models that allow calculation of an individual's prognosis and are therefore labelled as usable are included. This entails that these models provide either a full formula, score with absolute risk table or (currently working) web calculator for a specified prediction time frame. For categories containing multiple models, the risk of bias combined with evidence of external validity was weighed in determining the model order, starting with the most valid and least biased models. Nevertheless, many of the models listed have significant shortcomings and should be used with caution.

relatively large sample size and has been externally validated twice. Although the ARR score was developed using questionable model building methods and with incomplete reporting of performance, this score has been externally validated the most times, and a recently updated version presented by Knoop *et al.* [21] shows great potential.

Clinical relevance proved to be largely lacking for many of the included models in the current review. Specifically, models for general CKD patients were often developed on prevalent patients with a wide range of disease severity and did not specify a specific time point when the model should be used. Prediction of renal failure can be extremely accurate when using a population with GFRs ranging from 10 to 60 mL/min/1.73 m^2. However, in practice, such tools would probably be employed for a more homogeneous group of patients in which it is clinically relevant to discuss prognosis. The predictive capacities of the model would be lower in such a population. This is exemplified in the KFRE validation performed by Peeters *et al.* [17], where the area under the curve of the four-variable KFRE dramatically decreased from 0.88 in the whole population (CKD Stages 3–5) to 0.71 in the more relevant population of CKD Stage 4 patients. Another factor limiting usability and interpretability is that the number of studies did not define the prediction time frame. Finally, the definition of outcome differs between studies. The use of composite endpoints is particularly problematic, as it limits the value of the model for clinicians, as each separate endpoint requires different interventions. In conclusion, an ideal model is developed for one clearly defined clinically meaningful and objective endpoint in a population for which prediction is clinically relevant. Few models included in this review met these recommendations and this lack of clinical relevance could be a large contributor to the slow uptake seen in practice.

Despite the limited uptake and discussed shortcomings of existing tools, risk prediction models for kidney failure have a great potential for improving patients' decision making, treatment and overall health. In future studies there is the need for improvement of the quality of reporting and methodology used. As the majority of models included had a high risk of bias, these models should not be implemented unless their validity is proven in unbiased external validation studies. Hopefully efforts such as the TRIPOD guidelines will correct these inadequacies and result in more robust, usable and unbiased prognostic tools [9]. To limit research waste and improve clinical uptake, it is of crucial importance that development studies provide enough model information (formula/score with absolute risk table) to enable their use. For specific renal diseases and homogeneous patient populations, there certainly appears to be room for improvement in model development. For populations in which multiple models are available, we advise that future research should focus on the updating, validation and implementation of these existing prognostic tools. Previous studies have shown that the combination of well-established clinical risk factors and kidney disease markers can accurately predict renal failure in a general CKD population. Therefore one might advise focusing resources on updating models for more clinically relevant populations in an unbiased fashion. In this step, external validation of multiple models in the same population is of key importance. Additionally, translation of mathematical model formulas to simple tools such as web calculators and enabling automated uptake is of great importance for integration into daily clinical routine. Ultimately, impact studies will be necessary to determine whether the implementation of such tools truly improves patient outcomes. Ideally, such impact studies would be randomized controlled trials and would assess the effect of implementing a prediction model in

clinical practice. Different outcomes might be considered as endpoints in such studies, partly dependent on the time of prediction. Relevant outcomes might be timely referral to nephrologists, timely placement of vascular access, better informed patients, improved quality of life and possibly even improved survival.

The current review has a number of strengths. First, we expect to have included a complete overview of existing models. Furthermore, this is the first study on kidney failure models to perform a formal risk of bias assessment aimed specifically at prediction research. The study is limited by the inclusion of only English-language articles. Also, the differences in case mix and characteristics of included studies make it difficult to directly compare their performances. Herein we are limited by the lack of validation studies that compare multiple models in the same cohort. Finally, we limited the scope of this review to models predicting kidney failure, although other outcomes such as death or cardiovascular events may also have significant clinical value.

In conclusion, this study provides a systematic overview of existing models for predicting progression to kidney failure in CKD patients. The results may be used as a tool to select the most appropriate and robust prognostic model for various settings. Finally, we hope the current review motivates researchers in this field to decrease the generation of new models and combine efforts to explore, analyse and update existing models in clinically relevant settings to ultimately stimulate clinical uptake and improve patient outcomes.

SUPPLEMENTARY DATA

Supplementary data are available at ndt online.

FUNDING

AUTHORS' CONTRIBUTIONS

All authors have made substantial contributions to the conception of the work and the acquisition and interpretation of data. C.L.R. drafted the work and Y.J., F.W.D. and M.D. critically revised the work. The final version of this manuscript was approved by all the authors and all authors agree to be accountable for all aspects of the work.

CONFLICT OF INTEREST STATEMENT

The authors have no competing interests to declare. All authors declare no support from any organization for the submitted work, no financial relationships with any organizations that might have an interest in the submitted work in the previous 3 years and no other relationships or activities that could appear to have influenced the submitted work.

REFERENCES

1. KDIGO-group. Chapter 2: definition, identification, and prediction of CKD progression. *Kidney Int Suppl* 2013; 3: 63–72
2. Kadatz MJ, Lee ES, Levin A. Predicting progression in CKD: perspectives and precautions. *Am J Kidney Dis* 2016; 67: 779–786
3. Tangri N, Kitsios GD, Inker LA *et al*. Risk prediction models for patients with chronic kidney disease. *Ann Intern Med* 2013; 158: 596–603
4. Echouffo-Tcheugui JB, Kengne AP. Risk models to predict chronic kidney disease and its progression: a systematic review. *PLoS Med* 2012; 9: e1001344
5. Schell JO, Patel UD, Steinhauser KE *et al*. Discussions of the kidney disease trajectory by elderly patients and nephrologists: a qualitative study. *Am J Kidney Dis* 2012; 59: 495–503
6. KDIGO-group. Chapter 5: referral to specialists and models of care. *Kidney Int Suppl* 2013; 3: 112–119
7. Collins GS, Omar O, Shanyinde M et al. A systematic review finds prediction models for chronic kidney disease were poorly reported and often developed using inappropriate methods. *J Clin Epidemiol* 2013; 66: 268–277
8. Liberati A, Altman DG, Tetzlaff J *et al*. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009; 339: b2700
9. Collins GS, Reitsma JB, Altman DG *et al*. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ* 2015; 350: g7594
10. Moons KG, de Groot JA, Bouwmeester W *et al*. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med* 2014; 11: e1001744
11. Wolff RCG, Kleijnen J, Mallett S *et al*. PROBAST: a risk of bias tool for prediction modelling studies. In: *Challenges to Evidence-Based Health Care and Cochrane*. Abstracts of the 24th Cochrane Colloquium. Seoul, Korea: John Wiley & Sons, 2016
12. Royston P, Moons KG, Altman DG *et al*. Prognosis and prognostic research: Developing a prognostic model. *BMJ* 2009; 338: b604
13. Vergouwe Y, Moons KG, Steyerberg EW. External validity of risk models: use of benchmark values to disentangle a case-mix effect from incorrect coefficients. *Am J Epidemiol* 2010; 172: 971–980
14. Wynants L, Bouwmeester W, Moons KG *et al*. A simulation study of sample size demonstrated the importance of the number of events per variable to develop prediction models in clustered data. *J Clin Epidemiol* 2015; 68: 1406–1414
15. Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. *Stats Med* 2016; 35: 214–226
16. Grams ME, Li L, Greene TH *et al*. Estimating time to ESRD using kidney failure risk equations: results from the African American Study of Kidney Disease and Hypertension (AASK). *Am J Kidney Dis* 2015; 65: 394–402
17. Peeters MJ, van Zuilen AD, van den Brand JAJG *et al*. Validation of the kidney failure risk equation in European CKD patients. *Nephrol Dial Transplant* 2013; 28: 1773–1779
18. Tangri N, Grams ME, Levey AS *et al*. Multinational assessment of accuracy of equations for predicting risk of kidney failure: a meta-analysis. *JAMA* 2016; 315: 164–174
19. Lennartz CS, Pickering JW, Seiler MS *et al*. External validation of the kidney failure risk equation and re-calibration with addition of ultrasound parameters. *Clin J Am Soc Nephrol* 2016; 11: 609–615
20. Bjorneklett R, Vikse BE, Bostad L *et al*. Long-term risk of ESRD in IgAN; validation of Japanese prognostic model in a Norwegian cohort. *Nephrol Dial Transplant* 2012; 27: 1485–1491
21. Knoop T, Vagane AM, Vikse BE *et al*. Addition of eGFR and age improves the prognostic absolute renal risk-model in 1,134 Norwegian patients with IgA nephropathy. *Am J Nephrol* 2015; 41: 210–219

22. Mohey H, Laurent B, Mariat C *et al.* Validation of the absolute renal risk of dialysis/death in adults with IgA nephropathy secondary to Henoch–Schonlein purpura: a monocentric cohort study. *BMC Nephrol* 2013; 14: 169

23. Cheng LC, Hu YH, Chiou SH. Applying the temporal abstraction technique to the prediction of chronic kidney disease progression. *J Med Syst* 2017; 41: 85

24. Schroeder EB, Yang X, Thorp ML *et al.* Predicting 5-year risk of RRT in stage 3 or 4 CKD: development and external validation. *Clin J Am Soc Nephrol* 2017; 12: 87–94

25. Hsu CY, Xie D, Waikar SS *et al.* Urine biomarkers of tubular injury do not improve on the clinical model predicting chronic kidney disease progression. *Kidney Int* 2017; 91: 196–203

26. Tangri N, Inker LA, Hiebert B *et al.* A dynamic predictive model for progression of CKD. *Am J Kidney Dis* 2017; 69: 514–520

27. Xie Y, Maziarz M, Tuot DS *et al.* Risk prediction to inform surveillance of chronic kidney disease in the US healthcare safety net: a cohort study. *BMC Nephrol* 2016; 17: 57

28. Marks A, Fluck N, Prescott GJ *et al.* Looking to the future: predicting renal replacement outcomes in a large community cohort with chronic kidney disease. *Nephrol Dial Transplant* 2015; 30: 1507–1517

29. Maziarz M, Black RA, Fong CT *et al.* Evaluating risk of ESRD in the urban poor. *J Am Soc Nephrol* 2015; 26: 1434–1442

30. Levin A, Rigatto C, Barrett B *et al.* Biomarkers of inflammation, fibrosis, cardiac stretch and injury predict death but not renal replacement therapy at 1 year in a Canadian chronic kidney disease cohort. *Nephrol Dial Transplant* 2014; 29: 1037–1047

31. Maziarz M, Chertow GM, Himmelfarb J et al. Homelessness and risk of end-stage renal disease. *J Health Care Poor Underserved* 2014; 25: 1231–1244

32. Drawz PE, Goswami P, Azem R *et al.* A simple tool to predict end-stage renal disease within 1 year in elderly adults with advanced chronic kidney disease. *J Am Geriatr Soc* 2013; 61: 762–768

33. Smith ER, Lee D, Cai MM *et al.* Urinary neutrophil gelatinase-associated lipocalin may aid prediction of renal decline in patients with non-proteinuric stages 3 and 4 chronic kidney disease (CKD). *Nephrol Dial Transplant* 2013; 28: 1569–1579

34. Tangri N, Stevens LA, Griffith J *et al.* A predictive model for progression of chronic kidney disease to kidney failure. *JAMA* 2011; 305: 1553–1559

35. Landray MJ, Emberson JR, Blackwell L *et al.* Prediction of ESRD and death among people with CKD: the Chronic Renal Impairment in Birmingham (CRIB) prospective cohort study. *Am J Kidney Dis* 2010; 56: 1082–1094

36. Johnson ES, Thorp ML, Platt RW et al. Predicting the risk of dialysis and transplant among patients with CKD: a retrospective cohort study. *Am J Kidney Dis* 2008; 52: 653–660

37. Johnson ES, Thorp ML, Yang X *et al.* Predicting renal replacement therapy and mortality in CKD. *Am J Kidney Dis* 2007; 50: 559–565

38. Dimitrov BD, Ruggenenti P, Stefanov R *et al.* Chronic nephropathies: individual risk for progression to end-stage renal failure as predicted by an integrated probabilistic model. *Nephron Clin Pract* 2003; 95: c47–c59

39. Bidadkosh A, Lambooy SPH, Heerspink HJ *et al.* Predictive properties of biomarkers GDF-15, NTproBNP, and hs-TnT for morbidity and mortality in patients with type 2 diabetes with nephropathy. *Diabetes Care* 2017; 40: 784–792

40. Tang Y, Qin W, Peng W et al. Development and validation of a prediction score system in lupus nephritis. *Medicine (Baltimore)* 2017; 96: e8024

41. Barbour SJ, Espino-Hernandez G, Reich HN *et al.* The MEST score provides earlier risk prediction in lgA nephropathy. *Kidney Int* 2016; 89: 167–175

42. Li HY, Lin HA, Nien FJ *et al.* Serum vascular adhesion protein-1 predicts end-stage renal disease in patients with type 2 diabetes. *PLoS One* 2016; 11: e0147981

43. Pesce F, Diciolla M, Binetti G *et al.* Clinical decision support system for end-stage kidney disease risk estimation in IgA nephropathy patients. *Nephrol Dial Transplant* 2016; 31: 80–86

44. Diciolla M, Binetti G, Di Noia T *et al.* Patient classification and outcome prediction in IgA nephropathy. *Comput Biol Med* 2015; 66: 278–286

45. Hoshino J, Mise K, Ueno T *et al.* A pathological scoring system to predict renal outcome in diabetic nephropathy. *Am J Nephrol* 2015; 41: 337–344

46. Tanaka S, Ninomiya T, Katafuchi R *et al.* Development and validation of a prediction rule using the Oxford classification in IgA nephropathy. *Clin J Am Soc Nephrol* 2013; 8: 2082–2090

47. Xie J, Kiryluk K, Wang W *et al.* Predicting progression of IgA nephropathy: new clinical progression risk score. *PLoS One* 2012; 7: e38904

48. Berthoux F, Mohey H, Laurent B *et al.* Predicting the risk for dialysis or death in IgA nephropathy. *J Am Soc Nephrol* 2011; 22: 752–761

49. Desai AS, Toto R, Jarolim P *et al.* Association between cardiac biomarkers and the development of ESRD in patients with type 2 diabetes mellitus, anemia, and CKD. *Am J Kidney Dis* 2011; 58: 717–728

50. Day CJ, Howie AJ, Nightingale P *et al.* Prediction of ESRD in pauci-immune necrotizing glomerulonephritis: quantitative histomorphometric assessment and serum creatinine. *Am J Kidney Dis* 2010; 55: 250–258

51. Goto M, Wakai K, Kawamura T *et al.* A scoring system to predict renal outcome in IgA nephropathy: a nationwide 10-year prospective cohort study. *Nephrol Dial Transplant* 2009; 24: 3068–3074

52. Kent DM, Jafar TH, Hayward RA *et al.* Progression risk, urinary protein excretion, and treatment effects of angiotensin-converting enzyme inhibitors in nondiabetic kidney disease. *J Am Soc Nephrol* 2007; 18: 1959–1965

53. Keane WF, Zhang Z, Lyle PA *et al.* Risk scores for predicting outcomes in patients with type 2 diabetes and nephropathy: the RENAAL study. *Clin J Am Soc Nephrol* 2006; 1: 761–767

54. Magistroni R, Furci L, Leonelli M *et al.* A validated model of disease progression in IgA nephropathy. *J Nephrol* 2006; 19: 32–40

55. Wakai K, Kawamura T, Endoh M *et al.* A scoring system to predict renal outcome in IgA nephropathy: from a nationwide prospective study. *Nephrol Dial Transplant* 2006; 21: 2800–2808

56. Frimat L, Briancon S, Hestin D *et al.* IgA nephropathy: prognostic classification of end-stage renal failure. L'Association des Nephrologues de l'Est. *Nephrol Dial Transplant* 1997; 12: 2569–2575

57. Beukhof JR, Kardaun O, Schaafsma W *et al.* Toward individual prognosis of IgA nephropathy. *Kidney Int* 1986; 29: 549–556

58. van Diepen M, Ramspek CL, Jager KJ *et al.* Prediction versus aetiology: common pitfalls and how to avoid them. *Nephrol Dial Transplant* 2017; 32(Suppl 2): ii1–ii5

APPENDIX A1

Search strategies used on 31 December 2017.

PubMed:

('ESRD'[ti] OR 'ESKD'[ti] OR ((end stage*[ti] OR endstage*[ti]) AND ('renal'[ti] OR kidney*[ti])) OR 'Kidney Failure, Chronic'[majr] OR 'Chronic Kidney Failure'[ti] OR 'Chronic Renal Failure'[ti] OR 'Renal Insufficiency, Chronic'[majr] OR 'chronic Renal Insufficiency'[ti] OR 'chronic kidney Insufficiency'[ti] OR 'CKD'[ti] OR 'chronic kidney disease'[ti] OR 'chronic kidney diseases'[ti] OR nephropath*[ti]) AND ('predictive model'[ti] OR 'predictive models'[ti] OR predictive model*[ti] OR 'prediction model'[ti] OR 'prediction models'[ti] OR prediction model*[ti] OR 'prediction rule'[ti] OR 'prediction rules'[ti] OR 'predictive rule'[ti] OR 'predictive rules'[ti] OR 'prognostic model'[ti] OR 'prognostic models'[ti] OR prognostic model*[ti] OR 'risk score'[ti] OR 'risk scores'[ti] OR 'score'[ti] OR 'scoring'[ti] OR 'predictive'[ti] OR 'predicting'[ti] OR 'predict' [ti] OR 'predicts' [ti] OR 'prediction'[ti] OR 'Risk Assessment'[Majr] OR 'risk assessment'[ti] OR 'risk assessments'[ti]) AND English[lang]

EMBASE:

('ESRD'.ti. OR 'ESKD'.ti. OR ((end stage*.ti. OR endstage*.ti.) AND ('renal'.ti. OR kidney*.ti.)) OR exp *chronic kidney failure/OR 'Chronic Kidney Failure'.ti. OR 'Chronic Renal

Failure'.ti. OR 'chronic Renal Insufficiency'.ti. OR 'chronic kidney Insufficiency'.ti. OR 'CKD'.ti. OR 'chronic kidney disease'.ti. OR 'chronic kidney diseases'.ti. OR nephropath*.ti.) AND ('predictive model'.ti. OR 'predictive models'.ti. OR predictive model*.ti. OR 'prediction model'.ti. OR 'prediction models'.ti. OR prediction model*.ti. OR 'prediction rule'.ti. OR 'prediction rules'.ti. OR 'predictive rule'.ti. OR 'predictive rules'.ti. OR

'prognostic model'.ti. OR 'prognostic models'.ti. OR prognostic model*.ti. OR 'risk score'.ti. OR 'risk scores'.ti. OR 'score'.ti. OR 'scoring'.ti. OR 'predictive'.ti. OR 'predicting'.ti. OR 'predict'.ti. OR 'predicts'.ti. OR 'prediction'.ti. OR exp *'Risk Assessment'/ OR 'risk assessment'.ti. OR 'risk assessments'.ti.) AND English.lg. NOT (conference OR conference abstract OR conference paper OR 'conference review').pt.

Post-partum acute kidney injury: sorting placental and non-placental thrombotic microangiopathies using the trajectory of biomarkers

Fleuria Meibody[1,*], Matthieu Jamme[2,*], Vassilis Tsatsaris[3], François Provot[4,5], Jérôme Lambert[6], Véronique Frémeaux-Bacchi[7], Anne-Sophie Ducloy-Bouthors[8], Mercédès Jourdain[9], Yahsou Delmas[5,10], Pierre Perez[1,5], Julien Darmian[11], Alain Wynckel[12], Jean-Michel Rebibou[13], Paul Coppo[5,14], Cédric Rafat[2], Eric Rondeau[2,5], Luc Frimat[1] and Alexandre Hertig[2]

[1]Department of Nephrology and Kidney Transplantation, University Hospital of Nancy, Vandoeuvre-les-Nancy, France, [2]Sorbonne Université, Urgences Néphrologiques et Transplantation Rénale, Assistance Publique-Hôpital de Paris (APHP), Hôpital Tenon, Paris, France, [3]APHP, Department of Obstetrics and Gynecology, Port-Royal Maternity, University Hospital Center Cochin Broca Hôtel Dieu, Groupe Hospitalier Universitaire Ouest, Paris, France, [4]Department of Nephrology, Transplantation, Dialysis and Apheresis, Claude-Huriez Hospital, CHRU de Lille, Lille, France, [5]French Reference Center for Thrombotic Microangiopathies, APHP, Hôpital Saint-Antoine, Paris, France, [6]Biostatistics Department, Saint Louis Teaching Hospital, APHP, Paris, France, [7]INSERM UMRS-1138, Cordeliers Research Center, Complement and Diseases Team, Paris, France, [8]Department of Anesthesia and Intensive Care, Jeanne-de-Flandre Hospital, CHRU de Lille, Lille, France, [9]Intensive Care Unit, Pôle de Réanimation, University of Lille, CHU Lille, U1190, Lille, France, [10]Department of Nephrology Transplantation-Dialysis, Centre Hospitalier Universitaire de Bordeaux, Bordeaux, France, [11]Department of Intensive Care, Centre Hospitalier Régional Metz-Thionville, Ars-Laquenexy, France, [12]Department of Nephrology, Centre Hospitalier Universitaire, Reims, France, [13]Department of Nephrology, Dijon University Hospital, Dijon, France and [14]Sorbonne Université, Hematology, APHP, Hôpital Saint-Antoine, Paris, France

Correspondence and offprint requests to: Alexandre Hertig; E-mail: alexandre.hertig@sorbonne-universite.fr
*These authors contributed equally to this work.

ABSTRACT

Background. Among the severe complications of preeclampsia (PE), acute kidney injury (AKI) is problematic if features of thrombotic microangiopathy (TMA) are present. Although a haemolysis enzyme liver low-platelets syndrome is considerably more frequent, it is vital to rule out a flare of atypical haemolytic and uraemic syndrome (aHUS). Our objective was to improve differential diagnosis procedures in post-partum AKI.

Methods. A total of 105 cases of post-partum AKI, admitted to nine different regional French intensive care units from 2011 to 2015, were analysed. Analysis included initial and final diagnosis, renal features, haemostasis and TMA parameters, with particular focus on the dynamics of each component within the first days following delivery. A classification and regression tree (CART) was used to construct a diagnostic algorithm.

Results. AKI was attributed to severe PE ($n = 40$), post-partum haemorrhage ($n = 33$, including 13 renal cortical necrosis) and 'primary' TMA ($n = 14$, including 10 aHUS and 4 thrombotic thrombocytopenic purpura). Congruence between initial and final diagnosis was low (63%). The dynamics of haemoglobin, haptoglobin and liver enzymes were poorly discriminant. In contrast, the dynamic pattern of platelets was statistically different between primary TMA-related AKI and other groups. CART analysis independently highlighted the usefulness of platelet trajectory in the diagnostic algorithm. Limitations of this study include that only the most severe cases were included in this retrospective study, and the circumstantial complexity is high.