Journal of Diabetes and Its Complications 33 (2019) 107406

Contents lists available at ScienceDirect

Journal of Diabetes and Its Complications

journal homepage: WWW.JDCJOURNAL.COM

Validation of a protein biomarker test for predicting renal decline in type 2 diabetes: The Fremantle Diabetes Study Phase II

Kirsten E. Peters [a,b], Wendy A. Davis [a], Jun Ito [b], Scott D. Bringans [b], Richard J. Lipscombe [b], Timothy M.E. Davis [a,*]

[a] Medical School, University of Western Australia, Western Australia, Australia
[b] Proteomics International, Perth, Western Australia, Australia

ARTICLE INFO

Article history:
Received 24 April 2019
Received in revised form 14 July 2019
Accepted 14 July 2019
Available online 27 August 2019

Keywords:
Type 2 diabetes
Diabetic nephropathy
Diabetic kidney disease
Biomarkers
Prognostic test

ABSTRACT

Aims: To validate the prognostic utility of a novel plasma biomarker panel, PromarkerD, for predicting renal decline in an independent cohort of people with type 2 diabetes.
Methods: Models for predicting rapid estimated glomerular filtration rate (eGFR) decline defined as i) incident diabetic kidney disease (DKD), ii) eGFR decline ≥30% over four years, and iii) annual eGFR decline ≥5 mL/min/1.73 m^2 were applied to 447 participants from the longitudinal observational Fremantle Diabetes Study Phase II. Model performance was assessed using discrimination and calibration.
Results: During 4.2 ± 0.3 years of follow-up, 5–10% of participants experienced a rapid decline in eGFR. A consensus model comprising apolipoprotein A-IV (apoA4), CD5 antigen-like (CD5L), insulin-like growth factor–binding protein 3 (IGFBP3), age, serum HDL-cholesterol and eGFR showed the best performance for predicting incident DKD (AUC = 0.88 (95% CI 0.84–0.93)); calibration Chi-squared = 5.6, $P = 0.78$). At the optimal score cut-off, this model provided 86% sensitivity, 78% specificity, 30% positive predictive value and 98% negative predictive value for four-year risk of developing DKD.
Conclusions: The combination of readily available clinical and laboratory features and the PromarkerD biomarkers (apoA4, CD5L, IGFBP3) proved an accurate prognostic test for future renal decline in an independent validation cohort of people with type 2 diabetes.

1. Introduction

Diabetes is the most frequent risk factor for chronic kidney disease (CKD). An estimated 1 in 3 adults with diabetes have CKD,[2] and diabetes is the leading cause of end stage renal disease (ESRD) with 44% of ESRD cases in the US attributable to diabetes.[3] Diabetes-associated CKD is the 16th leading cause of death in the US, accounting for 40,000 deaths per year.[4] Despite the high prevalence of CKD complicating diabetes, most patients are unaware they have kidney disease. Timely identification of those at risk is, therefore, an essential part of implementing interventions that can prevent progression of CKD.

Conventional tests for assessing renal function, namely the urinary albumin-to-creatinine ratio (ACR) and estimated glomerular filtration rate (eGFR), have limited accuracy in predicting CKD progression.[5] A large number of promising urinary and plasma biomarkers have been assessed in the context of CKD[6] but large-scale longitudinal studies are required to validate their value over and above that of known clinical risk factors. We recently identified a novel panel of six plasma protein biomarkers (PromarkerD) comprising apolipoprotein A-IV (apoA4), apolipoprotein C-III (apoC3), CD5 antigen-like (CD5L), complement C1q subcomponent subunit B (C1QB), complement factor H–related protein 2 (CFHR2), and insulin-like growth factor–binding protein 3 (IGFBP3) that have diagnostic and prognostic utility in diabetic kidney disease (DKD).[7–9] The PromarkerD biomarkers apoA4, CD5L, C1QB and IGFBP3 predicted future renal decline over four years independently of known clinical risk factors in community-based patients with type 2 diabetes (T2D).[8]

The aim of the present study was to validate the prognostic utility of PromarkerD for predicting rapid renal decline over a four-year follow-up period defined as i) incident DKD, ii) eGFR decline ≥30%, and iii) eGFR decline ≥5 mL/min/1.73m^2/year, in a second, independent cohort of people with type 2 diabetes. The predictive performance of a simple consensus model which could be applied across the three definitions of renal decline was also assessed. It was hypothesized that PromarkerD would accurately predict renal outcomes in an independent group of participants drawn from the same community-based cohort that was used to develop the models, and that a simple consensus model would provide similar predictive performance to more comprehensive models.

* Corresponding author at: University of Western Australia Medical School, Faculty of Health and Medical Sciences, Fremantle Hospital, PO Box 480, Fremantle, Western Australia, 6959, Australia.
E-mail address: tim.davis@uwa.edu.au (T.M.E. Davis).

https://doi.org/10.1016/j.jdiacomp.2019.07.003
1056-8727/

2. Subjects, Materials and Methods

2.1. Study design and participants

Data from the longitudinal observational Fremantle Diabetes Study Phase II (FDS2), details of which have been published elsewhere,[10] were used in the present study. Of 1551 participants with type 2 diabetes recruited to FDS2 between 2008 and 2011, 792 had attended three biennial assessments (baseline, Year 2 and Year 4) between 2008 and 2014. The first 345 to complete follow-up at Year 4 were used for the development of PromarkerD prediction models[8] and are referred to as the "development cohort" in the present study. The remaining 447 participants formed the "validation cohort", an independent (temporally external) sample for assessing the utility of the prediction models. At each visit, all participants underwent a comprehensive assessment including questionnaires, a physical examination and biochemical tests.[10] Fasting plasma samples were stored at −80 °C and used for biomarker measurements in the present FDS2 sub-study. The FDS2 protocol was approved by the Southern Metropolitan Area Health Service Human Research Ethics Committee. All subjects gave informed consent before participation.

2.2. Renal outcomes

In the absence of a single clinically accepted surrogate for ESRD, three commonly used definitions of rapid eGFR decline were assessed i) incident DKD (eGFR <60 mL/min/1.73m^2 at Year 4 in individuals above this threshold at baseline), ii) ≥30% eGFR decline between study entry (baseline) and Year 4 (7.5%/year),[11] and iii) annual decline in eGFR ≥5 mL/min/1.73m^2 calculated as (baseline eGFR - Year 4 eGFR)/(follow-up time between baseline and Year 4).[12] The Chronic Kidney Disease Epidemiology Collaboration (CKD EPI) equation was used to calculate eGFR.[13] Microalbuminuria and macroalbuminuria were defined as a first-morning urinary ACR ≥3 mg/mmol and ≥ 30 mg/mmol, respectively.

2.3. Biomarker quantification

A targeted mass spectrometry platform utilizing multiple reaction monitoring (MRM) was used to measure baseline PromarkerD biomarker concentrations in the validation cohort as used previously for samples from the development cohort.[7,8] Changes in relative peptide abundances were measured against an ^{18}O-labelled reference plasma to give peak area ratios for each biomarker. The PromarkerD biomarkers measured included apolipoprotein A-IV (apoA4), CD5 antigen-like (CD5L) and insulin-like growth factor-binding protein 3 (IGFBP3).[1,8] Complement C1q subcomponent subunit B (C1QB) was not analyzed as it predicted a group-based rapidly declining eGFR trajectory in a previous study,[8] and this outcome definition was not included in the present study.

2.4. Statistical analyses

Data are presented as proportions, mean ± standard deviation (SD), geometric mean (SD range), or, in the case of variables which did not conform to a normal or log$_e$-normal distribution (ln), median and inter-quartile range [IQR]. All biomarker concentrations were ln-transformed prior to analysis. For independent samples, two-way comparisons for proportions were by Fisher's exact test, for normally distributed variables by Student's *t*-test, and for non-normally distributed variables by Mann-Whitney *U* test. All statistical analyses were performed in SPSS for Windows (version 22; SPSS Inc., Chicago, IL) and R (version 3.5.1)[14] using RStudio software (version 1.1.456). A two-tailed level of significance of $P < 0.05$ was used throughout.

A prediction model specific for each of the three definitions of rapid kidney decline and a consensus model were evaluated. The development of PromarkerD prediction models has been described in detail elsewhere.[8] Briefly, the most parsimonious model (referred to as "clinical plus biomarkers model 1" in the earlier study[8]) for each definition of renal decline was defined by considering all clinically plausible variables using a forward stepwise multiple logistic regression approach, before the addition of relative plasma biomarker concentrations. For each definition of renal decline, the final model algorithm (referred to as "parsimonious" in the present study) obtained from the development cohort was applied to individuals in the validation cohort. First, the linear predictor (L) for each participant was calculated using the model intercept (α) and regression coefficients (β-coefficients, $\beta_1 \ldots \beta_m$) for m predictors ($x_1 \ldots x_m$) and the respective participant data at each predictor ($L = \alpha + \beta_1 x_1 + ... + \beta_m x_m$).[8] Next, the predicted probability of renal decline was determined using $e^L / (1 + e^L)$, where e is the exponential function. For each model, only participants with complete data were included.

In addition to the previously defined most parsimonious models, a simple consensus model comprising key baseline clinical predictors (age, serum HDL-cholesterol and eGFR) and relative concentrations of PromarkerD biomarkers at baseline (apoA4, CD5L and IGFBP3) shared across the three renal outcomes was also assessed. The consensus model algorithm obtained from the development cohort was applied to individuals in the validation cohort using the same approach as for the most parsimonious models.

Model performance in the validation cohort was assessed using indices of discrimination and calibration. Model discrimination was assessed by the area under the receiver operating characteristic curve (AUC) which provides the overall accuracy of the test for differentiating individuals at high risk of future renal decline from those at low risk. The sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of the models were determined at the maximum Youden Index, as well as for 5%, 10% and 20% four-year probability of renal decline. The maximum Youden Index (sensitivity+specificity-1) was used to determine the optimal score cut-off at which maximum sensitivity and specificity could be achieved. Model calibration was determined graphically by plotting observed and predicted numbers of participants who experienced a renal outcome across deciles of risk (from the Hosmer-Lemeshow (H-L) goodness-of-fit test). The H-L $\chi 2$ test had eight degrees of freedom for model development and nine at external validation.[15]

The generalizability of the proposed prediction models was evaluated by comparing the relatedness of individuals in the development and validation cohorts.[16] Briefly, a simple logistic regression model was used to predict membership of the development or validation cohort by incorporating as independent variables all predictors defined in the consensus model, i.e. age, serum HDL-cholesterol, eGFR, and plasma concentrations of apoA4, CD5L and IGFBP3, as well as outcome status (whether there was a decline in eGFR or not). If the model discriminated well (AUC > 0.80), then individuals in the two cohorts were said to be "not related" and the transportability rather than reproducibility of the prediction model could be determined.

3. Results

3.1. Participant characteristics and outcomes

At baseline, the 792 participants included in the present study had a mean ± SD age of 65.6 ± 10.3 (range 34.9–93.8) years, 54.3% were male, and their median [IQR] diabetes duration was 7.0 [2.0–15.0] years. The included participants did not differ significantly in age, gender or body mass index (BMI) from the 759 remaining type 2 FDS2 participants (all $P \geq 0.06$), but had shorter duration of diabetes and better renal function (all $P < 0.001$). There were no significant differences in age, sex, BMI, diabetes duration, HbA_{1c} or renal function at baseline between individuals with complete data included in models versus those with missing data (data not shown).

The baseline characteristics of participants in the development and validation cohorts are compared in Table 1. The two cohorts had similar sex distributions, BMIs and renal function at baseline, but those in the validation cohort were significantly younger, had shorter diabetes

duration and a higher fasting plasma glucose ($P \leq 0.002$). During a mean ± SD follow-up of 4.2 ± 0.3 years, 39 (9.8%) individuals in the validation cohort developed DKD, 24 (5.4%) experienced a ≥ 30% decline in eGFR, and 32 (7.2%) had an annual eGFR decline ≥5 mL/min/1.73 m^2 (see Supplementary Fig. 1 for overlap between groups and Supplementary Fig. 2 for mean changes in eGFR over time by group). The number of renal outcomes and durations of follow-up in the two cohorts were not significantly different (see Table 1).

Analysis of the relatedness of the two cohorts and downstream generalizability of the prediction models showed that individuals were not related in terms of predictor variables (based on the consensus model) or outcome status across all definitions of renal decline analyzed (incident DKD: AUC = 0.84 (95% CI 0.81–0.87), eGFR decline ≥30%: AUC = 0.85 (95% CI 0.82–0.87), eGFR decline ≥5 mL/min/1.73 m^2/yr: AUC = 0.85 (95% CI 0.82–0.88)), suggesting the prediction models are transportable to other populations.

3.2. Most parsimonious prediction models

When applied to the validation cohort, each parsimonious model provided moderate to good discrimination (AUC range 0.63 to 0.83; see Supplementary Fig. 3), but calibration was poor ($P \leq 0.01$) (see Table 2 and Fig. 1). The model for incident DKD had the highest predictive ability to discriminate participants who did and did not develop DKD during follow-up (AUC = 0.83 (95%CI 0.77–0.88)). At the optimal score cut-off (3.3%), this model provided 88.9% sensitivity, 63.5%

Table 1
Baseline characteristics and renal outcomes of the development and validation cohorts from the Fremantle Diabetes Study Phase II.

Variable	Development Cohort ($n = 345$)		Validation Cohort ($n = 447$)		*P*-value
	N	Mean ± SD[a]	N	Mean ± SD[a]	
Age (years)	345	67.0 ± 9.4	447	64.4 ± 10.9	<0.001
Gender, % male	345	51.9	447	56.2	0.25
BMI (kg/m^2)	345	31.0 ± 5.5	446	31.7 ± 6.2	0.11
Waist circumference (cm)	345	102.7 ± 13.5	446	105.0 ± 14.9	0.024
Ethnic background (% AC/SE/OE/Asian/Ab/other)	345	64.9/11.0/7.0/3.2/0.3/13.6	447	55.7/11.2/9.2/4.5/2.9/16.6	0.015
Age at diabetes diagnosis (years)	345	57.1 ± 10.9	447	56.0 ± 11.3	0.19
Diabetes duration (years)[c]	345	9.0 [3.0–15.2]	447	6.0 [1.3–14.0]	<0.001
Fasting plasma glucose (mmol/L)[b]	344	7.1 (5.5–9.2)	447	7.6 (5.7–10.1)	0.002
HbA_{1c} (%)[b]	345	6.9 (6.0–8.0)	447	7.0 (5.9–8.4)	0.085
Serum total cholesterol (mmol/L)	344	4.3 ± 1.0	445	4.3 ± 1.0	0.63
Serum HDL-cholesterol (mmol/L)	344	1.28 ± 0.31	445	1.22 ± 0.32	0.011
Serum triglycerides (mmol/L)[b]	344	1.5 (0.9–2.3)	445	1.5 (1.0–2.3)	0.41
Serum uric acid (mmol/L)[b]	344	0.34 (0.26–0.44)	447	0.33 (0.25–0.43)	0.053
Serum creatinine (μmol/L)[b]	345	75 (56–101)	447	75 (60–95)	0.85
Urinary ACR (mg/mmol)[b]	345	2.9 (0.9–8.8)	447	2.5 (0.9–7.1)	0.050
eGFR (mL/min/1.73 m^2)	345	80.6 ± 18.8	447	82.7 ± 16.9	0.091
eGFR categories (% G1/G2/G3a/G3b/G4)[d]	345	37.7/49.3/6.4/5.5/1.2	447	39.6/49.9/7.6/2.9/0.0	0.062
CKD Stage (% 0/1/2/3)[e]	345	57.1/30.7/6.1/6.1	447	59.1/32.4/6.9/1.6	0.008
Systolic blood pressure (mmHg)	345	147 ± 20	446	143 ± 20	0.003
Diastolic blood pressure (mmHg)	345	80 ± 12	446	80 ± 12	0.69
Neuropathy (%)	345	73.6	445	46.5	<0.001
PAD (%)	345	17.4	446	20.2	0.36
CVD (%)	345	5.5	447	5.4	1.00
IHD (%)	345	25.5	447	19.9	0.071
Alcohol consumption (standard drinks/day)[c]	326	0.1 [0.0–1.5]	436	0.1 [0.0–1.2]	0.94
Smoking status (% never/ex−/current)	345	47.2/47.0/5.8	447	45.0/46.5/8.5	0.34
Any physical activity (%)	341	94.4	431	92.6	0.38
Diabetes treatment (%):	345		447		
Diet		29.3		25.7	0.30
OHA		49.0		56.6	0.037
Insulin ± OHA		21.7		17.7	0.17
Anti-hypertensive medications (%):	345	79.7	447	70.5	0.003
Diuretic		34.8		26.6	0.015
ACE-I		44.3		34.5	0.005
ARB		33.9		29.8	0.22
β-Blocker		22.3		15.9	0.027
Calcium channel blocker		26.1		24.2	0.56
Other		4.3		3.6	0.59
Lipid-lowering medication (%):	345	73.9	447	67.3	0.050
Aspirin use (%)	345	43.8	447	34.5	0.008
Renal outcomes (N, %):					
Incident DKD	300	37 (12.3)	400	39 (9.8)	0.33
eGFR decline ≥30%	345	30 (8.7)	447	24 (5.4)	0.087
eGFR decline ≥5 mL/min/1.73 m^2/yr	345	28 (8.1)	447	32 (7.2)	0.69
Plasma biomarker concentrations[b]					
apoA4 (peak area ratio)	345	1.17 (0.57–2.42)	447	0.89 (0.44–1.79)	<0.001
CD5L (peak area ratio)	344	2.37 (1.17–4.79)	445	1.00 (0.56–1.79)	<0.001
IBP3 (peak area ratio)	335	0.97 (0.58–1.64)	416	1.04 (0.66–1.65)	0.049

BMI, body mass index; AC, Anglo-Celt; SE, southern European; OE, other European; Ab, Aboriginal; ACE—I, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; OHA, oral hypoglycemic agent; eGFR, estimated glomerular filtration rate by CKD Epidemiology Collaboration equation; PAD, peripheral arterial disease; CVD, cerebrovascular disease; IHD, ischemic heart disease.

[a] All values are mean ± SD (standard deviation) unless labelled otherwise.

[b] Geometric Mean (SD range).

[c] Median [IQR – interquartile range].

[d] eGFR categories G1 ≥ 90, G2 60–89, G3a 45–59, G3b 30–44, G4 15–29 mL/min/1.73 m^2.

[e] CKD (chronic kidney disease) stage defined by KDIGO 2012 guidelines.[21]

Table 2
Prognostic performance of the most parsimonious prediction models applied to the validation cohort.

	Incident DKD	Decline ≥30%	Decline ≥5 mL/min/1.73m^2/yr
Number of Subjects	362	413	413
Observed outcomes (%)	36 (9.9%)	24 (5.8%)	31 (7.5%)
Predicted outcomes (%)	39 (10.8%)	18 (4.4%)	35 (8.5%)
Discrimination			
AUC (95%CI)	0.83 (0.77–0.88)	0.74 (0.65–0.84)	0.63 (0.53–0.73)
At max YI cut-off:	(3.3%)	(7.9%)	(8.6%)
Sensitivity (%)	88.9	54.2	51.6
Specificity (%)	63.5	88.2	72.3
PPV (%)	21.2	22.1	13.1
NPV (%)	98.1	96.9	94.8
At 5% cut-off:			
Sensitivity (%)	80.6	54.2	71.0
Specificity (%)	68.4	81.5	50.3
PPV (%)	22.0	15.3	10.4
NPV (%)	97.0	96.6	95.5
At 10% cut-off:			
Sensitivity (%)	72.2	41.7	41.9
Specificity (%)	77.9	89.7	76.7
PPV (%)	26.5	20.0	12.7
NPV (%)	96.2	96.1	94.2
At 20% cut-off:			
Sensitivity (%)	52.8	20.8	19.4
Specificity (%)	88.0	95.9	92.7
PPV (%)	32.8	23.8	17.6
NPV (%)	94.4	95.2	93.4
Calibration			
H-L test (χ^2, P)[a]	46.9, <0.001	48.4, <0.001	21.2, 0.01

Only participants with complete data were included in each model. The most parsimonious prediction models are as follows: incident DKD – IHD, eGFR, total cholesterol, apoA4; decline ≥30% - age, diuretic use, IHD, DBP, total cholesterol, apoA4, IBP3; decline ≥5 mL/min/1.73 m^2/year – IHD, DBP, HbA_{1c}, IBP3.[8] Discrimination performance measures are given for 5%, 10%, and 20% 4-year renal decline risk cut-offs, as well as for the optimal cut-off (shown in parentheses) defined by maximum Youden Index (YI). AUC = area under the curve; H-L = Hosmer-Lemeshow; χ^2 = Chi squared test value, P = p-value, PPV = positive predictive value; NPV = negative predictive value.

[a] Hosmer-Lemeshow (H-L) goodness-of-fit test with 9 degrees of freedom.

specificity, 21.2% positive predictive value and 98.1% negative predictive value to predict four-year risk of developing DKD. For the other definitions of renal decline, eGFR decline ≥30% and annual eGFR decline ≥5 mL/min/1.73 m^2, discrimination was moderate (AUC = 0.74 (95% CI 0.65–0.84) and (AUC = 0.63 (95% CI 0.53–0.73), respectively). Using additional predicted risk cut-offs of 5%, 10% and 20%, allowed improvements in specificity, with loss of sensitivity as the cut-off increased (see Table 2). The predicted risk of decline in renal function was underestimated across the lower risk deciles and overestimated for those in the highest risk category for all three definitions (see Fig. 1).

3.3. Simple consensus prediction model

The simple consensus model performed well in both cohorts with moderate to good discrimination (AUC range 0.61 to 0.89; see Supplementary Fig. 3) and acceptable calibration ($P \geq 0.06$) across all three definitions of renal decline (see Table 3 and Fig. 1). The consensus model had the highest predictive ability for incident DKD, providing good discrimination in the development and validation cohorts (AUC = 0.89 (0.85–0.94) and (AUC = 0.88 (0.84–0.93), respectively). At the optimal score cut-off, sensitivities of 94.6% and 86.1%, and specificities of 68.0% and 78.2%, were observed for the development and validation cohorts, respectively. For eGFR decline ≥30% and annual eGFR decline ≥5 mL/min/1.73 m^2, the consensus model provided similar levels of discrimination in the validation cohort to that seen with the most parsimonious models (AUC = 0.73 (0.64–0.81) and (AUC = 0.61 (0.51–0.70), respectively). In the validation cohort, the consensus model overestimated risk by 1.2 fold across the three definitions of renal decline, but the Hosmer-Lemeshow test showed that there was no significant difference between predicted and observed outcomes ($P \geq 0.21$) (see Fig. 1 and Table 3).

4. Discussion

The present study provides the first validation of the PromarkerD test which confirms the prognostic utility of a novel panel of plasma proteins in DKD. Three biomarkers, apoA4, CD5L and IGFBP3, combined with a limited number of conventional clinical variables, accurately predicted rapid eGFR decline over a four-year period in a cohort of community-based patients with type 2 diabetes that was independent of the participants used to develop the test. The predictive performance of a series of models, including i) models specific to each definition of eGFR decline, and ii) a simple consensus model which could be applied across all definitions, had good discrimination and calibration.

The development of the PromarkerD test in an earlier study[8] showed that the plasma protein panel added significant prognostic benefit to known clinical risk factors, including eGFR and albuminuria, for predicting renal decline in participants with type 2 diabetes from FDS2. In the present study, the three most parsimonious definition-specific models were applied to an independent cohort of people with type 2 diabetes also sourced from the FDS2. The prognostic performance of these models was good in terms of discrimination (AUC range 0.65 to 0.83), but calibration was poor reflecting the small numbers of participants in lower deciles. It is not uncommon to observe poor calibration and/or discrimination in validation studies, often due to differences in case-mix, measurement and definition of predictor variables and event rates.[17] Indeed, the baseline characteristics of individuals in the validation cohort differed across most clinical variables included in the parsimonious models (age, diuretic use, history of IHD, eGFR and HbA_{1c}) and fewer had suffered a decline in renal function compared to those in the development cohort.

Given the range of clinical variables required for the three most parsimonious models, some of which may be difficult to obtain in usual care, a simple and easy to use consensus model was derived. This model comprised three clinical variables (age, serum HDL-cholesterol and eGFR) which were chosen based on accessibility in routine diabetes care and statistical significance across the different definitions of renal decline, and were combined with the three plasma proteins apoA4, CD5L and IGFBP3. The consensus model provided similar discriminative ability to the most parsimonious models, but improved on performance in terms of calibration across the three definitions of renal decline.

The development and validation of accurate prediction models that predict renal decline and future DKD risk could aid clinical decision making and support cost-effective individualized treatment. In addition to supporting the optimization of glycemic and blood pressure control in high risk patients, such tests could justify the early introduction of new treatments that have renal benefits beyond blood glucose lowering such as the sodium glucose transporter subtype 2 (SGLT-2) inhibitors empagliflozin and canagliflozin[18,19] By the same token, identifying low risk patients whose management could be rationalized should have benefits for limiting adverse effects and cost while improving adherence. Validated prognostic tests such as PromarkerD may also be useful for selective enrolment of high-risk patients into future clinical trials of interventions in DKD, allowing smaller and shorter trials because of relatively high background event rates. A formal cost-benefit analysis is yet to be conducted but tests such as PromarkerD have the potential to reduce CKD-related healthcare spending that currently exceeds US$100 billion annually.[20]

The present study had limitations. The sample size was not large but the participants were well characterized and they were followed over a

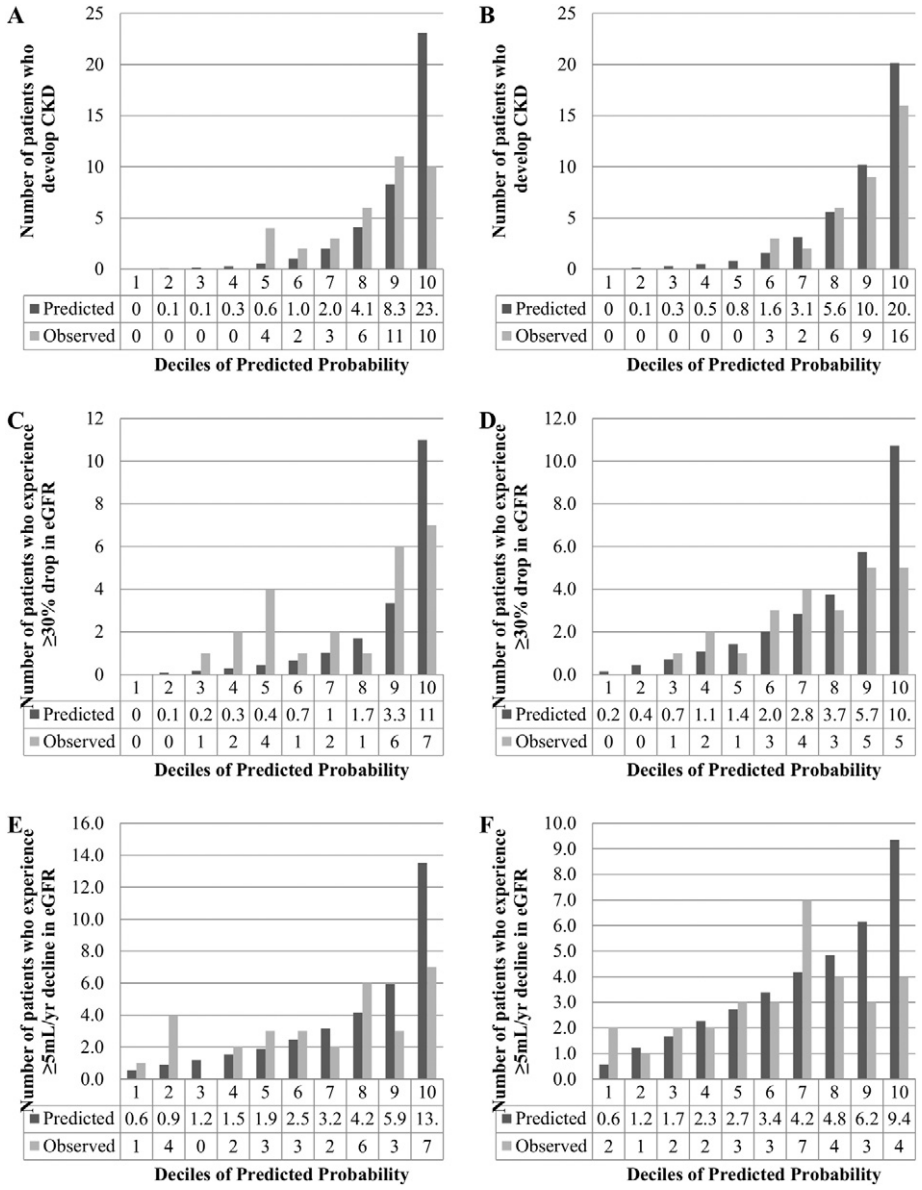

Fig. 1. Calibration plots of the validation cohort showing the observed and predicted number of subjects experiencing a renal outcome over 4 years by deciles of risk. A. Incident DKD (most parsimonious model), B. Incident DKD (consensus model), C. Decline ≥30% (most parsimonious model) D. Decline ≥30% (consensus model), E. Decline ≥5 mL/min/1.73 m^2 (most parsimonious model), F. Decline ≥5 mL/min/1.73m^2 (consensus model). Observed numbers are shown by the light grey bars and predicted numbers by the dark grey bars.

prognostic time horizon that is relevant to people with type 2 diabetes and clinicians. The majority of FDS2 participants were of Caucasian origin (79%), limiting the generalizability of the models to other racial and ethnic groups. Only baseline clinical and biomarker data were used to predict risk, and subsequent changes in biomarker concentrations or diabetes management were not considered. Nevertheless, the prediction models were validated temporally in this study as the relatedness of the two cohorts showed they were clinically distinct, suggesting the models could be applied to similar people with type 2 diabetes. To fully realize the generalizability of the models, additional external validation across different clinical settings and populations with a larger numbers of events, is warranted.

5. Conclusions

The present study assessed and validated the prognostic utility of PromarkerD, a novel diagnostic test that combines a panel of plasma biomarkers (apoA4, CD5L and IGFBP3) with clinical variables (age, HDL-cholesterol and eGFR) to accurately predict future renal decline in people with type 2 diabetes. PromarkerD may be useful for risk stratification in future clinical trials and has the potential to aid clinical decision-making by identifying at-risk individuals for earlier targeted personalised intervention and monitoring of disease progression, with the potential for improved patient outcomes.

Acknowledgements

We are grateful to FDS2 participants for their involvement and FDS2 staff for help with collecting and recording clinical information. We thank the staff at Proteomics International who performed all the biomarker assays. The MS analyses were performed in facilities provided by the Lotterywest State Biomedical Facility-Proteomics node and Bioplatforms Australia at the Harry Perkins Institute for Medical

Table 3
Prognostic performance of the consensus prediction model in the development and validation cohorts.

Performance Measure	Incident DKD		Decline ≥30%		Decline ≥5 mL/min/1.73m^2/yr	
	Development	Validation	Development	Validation	Development	Validation
Number of Subjects	290	362	333	413	333	413
Observed outcomes (%)	37 (12.8%)	36 (9.9%)	30 (9.0%)	24 (5.8%)	28 (8.4%)	31 (7.5%)
Predicted outcomes (%)	36 (12.4%)	42 (11.6%)	30 (9.0%)	28 (6.8%)	28 (8.4%)	36 (8.7%)
Discrimination						
AUC (95%CI)	0.89 (0.85–0.94)	0.88 (0.84–0.93)	0.81 (0.75–0.87)	0.73 (0.64–0.81)	0.70 (0.61–0.80)	0.61 (0.51–0.70)
Optimism-corrected AUC[b]	0.87		0.77		0.64	
At max YI cut-off:	(7.1%)	(12.0%)	(6.1%)	(4.4%)	(10.1%)	(8.3%)
Sensitivity (%)	94.6	86.1	96.7	83.3	60.7	67.7
Specificity (%)	68.0	78.2	61.7	54.0	72.8	57.6
PPV (%)	30.2	30.4	20.0	10.0	17.0	11.5
NPV (%)	98.9	98.1	99.5	98.1	95.3	95.6
At 5% cut-off:						
Sensitivity (%)	97.3	97.2	100.0	70.8	85.7	83.9
Specificity (%)	61.3	63.8	53.1	58.1	38.0	34.0
PPV (%)	26.9	22.9	17.4	9.4	11.3	9.4
NPV (%)	99.4	99.5	100.0	97.0	96.7	96.3
At 10% cut-off:						
Sensitivity (%)	78.4	86.1	60.0	41.7	60.7	45.2
Specificity (%)	74.7	73.9	74.3	79.9	72.5	66.2
PPV (%)	31.2	26.7	18.8	11.4	16.8	9.8
NPV (%)	95.9	98.0	94.9	95.7	95.3	93.7
At 20% cut-off:						
Sensitivity (%)	75.7	72.2	33.3	20.8	17.9	12.9
Specificity (%)	85.4	84.7	90.4	93.8	94.8	94.8
PPV (%)	43.1	34.2	25.6	17.2	23.8	16.7
NPV (%)	96.0	96.5	93.2	95.1	92.6	93.1
Calibration						
H-L test (χ^2, P)[a]	8.5, 0.39	5.6, 0.78	14.7, 0.06	7.1, 0.62	14.3, 0.07	12.0, 0.21

Only participants with complete data were included in each model. The consensus prediction model includes age, serum HDL-cholesterol, eGFR, apoA4, CD5L, IBP3. Discrimination performance measures are given for 5%, 10%, and 20% 4-year renal decline risk cut-offs, as well as for the optimal cut-off (shown in parentheses) defined by maximum Youden Index (YI). AUC = area under the curve; χ^2 = Chi squared test value, P = p-value, PPV = positive predictive value; NPV = negative predictive value.

[a] Hosmer-Lemeshow (H-L) goodness-of-fit test with 8 (development cohort) or 9 (validation cohort) degrees of freedom.

[b] Based on internal validation by bootstrap resampling (development cohort).

Research. Partial funding for the proteomics analysis was provided by the Federal Government of Australia under a Commercialization Australia Grant (CAU05922). FDS2 has been supported by the National Health and Medical Research Council of Australia project grants 513781 and 1042231. TMED is supported by a National Health and Medical Research Council of Australia Practitioner Fellowship (1058260).

Author contributions

KEP collected data, performed the statistical analyses and wrote the manuscript, WAD performed the statistical analyses and is a CI of the FDS, JI performed MRM experiments and sample processing, SB and RL were involved in experimental design, data analysis and interpretation. TD is PI of the FDS, provided the FDS cohort plasma samples, wrote the manuscript and was responsible for clinical interpretation. All authors reviewed/edited the manuscript. KEP is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Declaration of competing interest

Proteomics International and the University of Western Australia are stakeholders in patent PCT/AU2011/001212 that relates to biomarkers described in this manuscript.[1]

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jdiacomp.2019.07.003.

References

1. Stoll T, Bringans S, Winfield K, et al. *Method of Assessing Diabetic Nephropathy Using CD5 Antigen-like. US Patent and Trademark Office (USPTO)*. 2015.
2. Centers for Disease Control and Prevention. National Diabetes Statistics Report: Estimates of Diabetes and Its Burden in the United States; 2014. , https://www.cdc.gov/diabetes/pubs/pdf/kidney_factsheet.pdf. Accessed November 2017.
3. Centers for Disease Control and Prevention. National Chronic Kidney Disease Fact Sheet; 2017. , https://www.cdc.gov/diabetes/pubs/pdf/kidney_factsheet.pdf. Accessed July 2018.
4. Mokdad AH, Ballestros K, Echko M, et al. The state of US health, 1990-2016: burden of diseases, injuries, and risk factors among US states. JAMA 2018;319:1444-72.
5. Dunkler D, Gao P, Lee SF, et al. Risk prediction for early CKD in type 2 diabetes. Clin J Am Soc Nephrol 2015;10:1371-9.
6. Lin CH, Chang YC, Chuang LM. Early detection of diabetic kidney disease: present limitations and future perspectives. World J Diabetes 2016;7:290-301.
7. Bringans SD, Ito J, Stoll T, et al. Comprehensive mass spectrometry based biomarker discovery and validation platform as applied to diabetic kidney disease. EuPA Open Proteom 2017;14.
8. Peters KE, Davis WA, Ito J, et al. Identification of novel circulating biomarkers predicting rapid decline in renal function in type 2 diabetes: the Fremantle diabetes study phase II. Diabetes Care 2017;40:1548-55.
9. Stoll T, Bringans S, Winfield K, et al. Method of assessing a subject for abnormal kidney function. US Patent and Trademark Office (USPTO); 2017.
10. Davis TM, Bruce DG, Davis WA. Cohort profile: the Fremantle diabetes study. Int J Epidemiol 2013;42:412-21.
11. Levey AS, Inker LA, Matsushita K, et al. GFR decline as an end point for clinical trials in CKD: a scientific workshop sponsored by the National Kidney Foundation and the US Food and Drug Administration. Am J Kidney Dis 2014;64:821-35.
12. Kidney Disease. Improving global outcomes chronic kidney disease guideline development work group M. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. Kidney Int Suppl 2013;3:1-150.
13. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med 2009;150:604-12.
14. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria; 2012.
15. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating.* New York: Springer Science and Business Media. 2009.

16. Debray TP, Vergouwe Y, Koffijberg H, Nieboer D, Steyerberg EW, Moons KG. A new framework to enhance the interpretation of external validation studies of clinical prediction models. J Clin Epidemiol 2015;68:279-89.
17. Van Calster B, Van Hoorde K, Vergouwe Y, et al. Validation and updating of risk models based on multinomial logistic regression. Diagnostic and Prognostic Research 2017;1:2.
18. Wanner C, Inzucchi SE, Lachin JM, et al. Empagliflozin and progression of kidney disease in type 2 diabetes. N Engl J Med 2016;375:323-34.
19. Perkovic V, de Zeeuw D, Mahaffey KW, et al. Canagliflozin and renal outcomes in type 2 diabetes: results from the CANVAS program randomised clinical trials. Lancet Diabetes Endocrinol 2018;6:691-704.
20. National Institute of Diabetes and Digestive and Kidney Diseases NIoH. United States Renal Data System 2017 Annual Data Report; 2017. , https://www.usrds.org/2017/view/. Accessed July 2018.
21. Levey AS, de Jong PE, Coresh J, et al. The definition, classification, and prognosis of chronic kidney disease: a KDIGO controversies conference report. Kidney Int 2011;80:17-28.