Received: 14 August 2023 | Revised: 28 August 2023 | Accepted: 29 August 2023

DOI: 10.1111/dom.15273

ORIGINAL ARTICLE

WILEY

Derivation and independent validation of kidneyintelX.dkd: A prognostic test for the assessment of diabetic kidney disease progression

Girish N. Nadkarni MD[1,2,3] | Sharon Stapleton PhD[4] | Dipti Takale MSc[5] | Katherine Edwards PhD[4] | Kara Moran PhD[4] | Gohar Mosoyan MSc[1] | Michael K. Hansen PhD[6] | Michael J. Donovan PhD[4] | Hiddo J. L. Heerspink PhD[7] | Fergus Fleming BSc[4] | Steven G. Coca DO[1]

[1]Barbara T Murphy Division of Nephrology, Icahn School of Medicine at Mount Sinai, New York, New York, USA

[2]Division of Digital and Data Driven Medicine, Icahn School of Medicine at Mount Sinai, New York, New York, USA

[3]The Charles Bronfman Institute of Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, New York, USA

[4]Renalytix AI, PLC, New York, New York, USA

[5]Persistent Systems, Pune, India

[6]Janssen Research & Development, LLC, Spring House, Pennsylvania, USA

[7]Department of Clinical Pharmacy and Pharmacology, University of Groningen, Groningen, The Netherlands

Correspondence
Girish N. Nadkarni and Steven G. Coca, Department of Internal Medicine, Icahn School of Medicine at Mount Sinai One Gustave L. Levy Place, Box 1243, New York, NY 10029, USA.
Email: girish.nadkarni@mountsinai.org and steven.coca@mssm.edu

Funding information
Renalytix

Abstract

Aims: To develop and validate an updated version of KidneyIntelX (kidneyintelX.dkd) to stratify patients for risk of progression of diabetic kidney disease (DKD) stages 1 to 3, to simplify the test for clinical adoption and support an application to the US Food and Drug Administration regulatory pathway.

Methods: We used plasma biomarkers and clinical data from the Penn Medicine Biobank (PMBB) for training, and independent cohorts (Bio*Me* and CANVAS) for validation. The primary outcome was progressive decline in kidney function (PDKF), defined by a ≥40% sustained decline in estimated glomerular filtration rate or end-stage kidney disease within 5 years of follow-up.

Results: In 573 PMBB participants with DKD, 15.4% experienced PDKF over a median of 3.7 years. We trained a random forest model using biomarkers and clinical variables. Among 657 Bio*Me* participants and 1197 CANVAS participants, 11.7% and 7.5%, respectively, experienced PDKF. Based on training cut-offs, 57%, 35% and 8% of Bio*Me* participants, and 56%, 38% and 6% of CANVAS participants were classified as having low-, moderate- and high-risk levels, respectively. The cumulative incidence at these risk levels was 5.9%, 21.2% and 66.9% in Bio*Me* and 6.7%, 13.1% and 59.6% in CANVAS. After clinical risk factor adjustment, the adjusted hazard ratios were 7.7 (95% confidence interval [CI] 3.0-19.6) and 3.7 (95% CI 2.0-6.8) in Bio*Me*, and 5.4 (95% CI 2.5-11.9) and 2.3 (95% CI 1.4-3.9) in CANVAS, for high- versus low-risk and moderate- versus low-risk levels, respectively.

Conclusions: Using two independent cohorts and a clinical trial population, we validated an updated KidneyIntelX test (named kidneyintelX.dkd), which significantly enhanced risk stratification in patients with DKD for PDKF, independently from known risk factors for progression.

KEYWORDS
cohort study, diabetic nephropathy, machine learning, observational study, primary care

1 | INTRODUCTION

Chronic kidney disease (CKD) is a major cause of morbidity and mortality, affecting nearly 850 million individuals worldwide, including over 38 million in the United States.[1] The leading cause of CKD in the United States is type 2 diabetes (T2D) and up to 40% of individuals with T2D have CKD, known as diabetic kidney disease (DKD).[2] DKD is defined by the presence of elevated urinary albumin excretion (urinary albumin-to-creatinine ratio [uACR] ≥30 mg/g) or low estimated glomerular filtration rate (eGFR; <60 mL/min/1.73 m^2) in a person with T2D.[2] The clinical course of DKD is highly variable and includes fluctuating levels of albuminuria and a progressive loss of kidney function, represented first as a compensatory increase and then a gradual lowering in eGFR.[3,4]

For many patients the disease process, without early intervention, inevitably progresses to kidney failure, cardiovascular disease, and death.[5] Despite newer kidney-protective medications, only a small proportion of eligible patients are currently on these therapies.[6–8] The reasons for this are multifactorial and include a lack of readily available tools to identify those patients at the highest risk for progression of their kidney disease[9] combined with therapeutic inertia, due in part to the protracted time frame from treatment to an observed individual benefit.

Several blood-based biomarkers reflecting the underlying disease pathophysiology in DKD have been associated with the risk of a progressive decline in kidney function (PDKF).[10–14] Plasma tumour necrosis factor receptor (TNFR)1 and TNFR2 have been identified at the cellular level in kidney endothelial cells, podocytes, and renal tubular epithelial cells. TNFR1 and TNFR2 have been shown to be responsible for the upregulation of proapoptotic signals. Kidney injury molecule (KIM)-1 gene expression is upregulated in cases of ischaemia, hypoxia and cellular tubular injury, and has been implicated in biological mechanisms of CKD in the setting of T2D (eg, activation of phagocytic cells, autophagy, and immune cellular activation).[15] These three biomarkers represent the most studied plasma proteins for predicting CKD outcomes.[11] We have previously demonstrated that the combination of these three biomarkers with clinical variables provides effective risk stratification for kidney outcomes in individuals with T2D and early CKD stages 1 to 3.[16,17]

In the present study, we sought to optimize and validate an updated version of the KidneyIntelX test (kidneyintelX.dkd). The intent was to address specific changes including the removal of race from the eGFR equation,[18,19] and to balance the observed variability and frequently absent clinical features with the importance of the biomarkers, while focusing on time-to-event measures of PDKF. Only contemporary clinical trial biobanks with complete clinical data and follow-up were used to allow for optimal training and validation of models and ensure diverse patient generalizability.

2 | METHODS

In this study we analysed three prospectively collected biobanks linked to clinical data (two observational studies and one randomized control trial) that comprised geographically, demographically, clinically and socio-economically diverse adults with T2D and CKD stages 1 to 3. We aimed to train and validate a revised kidneyintelX.dkd model to predict PDKF within 5 years from baseline. The laboratory assay of the kidneyintelX.dkd test quantitatively measures K2EDTA plasma TNFR1, TNFR2 and KIM-1, and combines these measurements with clinical data to produce a level of risk (low, moderate or high) associated with PDKF for adult patients with T2D and CKD (ie, DKD). All patients had T2D and established DKD at baseline, with an eGFR of 30 to 59 mL/min/1.73 m^2 or an eGFR ≥60 mL/min/1.73 m^2 with albuminuria (uACR ≥30 mg/g).[17,20,21]

2.1 | Study populations

2.1.1 | Penn Medicine Biobank

The kidneyintelX.dkd algorithm was derived from the Penn Medicine Biobank (PMBB). The PMBB is a research cohort enrolled from the University of Pennsylvania Health System, with recruitment from 2008.[17] Participants consented to allow the linkage of biospecimens with their longitudinal electronic health records (EHRs).

2.1.2 | Bio*Me*

The Bio*Me* biobank is a plasma and DNA biorepository, with recruitment from 2008, which includes consented access to the patients' EHRs from a diverse local community in New York City.[17] The acquisition of Bio*Me* donor specimens and de-identified medical records were obtained and approved by institutional review boards for both the Icahn School of Medicine at Mount Sinai and the University of Pennsylvania.

Both Bio*Me* and PMBB are institutional biobanks representative of the outpatient populations of the institutions they serve. Patients are recruited from outpatient general medicine clinics and certain subspecialty clinics with limited preselection criteria.[22,23]

2.1.3 | CANagliflozin cardioVascular Assessment Study (CANVAS)

The CANagliflozin cardioVascular Assessment Study (CANVAS) was a multicentre, double-blinded, placebo-controlled, randomized trial to assess the effect of canagliflozin on primarily cardiovascular events, kidney function, and safety outcomes in patients with T2D who had a history of cardiovascular disease or multiple cardiovascular risk markers, as previously described.[16]

2.2 | Eligibility criteria

Adult patients (≥ 21 years old) from all three cohorts with T2D and a race-free eGFR value between 30 and 59.9 mL/min/1.73 m^2

or an eGFR ≥60 mL/min/1.73 m^2 with uACR ≥30 mg/g within 12 months prior to baseline (defined as time of biobank enrolment in PMBB and Bio*Me* or time of randomization in CANVAS) were included in the analyses. For the validation cohort, in order to eliminate the potential for bias, we excluded subjects who were included in prior validation studies for KidneyIntelX.[23] In all cases, those who did not donate plasma for measurement of biomarkers (n = 93), or were actively treated with etanercept (due to known interference with the assay for TNFR2; n = 2) were excluded. Additional limitations were applied to the Bio*Me* validation population to minimize variability from clinical input features and to best reflect the contemporary clinical status of the patients (Figure S1).

2.2.1 | Biomarker assays

Biomarker measurements were performed using the banked K_2EDTA plasma specimens derived from whole blood. Methods followed standard operating procedures within a Renalytix Clinical Laboratory Improvement Amendments (CLIA)-certified laboratory in accordance with ISO13485:2016 and US Food and Drug Administration (FDA) QSR 820 Quality Management System requirements. The three plasma biomarkers were measured using a proprietary, analytically validated multiplex MSD assay (Meso Scale Diagnostics, Gaithersburg, Maryland), which employs electrochemiluminescence detection methods combined with patterned arrays in combination with the MESO SECTOR S 600 instrument.[21] Each sample was run in duplicate, along with in-line quality control samples with known low, moderate and high concentrations of each biomarker on each plate. The laboratory personnel performing the biomarker assays were blinded to all clinical information.

2.2.2 | Ascertainment and definition of the kidney endpoint

We determined longitudinal eGFR values post-baseline using the new race-free CKD-EPI creatinine equation,[18] derived from serum creatinine, age and gender. The primary composite outcome, PDKF, included the following: a sustained decline in eGFR of ≥40% from baseline (defined as individuals who had ≥2 follow-up eGFRs that are ≥3 months apart with ≥40% decrease within 5 years from baseline), or end-stage kidney disease (defined as individuals who had ≥2 eGFR measurements of <15 mL/min/1.73 m^2 that are ≥30 days apart post enrolment). Follow-up time was censored after loss to follow-up or 5 years after baseline in PMBB and Bio*Me*. Additional follow-up for up to 7 years was applied for CANVAS data to allow for more complete ascertainment of sustained outcomes from available eGFR datapoints. The data analysts responsible for ascertainment of the kidney endpoint in the validation cohorts were blinded to baseline clinical and biomarker data.

2.2.3 | Model building: training cohort

Model training was completed in a cohort of 573 subjects from the PMBB who met the intended-use criteria. We considered demographics, laboratory variables, diagnostic codes, medication use, and concentrations of the three plasma biomarkers for training of the model. We used a random forest model and tuned the hyperparameters to estimate the 5-year risk for kidney outcome. We calculated the discrimination (c-statistics) and quantified the prediction model's ability to separate those who experience the kidney outcome from those who do not. Each incremental model's accuracy was assessed based on the 10-fold cross validation (70%/30% random split) to arrive at candidate algorithms comprising different feature sets.

2.2.4 | Hyperparameter determination

We used a standard machine-learning workflow to set hyperparameters. The following hyperparameters along with their description and values were utilized in the algorithm: (i) number of trees (n = 100), defined as the number of decision trees in the ensemble; (ii) node split (n = 2), defined as the number of variables tried at each split; (iii) terminal node side (n = 2), defined as the minimum number of observations in a terminal node; and (iv) node depth (n = 2), defined as the number of splits a tree in the ensemble can make before coming to a prediction.

The final set of input features included in the algorithm were the following: plasma TNFR1, plasma TNFR2, plasma KIM-1, baseline uACR, blood urea nitrogen (BUN) and glycated haemoglobin (HbA1c), and category cut-offs were defined and fixed prior to validation testing. Those with a predicted probability of ≤0.100 were categorized as low risk, those with a predicted probability >0.100 and <0.300 were categorized as moderate risk, and those with a predicted probability ≥0.300 were categorized as high risk.

2.3 | Statistical analyses

We then applied the fixed algorithm to the independent validation cohorts of Bio*Me* (n = 657) and CANVAS (n = 1197). We applied cumulative incidence analyses to determine the failure estimates for the outcome for each of the kidneyintelX.dkd test risk categories to account for censoring during the 5-year observation period. Proportional hazards assumptions were tested. Unadjusted and adjusted Cox proportional hazard models were applied to show the relative hazard of patients experiencing PDKF by kidneyintelX.dkd levels (high vs. low and moderate vs. low) over a 5-year period, with adjustment for well-known risk factors (age, gender, race, and baseline eGFR, uACR, systolic blood pressure, HbA1c). We then assessed calibration (predicted vs. actual risk) in all the validation cohorts using the Brier scores to assess concordance between predicted and actual risk. The Brier score is a scoring rule that measures the accuracy of

probabilistic predictions. The scores range from 0 to 1 and the lower the Brier score for a set of predictions, the better the predictions are calibrated. Finally, in exploratory analyses, we assessed the median eGFR slope per kidneyintelX.dkd risk category using mixed linear models and examined the performance of kidneyintelX.dkd by subgroups of age, race, CKD stage, and period of enrolment into the biobank.

For the Bio*Me* validation cohort, all statistical analyses were performed by independent statisticians (Advance Research Associates, Inc., California) according to a prespecified plan. To ensure study results were not subject to potential bias, blinding procedures were followed throughout the study. The estimated absolute event rate for sample size calculations was assumed for high-risk level >30% and for low-risk level <7.5%, and for prespecified target difference in event rates observed between the kidneyintelX.dkd levels.

3 | RESULTS

3.1 | Population characteristics

Baseline characteristics for study participants are summarized in Table S1. In the PMBB cohort, there were 573 participants; the mean (SD) age was 62 (10) years and 43% were women. The mean (SD) eGFR was 58 (20) mL/min/1.73 m^2 and the median (interquartile range [IQR]) uACR was 54 (14-221) mg/g. In the Bio*Me* cohort, 657 participants were included in the analysis based on prespecified inclusion/exclusion criteria; the mean (SD) age was 70 (11) years and 57% were women. The mean (SD) eGFR was 64 (22) mL/min/1.73 m^2 and the median (IQR) uACR was 52 (20-214) mg/g. Finally, in the CANVAS cohort there were 1197 participants with DKD at the time of randomization, all meeting the inclusion criteria based on baseline eGFR and uACR; the mean (SD) age was 64 (8) years and 31% were women. The mean (SD) eGFR was 75 (21) mL/min/1.73 m^2 and the median (IQR) uACR was 65 (34-189) mg/g.

3.2 | Model training

Over a median follow-up period of 3.2 years, there were 88 composite kidney events in the PMBB cohort. Using random forest optimization and hyperparameter tuning as described in the methods to derive the prediction model, the following variables were retained as independent risk factors for the composite kidney outcome in the PMBB cohort: plasma TNFR1, plasma TNFR2, plasma KIM-1, baseline uACR, BUN and HbA1c. In addition to their contribution as independent risk factors, the clinical features selected are all clinically relevant to the intended use and therefore contribute to generalizability to broad populations. The relative importance of each feature is shown in Figure S2. The c-statistic of the model for the discrimination of the kidney outcome was 0.83.

3.3 | Performance of the risk prediction model for the kidney outcome in the validation cohort (Bio*Me*)

Over a median follow-up period of 1385 days, there were 77 composite kidney events in Bio*Me*. The locked model derived from PMBB and applied to the Bio*Me* population scored 57% of the participants as low risk, 35% as moderate risk, and 8% as high risk. The cumulative incidence probabilities in the low-, moderate- and high-risk levels were 6%, 21% and 67%, respectively (Figure 1B and Table 1). The unadjusted hazard ratio (HR) for the kidney outcome was 18.3 (95% CI 9.7-34.6) for high versus low risk and 4.2 (95% CI 2.4-7.4) for moderate versus low risk. After adjustment for age, gender, race and baseline eGFR, uACR, systolic blood pressure and HbA1c, the adjusted HRs were 7.7 (95% CI 3.0-19.6) and 3.7 (95% CI 2.0-6.8), respectively (Table S2 and Figure 2). The magnitude of associations was consistent across key subgroups (Figure S3). The eGFR slopes for the low-, moderate- and high-risk levels were −0.91, −1.80 and −3.69 mL/min/1.73 m^2 per year, respectively (ANOVA test, $P < 0.001$; Table S3 and Figure 3).

3.4 | Performance of the risk prediction model for the kidney outcome in CANVAS

Over a median follow-up period of 6.1 years (73.4 months, n = 1197) there were 90 (7.5%) composite kidney events in those with baseline CKD G1 to G3b in CANVAS (n = 1197). The kidneyintelX.dkd model, derived from PMBB, scored 56% of the participants as low risk, 38% as moderate risk and 6% as high risk (Table 1). The cumulative incidence probabilities in the low-, moderate- and high-risk levels were 6.7%, 13.1% and 59.6%, respectively (Figure 1C and Table 1). The unadjusted HR for the kidney outcome was 13.8 (95% CI 7.9-24) for high versus low risk and 2.6 (95% CI 1.5-4.2) for moderate versus low risk. After adjustment for age, gender, race and baseline eGFR, uACR, systolic blood pressure and HbA1c, the HR for the kidney outcome was 5.4 (95% CI 2.5-11.9) in the high- versus low-risk levels and 2.3 (95% CI 1.2-4.4) in the moderate- versus low-risk levels (Table S3 and Figure 2). The eGFR slopes for low-, moderate- and high-risk levels were − 0.23, −0.60 and − 2.05 mL/min/1.73 m^2 per year, respectively (ANOVA test, $P < 0.001$; Table S3 and Figure 3). The absolute magnitudes of the eGFR slopes by risk category were more pronounced in the placebo group compared to the canagliflozin-treated group (−0.61, −0.95 and −3.45 vs. −0.04, −0.43 and −1.35 mL/min/1.73 m^2, respectively).

3.4.1 | Calibration in the validation cohorts

Brier scores were calculated in the validation cohorts (Table S4). The results demonstrated that the kidneyintelX.dkd model was well calibrated across the validation cohorts, with good concordance between the predicted and actual risks (Figures S4 and S5).

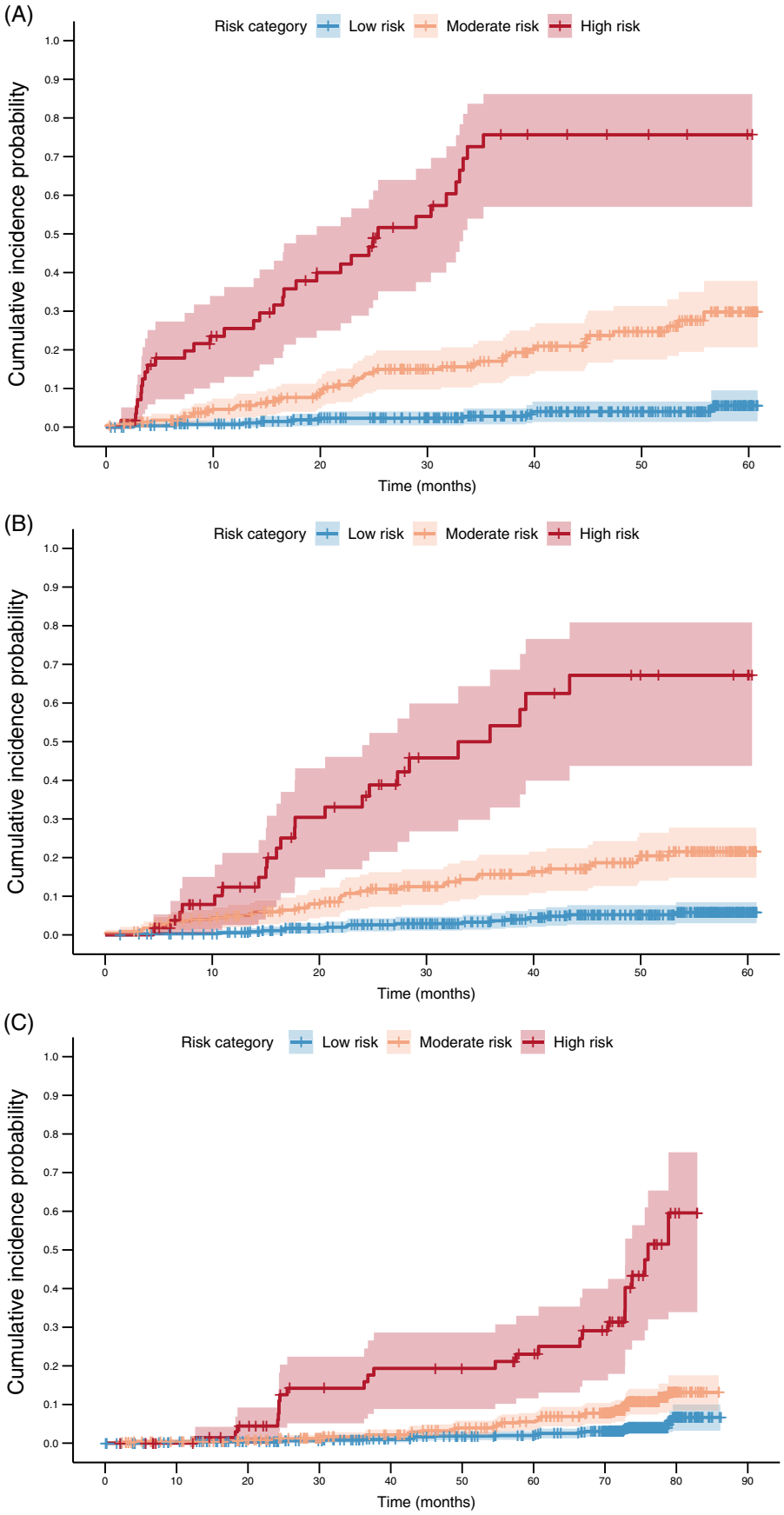

FIGURE 1 Cumulative incidence curves for the Penn Medicine Biobank (A), Bio*Me* (B) and CANVAS (C) cohorts across the follow-up period.

TABLE 1 Comparison of the performance characteristics across the three cohorts

	Number of subjects[a] n (%)			Cumulative incidence probability accelerated PDKF over 5 years (PMBB, Bio*Me*) uncensored data of CANVAS, % (95% CI)		
kidneyIntelX.dkd level	PMBB	Bio*Me*	CANVAS	PMBB	Bio*Me*	CANVAS
Low	289 (50.4)	374 (56.9)	666 (55.6)	5.6 (2.7–11.3)	5.9 (3.6–9.4)	6.7 (3.2–10.0)
Moderate	227 (39.6)	228 (34.7)	456 (38.1)	30.5 (22.7–40.1)	21.2% (15.6–28.3)	13.1 (8.4–17.6)
High	57 (10.0)	55 (8.4)	75 (6.3)	84.9 (70.0–96.1)	66.9% (49.3–83.5)	59.6 (33.9–75.3)

Abbreviations: CI, confidence interval; PDKF, progressive decline in kidney function; PMBB, Penn Medicine Biobank.
[a]Group sizes by the three risk levels in the validation cohorts determined by application of the predicted probabilities of ≤0.100 (low risk), >0.100 and <0.300 (moderate risk) and ≥0.300 (high risk).

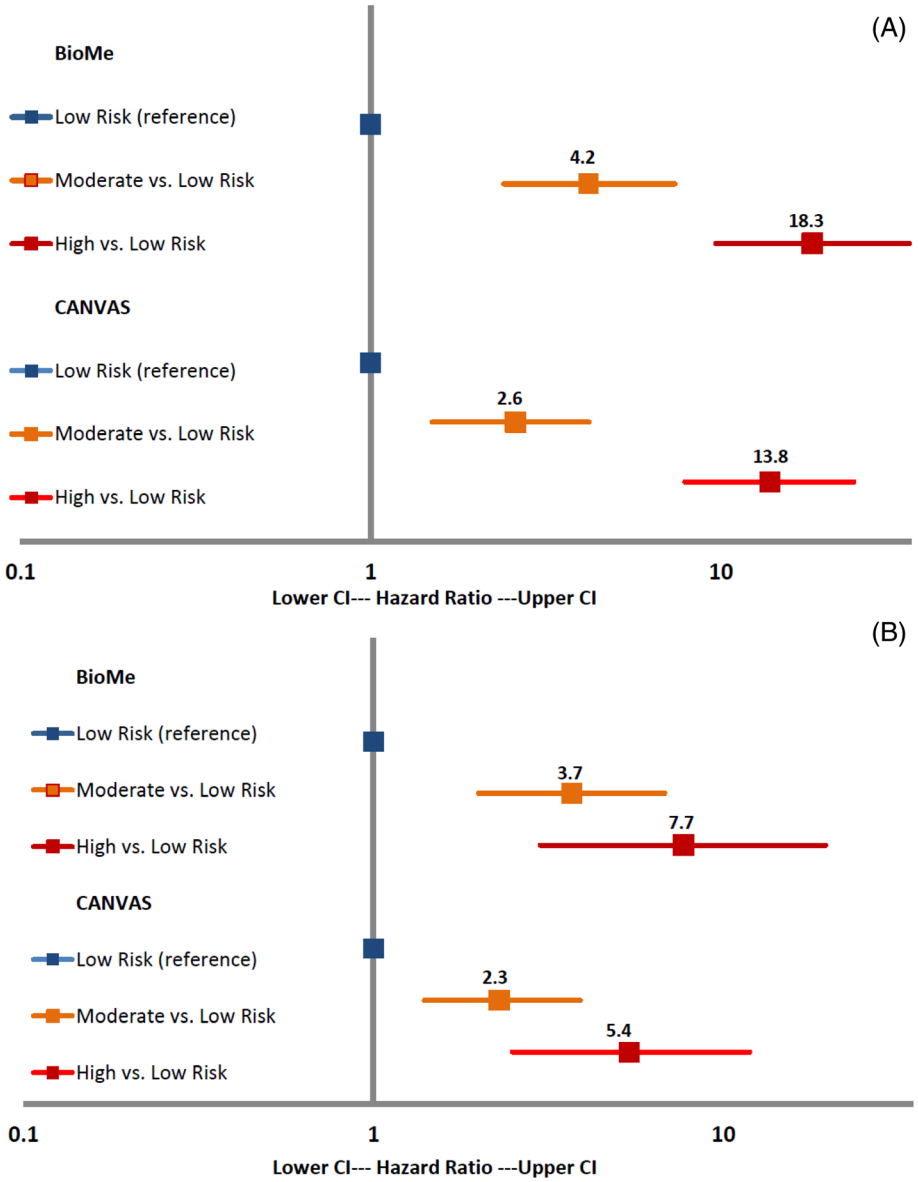

FIGURE 2 Unadjusted (A) and adjusted (B) hazard ratios by kidneyintelX.dkd level in Bio*Me* and CANVAS. Adjusted hazard ratios (compared to low-risk level) derived from Cox proportional hazard models with kidneyintelX.dkd levels adjusting for age, gender, race, and baseline estimated glomerular filtration rate, urinary albumin-to-creatinine ratio, systolic blood pressure and glycated haemoglobin.

3.5 | Supplementary analyses

We performed simulations to evaluate the generalizability of the performance of kidneyintelX.dkd to patients tested from a wide variety of health settings. All input variables for all participants in the clinical validation cohort (n = 657) were simultaneously and randomly varied based on the measured or expected precision profile for each feature (n = 100) for a total of 65 700 simulations. The ranges chosen for the

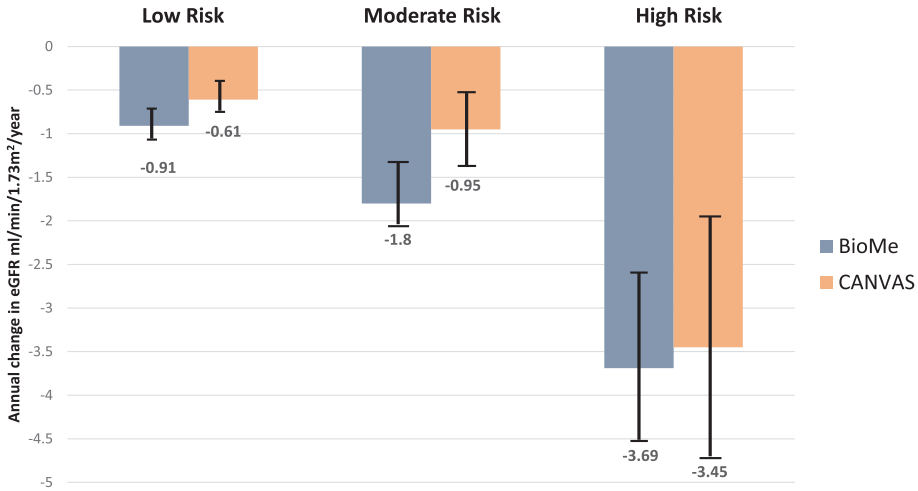

FIGURE 3 Annual rate of decline in estimated glomerular filtration rate (eGFR) by kidneyintelX.dkd risk level in Bio*Me* and CANVAS derivation and validation cohorts. Data shown are the mean annual change in eGFR along with standard deviations.

three biomarkers were based on imprecision studies performed at Renalytix. For the clinical features (uACR, HbA1c and BUN), data from published proficiency testing and from decision summaries from FDA-cleared devices were applied to reflect the expected imprecision in these features. The maximum level of re-categorization in the presence of this level of imprecision in the input features was 5.1% across all simulations for all risk categories, with no subjects moving by more than one kidneyintelX.dkd category (ie, high to low or low to high; Table S5) The robustness of the kidneyintelX.dkd result was further demonstrated by increasing the imprecision of the clinical features (uACR to 40%, HbA1c to 6% and BUN to 14%) based on reported College of American Pathologists survey data and some published literature for uACR intra-individual variability. The results showed a marginal increase in the overall classification rate (3.9%), with no change in the median coefficient of variation (1.1%; IQR 0.5%, 2.9%) across 100 simulations for each of the 657 participants.

We further performed sensitivity analysis to assess performance across multiple subgroups in the Bio*Me* validation population based on age, gender, race, baseline eGFR range, and period of enrolment in the Bio*Me* biobank (reflecting temporal patterns of clinical care). kidneyintelX.dkd performance in risk event stratification of low-, moderate- and high-risk patients was consistent across all subgroups.

4 | DISCUSSION

There are over 12 million individuals in the United States with DKD, but only a small proportion will experience progression over a multi-year timeframe.[5] With limited resources and healthcare services, it is critical to identify patients who have a higher risk of progression. Interventions, including novel kidney and cardioprotective medications, multi-drug combinations and specialist referral, can then be prioritized for these high-risk individuals to decrease the overall burden of advanced-stage kidney disease. However, current methods are inadequate to address this gap in knowledge. In response, we previously developed and validated the KidneyIntelX test, combining novel biomarkers and clinical risk factors.[17] Due to contemporary changes in clinical management, we further developed and then validated an updated version of KidneyIntelX, supporting a de novo marketing authorization from the FDA for the test, named kidneyintelX.dkd (DEN200052).

kidneyintelX.dkd is designed to be straightforward to implement and interpret in clinical practice, and has been shown to be generalizable to broad, diverse populations. The base performance of kidneyintelX.dkd is similar to that of KidneyIntelX,[17] the unadjusted HR for kidneyintelX.dkd high versus low risk was 18 for Bio*Me* and 14 for CANVAS, with cumulative incidence probabilities in the high-risk level that were 60% or greater. The gradients of risk between high- and low-risk strata were consistent in both validation cohorts despite differences in baseline eGFR, with that in CANVAS being significantly higher (adjusted HRs of 8 and 5 in Bio*Me* and CANVAS for high vs. low risk, respectively). Importantly, the models were well calibrated across the independent external validation cohorts, and moreover, through extensive simulation and sensitivity testing, kidneyintelX.dkd was shown to be robust against large variations in the standard clinical variables used in the model, and prognostic performance was consistent across patient subgroups. The ability to risk-stratify a large population of patients with DKD for such an important clinical outcome has the potential to result in significant clinical decisions and actions at both ends of the risk spectrum. For high-risk patients there is an opportunity to change and conceivably slow the current trajectory of PDKF through consult services and novel therapies, while low-risk patients could avoid polypharmacy and promote continued weight loss programmes and effective diabetes care. Indeed, those scored as low risk in the cohorts had an absolute risk of progression of approximately 5% over 5 years, with eGFR declines approximately equal to the rate due to normal aging, which are compelling datapoints to avoid aggressive care in those at low risk. Inclusion of time-to-event outcomes ensures that clinicians are provided with performance information to support the interpretation of test results. We performed additional analyses on KidneyIntelX which have demonstrated that cost savings are derived from deployment of KidneyIntelX testing for a large population of patients with DKD, compared to the standard of care.[24]

This work should be interpreted in the light of some limitations. First, we did not perform an a priori sample size determination for the

derivation cohort. However, for the validation cohorts, we empirically demonstrate a sufficient sample size to show risk difference between the high- and low-risk groups. Second, eGFR was not a final feature in the kidneyintelX.dkd model which may seem counterintuitive for a test to predict kidney function decline. We included both eGFR and BUN concentration, a surrogate measure of kidney function, in consecutive models and, when used in the composite kidneyintelX.dkd models, BUN contributed more to prediction of eGFR decline than eGFR itself. Third, we only required a single measurement of uACR for inclusion, however, usually two or more are needed for definitive diagnosis. Fourth, we did not adjust for diabetes duration, antihypertensive drug use or smoking in the fully adjusted models as these parameters are not captured accurately or routinely in patient health records. Finally, the overall event rate in CANVAS was lower than that in the two observational cohorts, which limited the absolute cumulative incidence probability in the high-risk group in CANVAS. However, this was expected for three reasons: (i) clinical trial population (selection bias); (ii) treatment effect 2:1 randomization to the very effective sodium-glucose cotransporter-2 inhibitor canagliflozin; (iii) only once-yearly ascertainment of eGFR, which limits association with outcome.

In conclusion, an updated version of the prior KidneyIntelX test, named kidneyintelX.dkd, was trained and validated in two contemporary cohorts. The assay demonstrated excellent prognostic performance for PDKF, independent of key demographics and clinical variables, with wide-ranging analytical stability, and consistent risk discrimination in the intended-use population of DKD overall, as well across key subgroups.

ACKNOWLEDGMENTS

We wish to thank Ya Chen Tonar for her invaluable contributions.

FUNDING INFORMATION

The study was funded by Renalytix, but validation was conducted by an independent third-party statistician according to a prespecified plan to meet FDA requirements.

CONFLICT OF INTEREST STATEMENT

Girish N. Nadkarni reports consultancy agreements with AstraZeneca, BioVie, GLG Consulting, Pensieve Health, Reata, Renalytix, Siemens Healthineers, G.S.K Pharma and Variant Bio, research funding from Goldfinch Bio and Renalytix, honoraria from AstraZeneca, BioVie, Lexicon and Reata, and patents or royalties for Renalytix, owns equity and stock options in Pensieve Health as a cofounder and in Renalytix, has received financial compensation as a scientific board member and advisor to Renalytix, serves on the advisory board of Neurona Health, and serves in an advisory or leadership role for Pensieve Health and Renalytix. Michael J. Donovan, Fergus Fleming, Sharon Stapleton, Katherine Edwards and Kara Moran are employees of Renalytix. Gohar Mosoyan serves as a consultant to Renalytix. Dipti Takale is an employee of Persistent Systems. Michael K. Hansen is an employee of Janssen Research & Development, LLC. Hiddo J. L. Heerspink is supported by a VIDI (917.15.306) grant from the Netherlands Organization for Scientific Research and has served as a consultant for AbbVie, Astellas, AstraZeneca, Bayer, Boehringer Ingelheim, Chinook, CSL Pharma, Fresenius, Gilead, Janssen, Merck, Mundipharma, Mitsubishi Tanabe and Retrophin, and has received grant support from AbbVie, AstraZeneca, Boehringer Ingelheim and Janssen. Steven G. Coca has received consulting fees from Renalytix, Takeda, Nuwellis, Vifor, Bayer, Boehringer-Ingelheim, Reprieve Cardiovascular, Axon and 3ive, and financial compensation as a scientific board member and advisor to Renalytix, and owns equity in Renalytix. The study was funded by Renalytix, but validation was conducted by an independent third-party statistician according to a prespecified plan to meet FDA requirements. Girish N. Nadkarni serves as the guarantor of this article.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/dom.15273.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Girish N. Nadkarni https://orcid.org/0000-0001-6319-4314
Hiddo J. L. Heerspink https://orcid.org/0000-0002-3126-3730

REFERENCES

1. Kibria GMA, Crispen R. Prevalence and trends of chronic kidney disease and its risk factors among US adults: an analysis of NHANES 2003-18. *Prev Med Rep*. 2020;20:101193.
2. Rossing P, Caramori ML, Chan JCN, et al. Executive summary of the KDIGO 2022 clinical practice guideline for diabetes Management in Chronic Kidney Disease: an update based on rapidly emerging new evidence. *Kidney Int*. 2022;102:990-999.
3. Oshima M, Shimizu M, Yamanouchi M, et al. Trajectories of kidney function in diabetes: a clinicopathological update. *Nat Rev Nephrol*. 2021;17:740-750.
4. Yang Y, Xu G. Update on pathogenesis of glomerular hyperfiltration in early diabetic kidney disease. *Front Endocrinol (Lausanne)*. 2022;13:872918.
5. Mima A. A narrative review of diabetic kidney disease: previous and current evidence-based therapeutic approaches. *Adv Ther*. 2022;39:3488-3500.
6. Eberly LA, Yang L, Eneanya ND, et al. Association of Race/ethnicity, gender, and socioeconomic status with sodium-glucose cotransporter 2 inhibitor use among patients with diabetes in the US. *JAMA Netw Open*. 2021;4:e216139.
7. Zhuo M, Li J, Buckley LF, et al. Prescribing patterns of sodium-glucose Cotransporter-2 inhibitors in patients with CKD: a cross-sectional registry analysis. *Kidney360*. 2022;3:455-464.
8. Gregg LP, Ramsey DJ, Akeroyd JM, et al. Predictors, disparities, and facility-level variation: SGLT2 inhibitor prescription among US veterans with CKD. *Am J Kidney Dis*. 2023;82:53-62.e1.
9. Vassalotti JA, Boucree SC. Integrating CKD into US primary care: bridging the knowledge and implementation gaps. *Kidney Int Reports*. 2022;7:389-396.
10. Sarnak MJ, Katz R, Ix JH, et al. Plasma biomarkers as risk factors for incident CKD. *Kidney Int Rep*. 2022;7:1493-1501.

11. Liu C, Debnath N, Mosoyan G, et al. Systematic review and meta-analysis of plasma and urine biomarkers for CKD outcomes. *J Am Soc Nephrol*. 2022;33:1657-1672.
12. Gutierrez OM, Shlipak MG, Katz R, et al. Associations of plasma biomarkers of inflammation, fibrosis, and kidney tubular injury with progression of diabetic kidney disease: a cohort study. *Am J Kidney Dis*. 2022;79:849-857.e1.
13. Zabetian A, Coca SG. Plasma and urine biomarkers in chronic kidney disease: closer to clinical application. *Curr Opin Nephrol Hypertens*. 2021;30:531-537.
14. Schrauben SJ, Shou H, Zhang X, et al. Association of Multiple Plasma Biomarker Concentrations with progression of prevalent diabetic kidney disease: findings from the chronic renal insufficiency cohort (CRIC) study. *J Am Soc Nephrol*. 2021;32:115-126.
15. Brooks CR, Bonventre JV. KIM-1/TIM-1 in proximal tubular cell immune response. *Oncotarget*. 2015;6:44059-44060.
16. Lam D, Nadkarni GN, Mosoyan G, et al. Clinical utility of KidneyIntelX in early stages of diabetic kidney disease in the CANVAS trial. *Am J Nephrol*. 2022;53:21-31.
17. Chan L, Nadkarni GN, Fleming F, et al. Derivation and validation of a machine learning risk score using biomarker and electronic patient data to predict progression of diabetic kidney disease. *Diabetologia*. 2021;64:1504-1515.
18. Inker LA, Eneanya ND, Coresh J, et al. New creatinine- and cystatin C-based equations to estimate GFR without race. *N Engl J Med*. 2021;385:1737-1749.
19. Delgado C, Baweja M, Crews DC, et al. A unifying approach for GFR estimation: recommendations of the NKF-ASN task force on reassessing the inclusion of race in diagnosing kidney disease. *Am J Kidney Dis*. 2022;79:268-288.e1.
20. Tokita J, Vega A, Sinfield C, et al. Real world evidence and clinical utility of KidneyIntelX on patients with early-stage diabetic kidney disease: interim results on decision impact and outcomes.*J Prim Care Community Health*. 2022;13: 21501319221138196.
21. Connolly P, Stapleton S, Mosoyan G, et al. Analytical validation of a multi-biomarker algorithmic test for prediction of progressive kidney function decline in patients with early-stage kidney disease. *Clin Proteomics*. 2021;18:26.
22. Tayo BO, Teil M, Tong L, et al. Genetic background of patients from a university medical center in Manhattan: implications for personalized medicine. *PloS One*. 2011;6:e19166.
23. Bajaj A, Ihegword A, Qiu C, et al. Phenome-wide association analysis suggests the APOL1 linked disease spectrum primarily drives kidney-specific pathways. *Kidney Int*. 2020;97:1032-1041.
24. Datar M, Burchenal W, Donovan MJ, Coca SG, Wang E, Goss TF. Payer budget impact of an artificial intelligence in vitro diagnostic to modify diabetic kidney disease progression. *J Med Econ*. 2021;24:972-982.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Nadkarni GN, Stapleton S, Takale D, et al. Derivation and independent validation of kidneyintelX. dkd: A prognostic test for the assessment of diabetic kidney disease progression. *Diabetes Obes Metab*. 2023;1-9. doi:10.1111/dom.15273