Validation of Risk Equations for Complications of Type 2 Diabetes (RECODe) Using Individual Participant Data From Diverse Longitudinal Cohorts in the U.S.

Sanjay Basu,[1,2] Jeremy B. Sussman,[3,4] Seth A. Berkowitz,[5,6] Rodney A. Hayward,[3,4] Alain G. Bertoni,[7] Adolfo Correa,[8] Stanford Mwasongwe,[9] and John S. Yudkin[10]

OBJECTIVE

We sought to validate Risk Equations for Complications of Type 2 Diabetes (RECODe) among diverse populations.

RESEARCH DESIGN AND METHODS

We compared risk predictions from RECODe equations and from two alternative risk models (UK Prospective Diabetes Study Outcomes Model 2 [UKPDS OM2] and American College of Cardiology/American Heart Association Pooled Cohort Equations) to observed outcomes in two studies: the Multi-Ethnic Study of Atherosclerosis (MESA, *n* = 1,555 adults with type 2 diabetes, median follow-up 9.1 years) and the Jackson Heart Study (JHS, *n* = 1,746 adults with type 2 diabetes, median follow-up 8.0 years). Outcomes included nephropathy by multiple measures (microalbuminuria, macroalbuminuria, renal failure, end-stage renal disease, and reduction in glomerular filtration rate), moderate to severe diabetic retinopathy by Airlie House classification, fatal or nonfatal myocardial infarction, fatal or nonfatal stroke, congestive heart failure, and all-cause mortality.

RESULTS

RECODe equations for microvascular and cardiovascular outcomes had C-statistics for discrimination ranging from 0.71 to 0.85 in MESA and 0.64 to 0.91 in JHS for alternative outcomes. Calibration slopes in MESA ranged from 0.62 for a composite nephropathy outcome, 0.83–1.04 for individual nephropathy outcomes, 1.07 for retinopathy, 1.00–1.05 for cardiovascular outcomes, and 1.03 for all-cause mortality. Slopes in JHS ranged from 0.47 for retinopathy, 0.97–1.16 for nephropathy, 0.72–1.05 for cardiovascular outcomes, and 1.01 for all-cause mortality. The alternative models had C-statistics 0.50–0.72 and calibration slopes 0.07–0.60.

CONCLUSIONS

RECODe equations improved risk estimation for diverse patients with type 2 diabetes, as compared with two commonly used alternatives.

[1]Center for Primary Care and Outcomes Research, Center for Population Health Sciences, Departments of Medicine and Health Research and Policy, Stanford University, Stanford, CA
[2]Center for Primary Care, Harvard Medical School, Boston, MA
[3]Division of General Medicine, University of Michigan, Ann Arbor, MI
[4]Center for Clinical Management Research, Veterans Affairs Ann Arbor Healthcare System, Ann Arbor, MI
[5]Division of General Internal Medicine and Diabetes Unit, Massachusetts General Hospital, Boston, MA
[6]Department of Medicine, Harvard Medical School, Boston, MA
[7]Department of Epidemiology and Prevention, Wake Forest School of Medicine, Winston-Salem, NC
[8]Departments of Medicine and Pediatrics, University of Mississippi Medical Center, Jackson, MS
[9]School of Public Health, Jackson State University, Jackson, MS
[10]Institute of Cardiovascular Science, Division of Medicine, University College London, London, U.K.

Corresponding author: Sanjay Basu, basus@stanford.edu.

Risk equations for complications of type 2 diabetes can help guide clinical decisions and population health management, such as to identify patients for intensive outreach (1,2), particularly as hemoglobin A_{1c} alone is a limited predictor of outcomes (3). Risk equations can facilitate decision making (1,4) and assist comparative effectiveness research on clinical guidelines (5–7).

Previous risk equations systematically misestimate both microvascular and cardiovascular complications among modern populations (8–10). We recently derived new Risk Equations for Complications of Type 2 Diabetes (RECODe) (11) (Supplementary Tables 1 and 2). RECODe equations were derived from the Action to Control Cardiovascular Risk in Diabetes (ACCORD) study (12–14) and validated with the Diabetes Prevention Program Outcomes Study (DPPOS) (15) and Look AHEAD (Action for Health in Diabetes) (16). We observed that predictions from the RECODe equations improved discrimination and calibration as compared with the UK Prospective Diabetes Study Outcomes Model 2 (UKPDS OM2; for microvascular and cardiovascular outcomes) (17) and the American College of Cardiology (ACC)/American Heart Association (AHA) Pooled Cohort Equations (PCEs; for atherosclerotic cardiovascular disease [ASCVD] events) (18).

Because the development and validation of the RECODe equations were performed using data from randomized clinical trials, the equations may have limited generalizability. Here, we sought to evaluate the validity of the RECODe equations against two longitudinal cohort studies: the Multi-Ethnic Study of Atherosclerosis (MESA) and the Jackson Heart Study (JHS). MESA and JHS were chosen because they included rigorous adjudication of both microvascular and cardiovascular outcomes; included modern populations among whom older equations are thought to be miscalibrated (8–10); included minorities; and have made de-identified individual participant data available for research use, which improves research transparency and reproducibility.

RESEARCH DESIGN AND METHODS

Sources of Data for Validation of Equations

We used data from two longitudinal, observational cohort studies: MESA (2000–2012, n = 1,555 people with type 2 diabetes) and JHS (2000–2012, n = 1,746 people with type 2 diabetes). See the Supplementary Data for the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) checklist (19).

Participants

MESA participants were selected from six general population geographic areas across the U.S. MESA participants experienced baseline examinations from July 2000 to July 2002, three subsequent examinations each 17–20 months apart, and a fifth follow-up examination between April 2010 and January 2012, with follow-up phone calls between examinations. Eligibility criteria included age 45–84 years and the absence of the following: prior physician diagnosis of angina, myocardial infarction (MI), stroke, or congestive heart failure (CHF); atrial fibrillation at enrollment; prior cardiovascular procedure; active cancer; pregnancy; serious medical condition preventing long-term participation; weight >136 kg (300 lb); cognitive incapacity; living in a nursing home or planning to leave the community within 5 years; inability to communicate in English, Spanish, Cantonese, or Mandarin; and chest computed tomography scan in the year prior to enrollment (20).

JHS participants were recruited via community-based sampling from the urban and rural African American population of three counties in Mississippi, U.S., and included a subset of participants previously in the Atherosclerosis Risk in Communities (ARIC) study. JHS participants experienced baseline examinations from 2000 to 2004, a second examination from 2005 to 2008, and a third from 2009 to 2012, with follow-up phone calls between examinations. Eligibility criteria included the following: age 35–84 years, except in a nested family cohort subset, for which eligibility included people 21–34 years old (21).

Outcomes

Separate RECODe equations were developed for each of several microvascular and cardiovascular outcomes, with unique equations for each alternative definition of each outcome. Participants with each of the outcomes at baseline were excluded, to focus on incident outcomes.

Microvascular outcomes included nephropathy and retinopathy; neuropathy was not assessed in MESA or JHS. Separate RECODe equations were evaluated for each of several outcome definitions of nephropathy: 1) development of microalbuminuria (albumin-to-creatinine ratio ≥30 mg albumin per gram creatinine in urine obtained in exams 1, 2, 3, and 5 in MESA and exams 1–3 in JHS); 2) development of macroalbuminuria (albumin-to-creatinine ratio ≥300 mg albumin per gram creatinine in urine); 3) renal failure or end-stage renal disease (ESRD; dialysis) or serum creatinine >291.7 μmol/L (3.3 mg/dL, obtained in exams 1, 3, 4, and 5 in MESA and 1–3 in JHS); 4) doubling of serum creatinine or >20 mL/min decrease in estimated glomerular filtration rate (eGFR) based on the Modification of Diet in Renal Disease (MDRD) study equation (22); 5) any of the outcomes 1, 3, or 4; and 6) an additional composite outcome for any of 1, 2, or 3. The RECODe equation for diabetic retinopathy was evaluated against the diagnosis of new moderate to severe diabetic retinopathy in either eye, adjudicated by ophthalmologists from bilateral retinal photographs using the Airlie House classification system modified for the Early Treatment of Diabetic Retinopathy Study (ETDRS) (23). The photographs were obtained through ancillary studies on a subset of participants in both MESA (a subset of n = 949 subjects with retinal photographs with exam 2 and n = 905 with exam 5) and JHS (a subset of n = 886 subjects with diabetes or impaired fasting glucose photographed with exam 3) (24,25). Further details about the outcome examinations, definitions, and timing are provided in Table 2. Our analysis was limited to incident microvascular events, meaning that participants who lacked evidence of the outcome at the first examination after diagnosis of diabetes and had the outcome at subsequent examinations were treated as incident. For the retinopathy outcome in JHS, people without a prior history of self-reported vision loss or evidence of prior photocoagulation or vitrectomy at retinal exam were included as incident cases if they had retinopathy on exam 3.

Separate RECODe equations were also evaluated for each of several cardiovascular outcome definitions: 1) composite ASCVD (defined as first fatal or nonfatal MI or stroke); 2) fatal or nonfatal MI; 3) fatal or nonfatal stroke; 4) CHF; or 5) death from any cardiovascular cause, each assessed during every MESA and JHS follow-up period, with exact date of events obtained from health care records. As with the microvascular outcomes, our analyses were limited to incident cardiovascular events after diabetes diagnosis.

We additionally assessed the RECODe equation for all-cause mortality (including both cardiovascular and noncardiovascular mortality).

Predictors

Predictors included age (in years), sex, race/ethnicity (black race and/or Hispanic ethnicity), current medication use (each of the following: statin, oral diabetes medications including metformin, and anticoagulants other than aspirin), current tobacco smoking, systolic blood pressure, and biomarkers (total cholesterol, HDL cholesterol, urine creatinine, urine microalbumin-to-creatinine ratio, and hemoglobin A_{1c}) (see Table 1). The first available value for each predictor after diagnosis of diabetes was entered into the RECODe equations. Time-varying covariates were not included because the RECODe equations were intended for use in clinical settings to assist initial treatment decisions.

Sample Size

The final sample size of participants (n = 1,555 from MESA and n = 1,746 from JHS) included those with type 2 diabetes. Type 2 diabetes diagnosis was based on prior physician diagnosis, use of insulin or oral hypoglycemia medications without a history of type 1 diabetes, or 2010 American Diabetes Association diagnostic criteria (fasting plasma glucose $\geq$7.0 mmol/L [126 mg/dL], 2-h plasma glucose $\geq$11.1 mmol/L [200 mg/dL] during oral glucose tolerance test, hemoglobin A_{1c} $\geq$48 mmol/mol [6.5%], or random plasma glucose $\geq$11.1 mmol/L [200 mg/dL] with accompanying hyperglycemia symptoms) (26). Further characteristics of people with type 2 diabetes in each sample are provided in Table 2.

Missing Data

In the MESA sample, 11% of participants (n = 172) with type 2 diabetes had one or more predictor variables missing. In the JHS sample, 7% of participants (n = 128) with type 2 diabetes had one or more predictor variables missing. Missing values for predictor variables in both samples were imputed with fivefold multiple imputation by chained equations using the predictive mean matching method, followed by bootstrap resampling across imputations for subsequent computation of confidence intervals around the statistical metrics described below (27).

Statistical Analysis Methods

Performance of the RECODe equations was compared with that of the UKPDS OM2 (17) and the ACC/AHA PCEs (18) by comparing model calibration, discrimination, and reclassification. We do not compare between cohorts (i.e., with the UKPDS cohort) but between equations (i.e., with the UKPDS risk equations). The RECODe equations are Cox proportional hazards models, as are the AHA/ACC PCEs, whereas the UKPDS equations take various functional forms specific to each outcome (Gompertz, Weibull, logistic, or exponential). Model calibration was assessed by the slope and intercept of the line between predicted and observed Kaplan-Meier event rates for each outcome over 10 years by deciles of risk (Greenwood-Nam-D'Agostino [GND] test) (28). The calibration slope is ideally 1 and intercept ideally 0, reflecting a perfect correspondence between predicted and observed event rates of each outcome. The GND test nonparametrically evaluates the distance between predicted and observed Kaplan-Meier outcome rates, so that higher P values, indicating greater concordance (less difference) between predicted and observed outcome rates, are desirable. Model discrimination was assessed by the C-statistic (area under the receiver operating characteristic curve) (29). Following the TRIPOD guidelines (19), time to event for outcomes was taken from the time of study enrollment or the time of diabetes diagnosis (in the event of new diabetes diagnosed during the study), whichever came later, to mimic application of the equations in clinical settings in which patients may be previously diagnosed before application of the equations, or newly diagnosed at the time of application of the equations. Reclassification referred to the ability of the older models versus the newer models to distinguish high-risk from low-risk patients (10-year risk $\geq$10% vs. <10% for nephropathy, retinopathy, ASCVD, and CHF; and $\geq$5% vs. <5% for MI and stroke), the cut points often suggested for consideration of more intensive blood pressure, lipid, or glycemic therapies (5,30). The net reclassification index (NRI) was calculated, which is the number of people who had an event who are correctly reclassified as high risk by a new

Table 1—Characteristics of the MESA (2000–2012, n = 1,555 people with type 2 diabetes) and JHS (2000–2012, n = 1,746 people with type 2 diabetes) study participants included for validation of RECODe equations

	Included sample	
	MESA (n = 1,555)	JHS (n = 1,746)
Demographics		
Age, mean (SD), years	63.0 (9.7)	58.8 (11.0)
Women	772 (49.6)	1,136 (65.1)
Race/ethnicity		
White race	369 (23.7)	0 (0)
Black race	548 (35.2)	1,746 (100.0)
Asian race	187 (12.0)	0 (0)
Hispanic or Latino ethnic group	451 (29.0)	0 (0)
Clinical features		
Tobacco smoking, current	195 (12.5)	197 (11.3)
Systolic blood pressure, mean (SD), mmHg	131.7 (21.4)	130.3 (16.9)
CVD history	0 (0)	191 (15.1)
Medication utilization		
Blood pressure treatment	857 (55.1)	1,223 (71.9)
Oral diabetes medication (including metformin)	664 (42.7)	573 (33.7)
Statin use	351 (22.6)	349 (20.5)
Anticoagulant use	11 (0.7)	Not assessed
Biomarkers, mean (SD)		
Hemoglobin A_{1c}, % [mmol/mol]	6.8 (1.5) [51 (11)]	7.1 (1.7) [53 (13)]
Total cholesterol, mmol/L [mg/dL]	4.9 (1.0) [191.1 (38.4)]	5.0 (1.0) [194.5 (39.7)]
Direct HDL cholesterol, mmol/L [mg/dL]	1.2 (0.3) [46.3 (12.5)]	1.3 (0.3) [51.3 (13.5)]
Serum creatinine, μmol/L [mg/dL]	88.4 (35.4) [1.0 (0.4)]	88.4 (62.0) [1.0 (0.7)]
Urine albumin-to-creatinine, mg/mmol [mg/g]	7.6 (35.7) [66.9 (316.2)]	14.7 (58.2) [129.0 (515.6)]

Data are n (%) unless stated otherwise.

Table 2—Characteristics of people with type 2 diabetes in each cohort, along with incident microvascular and cardiovascular outcome definitions and timing

Outcome	MESA (n = 1,555)	JHS (n = 1,746)
Prevalent diabetes (diagnosed at baseline exam or prior), count	859	1,152
Incident diabetes (diagnosed during course of study follow-up), count	696	594
On oral diabetes agents, count	661	806
On insulin, count	120	Not specified
Time to diabetes diagnosis, median (range) in days among those diagnosed during course of study (after baseline exam)	1,653 (319, 4,078)	2,230 (1,331, 4,403)
Incident microalbuminuria, count (new microalbuminuria not present on baseline laboratories or first laboratories after diabetes diagnosis)	291	176
Time to incident microalbuminuria, median (range) in days	721 (395, 4,003)	1,847 (878, 5,982)
Incident macroalbuminuria, count (new macroalbuminuria not present on baseline laboratories or first laboratories after diabetes diagnosis)	92	77
Time to incident macroalbuminuria, median (range) in days	1,126 (414, 3,780)	2,386 (1,125, 4,964)
Incident renal failure/ESRD, count (new renal failure/ESRD not present on baseline laboratories or first laboratories after diabetes diagnosis)	13	27
Time to incident renal failure/ESRD, median (range) in days	1,808 (1,038, 3,787)	1,848 (1,700, 2,346)
Incident retinopathy, count (new retinopathy not present on exam 2 photos but present on exam 5 photos)	34	94
Time to incident retinopathy, median (range) in days	2,837 (2,450, 3,259)	2,817 (1,909, 4,083)
Incident MI (fatal or nonfatal), count (new MI among people without history of MI on baseline exam or first exam after diabetes diagnosis)	92	151
Time to incident MI, median (range) in days	3,645 (29, 5,275)	1,786 (9, 4,385)
Incident stroke (fatal or nonfatal), count (new stroke among people without history of stroke on baseline exam or first exam after diabetes diagnosis)	89	142
Time to incident stroke, median (range) in days	3,636 (7, 5,275)	1,542 (17, 4,039)
Incident CHF, count (new CHF among people without history of CHF on baseline exam or first exam after diabetes diagnosis)	117	161
Time to incident CHF, median (range) in days	3,598 (29, 5,275)	3,896 (35, 4,844)
Incident cardiovascular death, count	88	64
Time to incident cardiovascular death, median (range) in days	2,764 (29, 5,045)	1,806 (9, 4,444)
Incident all-cause mortality, count	323	795
Time to incident all-cause mortality, median (range) in days	4,224 (29, 5,275)	3,944 (77, 5,209)

Time to an incident event is defined as number of days from diabetes diagnosis or baseline examination, whichever comes later, to the point at which the incident event was recorded. Time to event was censored at 10 years to correspond to the 10-year risk prediction from the RECODe equations. Urine albumin and creatinine were measured by nephelometry and the rate Jaffe reaction, respectively (40,41). Serum creatinine was measured in MESA using rate reflectance spectrophotometry using thin-film adaptation of the creatinine amidinohydrolase method on a Vitros analyzer calibrated to the Cleveland Clinic (42), and in JHS using a multipoint enzymatic spectrophotometric assay on a Vitros analyzer (43) with calibration against a National Institute of Standards and Technology standard (44). Retinopathy was evaluated by dilated, bilateral, seven-standard field fundus photographs including macular stereoscopic pairs, scored contemporaneously by two independent, masked ophthalmologist investigators with the ETDRS (23), including notation of focal laser treatment scars. Differences in diagnosis were arbitrated by a third masked ophthalmologist investigator and/or by joint review by the ophthalmologist investigators. Participants with another ocular disease that precluded photograph grading were excluded (25,45). CHF in JHS was defined by 1) a discharge diagnosis of ICD-9 code 428 and/or underlying cause of death I50, 2) radiographic findings consistent with CHF or increased venous pressure or dilated ventricle/left ventricular function <40% by echo/multiple gated acquisition scan, or 3) autopsy finding of pulmonary edema/CHF (46). CHF in MESA was defined by 1) CHF diagnosed by a physician and patient receiving medical treatment for CHF, 2) pulmonary edema/congestion seen on a chest radiograph, and 3) dilated ventricle or poor left ventricular systolic function by echocardiography or ventriculography or evidence of left ventricular diastolic dysfunction by echocardiography (47).

model minus those incorrectly reclassified to low risk, plus the number who did not have an event who are correctly reclassified as low risk by a new model minus those incorrectly reclassified to high risk (i.e., a positive NRI indicates improvement of the new model versus the older model) (30).

Analyses were performed in R and were approved by the Stanford University Institutional Review Board (e-Protocol ID 39274).

RESULTS

Participants

The eligible participant sample from MESA (Table 1) averaged 63.0 years old, was 49.6% female, and was followed for a median of 9.1 years. Of eligible participants,

396 (25%) developed microalbuminuria, macroalbuminuria, or renal failure/ESRD; 34 (2%) developed retinopathy; 92 (6%) had an MI (fatal or nonfatal); 89 (6%) had a stroke (fatal or nonfatal); 117 (8%) developed CHF; 88 (6%) died of cardiovascular disease (CVD); and 323 (21%) died of all causes during the study.

The eligible participant sample from JHS (Table 1) averaged 57.5 years old, was 64.7% female, and was followed for a median of 8.0 years. Of the eligible participants, 280 (16%) developed microalbuminuria, macroalbuminuria, or renal failure/ESRD; 94 (5%) developed retinopathy; 151 (5%) had an MI (fatal or nonfatal); 142 (8%) had a stroke (fatal or nonfatal); 161 (9%) developed CHF; 64 (4%) died of CVD; and 795 (46%) died of all causes during the study. Further details about the disaggregated incident outcomes, timing of incident outcomes, and outcome adjudication in both MESA and JHS are provided in Table 2.

The eligible participant samples from MESA and JHS were of similar age and sex distribution but had greater race/ethnic diversity compared with the ACCORD participant sample (from whom the RECODe equations were derived) and the DPPOS and Look AHEAD samples (among whom the RECODe equations were initially validated) (Supplementary Table 3). The eligible participant sample from MESA and JHS also had fewer participants with a history of pharmaceutical therapy for CVD risk factors (blood pressure and/or statin therapy) as compared with the ACCORD, DPPOS, or Look AHEAD samples (31) (Supplementary Table 3).

Model Performance

In external validation against MESA data (Fig. 1 and Table 3), RECODe equations predicted retinopathy with a calibration slope of 1.07 between expected and observed event rates, a calibration intercept of −0.007, calibration GND test P value of 0.05, and a C-statistic of 0.76. RECODe equations for nephropathy performed best on the calibration and discrimination metrics for the outcomes of microalbuminuria, macroalbuminuria, or the composite outcome of microalbuminuria, macroalbuminuria, or renal failure/ESRD; for these outcomes, calibration slopes were between 0.99 and 1.04, intercepts between 0.004 and 0.032, GND P values 0.08–0.12, and C-statistics 0.76–0.85 (Fig. 1

and Table 3). For the composite outcome of either doubling of serum creatinine or >20 mL/min decrease in eGFR, the RECODe equations had a calibration slope of 0.83, intercept of 0.080, GND P value of 0.05, and C-statistic of 0.77. The RECODe equation performed more poorly for the composite outcome of macroalbuminuria, renal failure, doubling of creatinine, or >20 mL/min decrease in eGFR, which had a calibration slope of 0.62, intercept of 0.157, GND P value of <0.001, and C-statistic of 0.71. The MESA sample did not have an adequate number of patients with ESRD events to conduct calibration testing on ESRD as an isolated outcome ($n = 13$). RECODe equations predicted cardiovascular outcomes with calibration slopes of 1.00–1.05, calibration intercepts of −0.005 to 0.024, GND P values >0.05 for all outcomes except stroke, and C-statistics varying from 0.73 (for MI) to 0.81 (for CVD death) (Fig. 1 and Table 3). The RECODe equation for all-cause mortality applied to MESA had a calibration slope of 1.03, intercept of −0.009, GND P value of 0.08, and C-statistic of 0.81.

In external validation against JHS data (Fig. 1 and Table 3), the RECODe equation for retinopathy had a calibration slope of 0.47, intercept of 0.119, GND P value of 0.14, and C-statistic of 0.64 (Fig. 1 and Table 3). The RECODe equations for nephropathy performed best for the outcomes of macroalbuminuria and the composite outcome of microalbuminuria, macroalbuminuria, or renal failure/ESRD, with calibration slopes of 0.98–1.16, intercepts −0.029 to 0.008, GND P values 0.05–0.09, and C-statistics 0.71–0.77 (Fig. 1 and Table 3). For the other nephropathy outcomes, calibration slopes varied from 0.97 (for microalbuminuria) to 1.16 (for the composite of microalbuminuria, macroalbuminuria, or renal falure/ESRD), intercepts from −0.029 to 0.106, GND P values <0.001, and C-statistics from 0.62 (for the composite outcome of macroalbuminuria, renal failure/ESRD, doubling of creatinine, or >20 mL/min decrease in eGFR) to 0.91 for renal failure/ESRD (Fig. 1 and Table 3). The JHS sample did not have adequate ESRD events to conduct calibration testing on ESRD as an isolated outcome ($n = 27$). RECODe equations predicted cardiovascular outcomes with calibration slopes of 0.94–1.10; calibration intercepts of −0.023 to 0.091; GND P values >0.05

for MI, CHF, and CVD death; and C-statistics varying from 0.72 (for stroke) to 0.87 (for CVD death) (Fig. 1 and Table 3). The RECODe equation for all-cause mortality applied to JHS had a calibration slope of 1.01, intercept of 0.007, GND P value of 0.95, and C-statistic of 0.78 (Fig. 1 and Table 3).

Model performance did not vary substantially when evaluated among subgroups of only participants who were diagnosed with diabetes prior to or at the baseline study visit, or only participants who were diagnosed during the course of the study (Supplementary Fig. 1).

Comparison With Alternative Risk Equations

Compared with the RECODe equations, the UKPDS OM2 equations and the ACC/AHA PCEs had lower C-statistics and calibration slopes in both the MESA and JHS samples (Table 3). For the outcome of retinopathy, the UKPDS OM2 predicted outcomes less well than random chance, with calibration slopes of 0.17 and 0.60 in MESA and JHS, respectively, calibration intercepts of 0.042 and 0.180, GND P values 0.17 and <0.001, and C-statistics of 0.50 and 0.52 (Table 3). For the outcome of ESRD, the UKPDS OM2 had a C-statistic of 0.72 in MESA and 0.56 in JHS; calibration could not be tested due to insufficient events. The UKPDS OM2 lacked equations for the remaining nephropathy outcomes. For cardiovascular outcomes, the UKPDS OM2 equations had calibration slopes ranging from 0.07 to 0.27, intercepts from 0.003 to 0.163, GND P values all <0.001, and C-statistics ranging from 0.54 to 0.61 across the two data sets. For the cardiovascular outcome of ASCVD, the ACC/AHA PCEs had calibration slopes of 0.19 and 0.42, intercepts of 0.019 and 0.0002, GND P values <0.001, and C-statistics of 0.60 and 0.63 in MESA and JHS, respectively (Table 3).

The reclassification of high- versus low-risk people by the RECODe equations as compared with the older sets of equations is shown in Supplementary Table 4 for each outcome. RECODe equations were more likely than the UKPDS OM2 to correctly distinguish high-risk from low-risk patients for each of the microvascular and cardiovascular outcomes (10-year risk ≥10% vs. <10% for nephropathy [NRI = 0.01 in MESA and 0.72 in JHS], retinopathy [NRI = 0.02 in MESA and
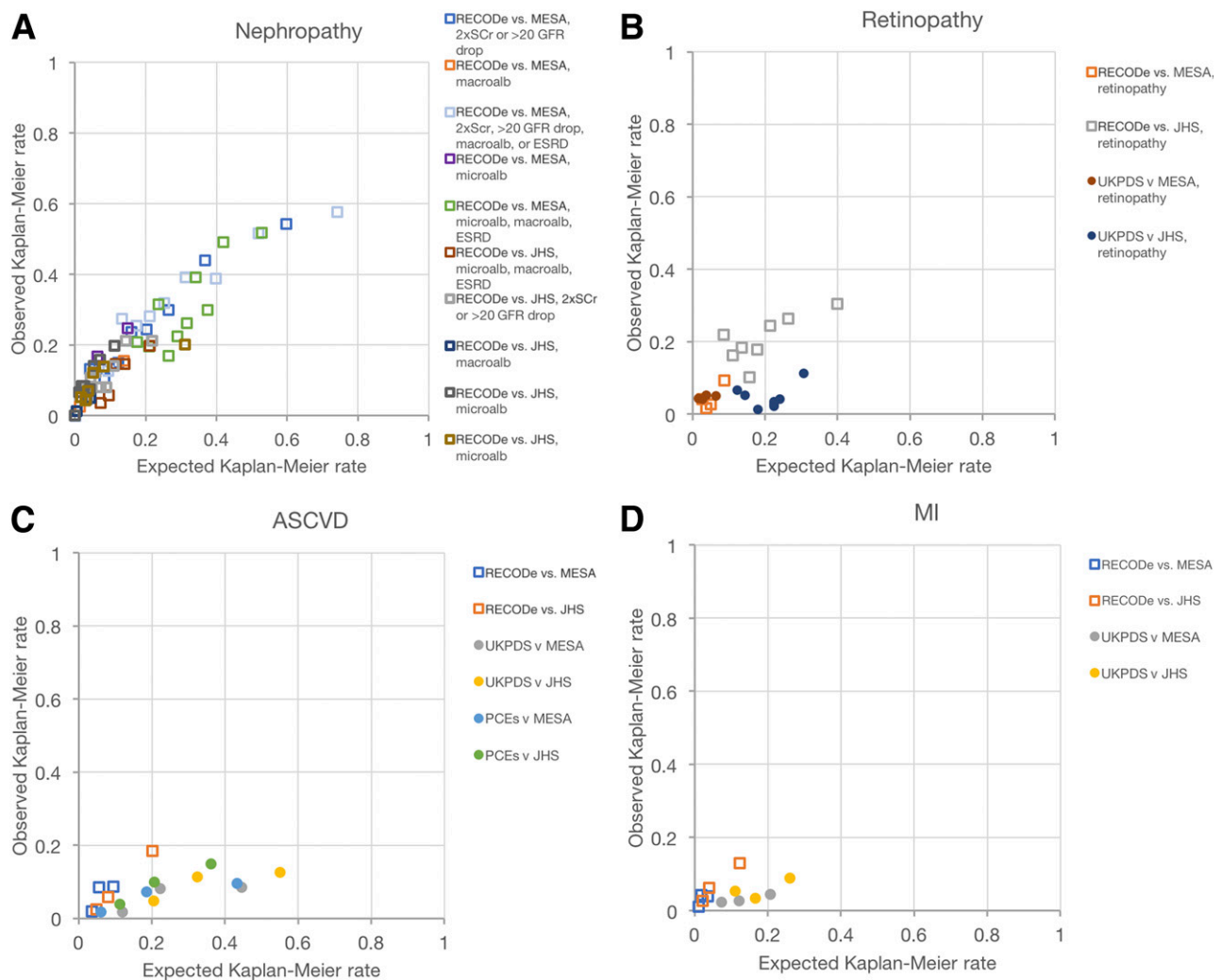
Figure 1—Calibration plots. Plots display Kaplan-Meier event rates over 10 years predicted by the RECODe equations vs. observed rates among MESA (2000–2012, *n* = 1,555 people with type 2 diabetes) and JHS (2000–2012, *n* = 1,746 people with type 2 diabetes). Also displayed, where available for each outcome, are predictions from UKPDS OM2. Each set of equations was applied to each data set to provide the comparisons shown here. Points are displayed for deciles of predicted and observed Kaplan-Meier event rates, with lower numbers of centiles than deciles used if less than five events are observed per group, to prevent unstable inferences per current guidelines. *A*: Nephropathy. *B*: Retinopathy. *C*: ASCVD. *D*: MI. *E*: Stroke. *F*: CHF. *G*: Mortality. 2xScr, doubling of serum creatinine; macroalb, macroalbuminuria; microalb, microalbuminuria.

0.08 in JHS], ASCVD [NRI = 1.87 in MESA and 9.00 in JHS], and CHF [0.17 in MESA and 0.19 in JHS]; and ≥5% vs. <5% for MI [NRI = 8.09 in MESA and 26.00 in JHS] and stroke [0.51 in MESA and 6.14 in JHS]). RECODe equations were also more likely than the ACC/AHA PCEs to correctly distinguish high-risk from low-risk patients for ASCVD (10-year risk ≥10% vs. <10% for ASCVD; NRI = 0.09 in MESA and 0.19 in JHS).

CONCLUSIONS

In this work, we validated the RECODe equations for predicting microvascular and macrovascular outcomes in type 2 diabetes among sociodemographically diverse, community-based participants in

two different cohort studies. We have previously published a validation of the RECODe equation in two randomized trials, the DPPOS and Look AHEAD studies, comprising mainly non-Hispanic white subjects with type 2 diabetes (11). The RECODe equations improved upon prior risk equations for the prediction of key outcomes of interest among people with type 2 diabetes. The RECODe equations outperformed the UKPDS OM2 and ACC/AHA PCE risk equations based on metrics for discrimination and calibration, whereas the latter two risk equation sets exhibited substantial overprediction of both macrovascular and microvascular events in the studied cohorts. The RECODe equations nevertheless did not

perform ideally; for the outcome of retinopathy in particular, the equations still overpredicted risk. The RECODe equations also performed well for most nephropathy outcomes, but not in MESA for the composite ACCORD outcome of macroalbuminuria or renal failure/ ESRD or doubling of serum creatinine or >20 mL/min decrease in eGFR. The RECODe equations were validated here against data available to other researchers, were derived through open-source statistical code, and have an online risk calculator (32). Unlike prior studies of the UKPDS OM2, our external validation used individual patient data rather than only aggregate study outcomes (8).
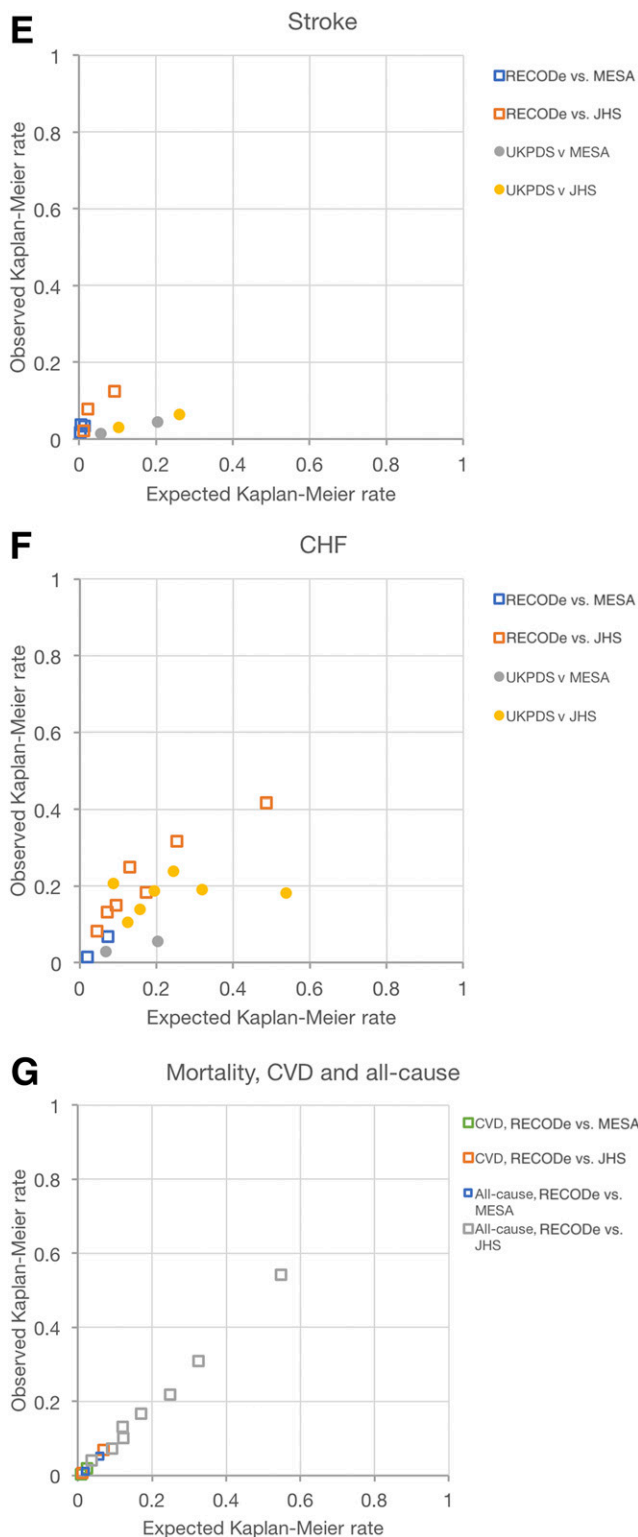
E

Stroke

Legend:
- RECODe vs. MESA
- RECODe vs. JHS
- UKPDS v MESA
- UKPDS v JHS

F

CHF

Legend:
- RECODe vs. MESA
- RECODe vs. JHS
- UKPDS v MESA
- UKPDS v JHS

G

Mortality, CVD and all-cause

Legend:
- CVD, RECODe vs. MESA
- CVD, RECODe vs. JHS
- All-cause, RECODe vs. MESA
- All-cause, RECODe vs. JHS

Figure 1—*Continued*.

The RECODe equations greatly mitigated problems with older equations. The UKPDS equations typically overestimated risk by four- to sevenfold in this assessment and reflected the health and racial diversity of the U.K., not the U.S.

The ACC/AHA equations did not include microvascular complications, omitted important covariates not widely available in their development data sets (e.g., statin treatment), and substantially overestimated risk, consistent with prior reports

(9,33,34). This may be due to the younger age of diagnosis and earlier diagnosis (i.e., diagnosis at a lower baseline hemoglobin A_{1c}) among modern U.S. populations than the original UKPDS cohort. In particular, we found that this study verifies that the RECODe equations are more accurate for culturally diverse, American patients with diabetes, including those both previously diagnosed and newly diagnosed. Risk equations should be valid for clinical assessments immediately upon diagnosis and when physicians care for a patient who was previously diagnosed but with new therapeutic considerations; the latter patients will have different input parameters and risk estimates from the equations.

The equations have direct clinical use, as the American Diabetes Association and other groups already recommend that clinicians consider overall risk of diabetes complications and "shared decision making" when deciding upon therapy. Shared decision making necessarily involves discussing the benefits and risks of a therapy with the patient (4,35). As some populations in the U.S. appear to be at particularly high risk of overtreatment, leading to serious adverse events (36,37), having more accurate risk estimates could help avert drug-related morbidity and mortality among populations for whom older equations overestimated risk; by contrast, more aggressive but potentially higher-risk treatments for blood pressure, lipids, and glucose can be concentrated among those patients for whom the risk of complications and potential benefits of more aggressive therapy might be more favorable (38). The equations can also inform comparative effectiveness and cost-effectiveness research for population health interventions for which accurate risk models for modern populations are necessary.

The RECODe equations had good discrimination and were well calibrated as compared with alternatives, indicating that both higher- and lower-risk subjects would have more accurate risk estimates when using RECODe.

Limitations

The available studies lacked neuropathy as an outcome variable. Relatedly, due to differences in study end point definitions, we could only validate a subset of possible microvascular end point definitions and notably did not have a sufficient

Table 3—Validation statistics for microvascular and macrovascular RECODe equations

Equation	External validation: MESA		External validation: JHS		Alternative risk equations: UKPDS OM2 in MESA		Alternative risk equations: UKPDS OM2 in JHS	
	Discrimination: C-statistic	Calibration: slope/intercept/ χ^2, P value	Discrimination: C-statistic	Calibration: slope/intercept/ χ^2, P value	Discrimination: C-statistic	Calibration: slope/intercept/ χ^2, P value	Discrimination: C-statistic	Calibration: slope/intercept/ χ^2, P value
Microvascular								
Retinopathy								
1) Retinopathy	0.76	1.07/−0.007/7.7, 0.05	0.64	0.47/0.119/10.9, 0.14	0.50	0.17/0.042/5.0, 0.17	0.52	0.60/0.180/65.6, <0.001
Nephropathy								
1) Microalbuminuria	0.85	1.04/−0.032/10.8, 0.10	0.64	0.97/0.106/26.3, <0.001				
2) Macroalbuminuria	0.84	1.01/0.018/4.2, 0.12	0.77	0.98/0.008/2.9, 0.09				
3) Renal failure/ESRD	0.78	—*	0.91	—*	0.72	—*	0.56	—*
4) Doubling of serum creatinine or >20 mL/min decrease in eGFR	0.77	0.83/0.080/13.9, 0.05	0.68	1.02/0.006/42.1, <0.001				
5) Any of 2–4	0.71	0.62/0.157/36.5, <0.001	0.62	0.98/0.007/55.4, <0.001				
6) Any of 1–3	0.76	0.99/−0.004/15.4, 0.08	0.71	1.16/−0.029/9.6, 0.05				
Cardiovascular								
1) ASCVD	0.74	1.00/0.004/6.2, 0.005	0.77	1.04/−0.023/25.5, <0.001	0.60 (UKPDS OM2); 0.60 (ACC/AHA)	0.17/0.016/725.5, <0.001 (UKPDS OM2); 0.19/0.019/442.4, <0.001 (ACC/AHA)	0.61 (UKPDS OM2); 0.63 (ACC/AHA)	0.20/0.024/901.9, <0.001 (UKPDS OM2); 0.42/0.0002/214.8, <0.001 (ACC/AHA)
2) MI (fatal/nonfatal)	0.73	1.00/0.010/4.4, 0.11	0.74	0.95/0.016/4.3, 0.11	0.54	0.17/0.009/295.0, <0.001	0.57	0.27/0.010/325.4, <0.001
3) Stroke (fatal/nonfatal)	0.75	1.00/0.024/17.3, <0.001	0.72	1.05/0.032/22.9, <0.001	0.60	0.21/0.003/287.8, <0.001	0.60	0.22/0.008/42.0, <0.001
4) CHF	0.80	1.01/−0.005/6, 0.42	0.73	0.72/0.091/11.7, 0.07	0.57	0.19/0.017/189.7, <0.001	0.54	0.07/0.163/154.3, <0.001
5) CVD death	0.81	1.05/−0.003/0.7, 0.40	0.87	1.10/−0.006/1.4, 0.24				
All-cause mortality	0.81	1.03/−0.009/3.0, 0.08	0.78	1.01/0.007/2.15, 0.95				

See RESEARCH DESIGN AND METHODS for definitions of outcomes. Each set of equations was applied to each data set to provide the comparisons shown here. Calibration slopes/intercepts are calculated between deciles of predicted and observed Kaplan-Meier event rates, with lower numbers of centiles than deciles used if less than five events are observed per group, to prevent unstable inferences per current guidelines. P values <0.05 reflect larger difference between predicted and observed Kaplan-Meier event rates by the GND test (see Fig. 1 for calibration plots). *Insufficient events to perform GND test (two or less events per risk group).

sample to validate against the important outcome of ESRD. Retinopathy, in particular, occurred in a limited set of patients, and in the JHS, only one retinal examination was available, such that baseline prior retinopathy definition was imprecise due to self-report and in turn incident retinopathy was imprecise as well. Further study should evaluate the utility of the equations for retinopathy assessment among blacks. Finally, longer periods of follow-up may enable the extension of equations to the full life-course, enabling linkage to long-term life expectancy outcomes for which we lack sufficient statistical power at present.

A key set of limitations to this validation concerns the populations against which we validated. Because the RECODe equations were validated previously in predominantly white populations (n = 11,036 white, 2,854 black, and 1,523 Hispanic) (11), we validated them here in predominantly minority populations (1,746 black in JHS; 369 white, 548 black, 187 Asian, and 451 Hispanic in MESA). Hence, our findings do not generalize to that of the broad U.S. population, although there are disproportionately black adults represented in the U.S. population of patients with type 2 diabetes. All risk equation evaluations should consider the social and historical conditions against which the validation populations are assessed; in our case, the JHS black population in particular is from Mississippi and may face high competing risks of other disease, and which differs markedly from the social and historical context of the UKPDS cohort. Hence, we used calibration and discrimination tests that account for censorship due to other conditions (28,39) and validated both CVD and all-cause mortality equations. Nevertheless, clinical practitioners should use risk equations that validated in populations most similar to their practice population. We also highlight that readers should not wrongly interpret that risk is lower for all populations when using RECODe than older risk equations. We found, instead, that although older equations may produce inflated risk estimates on average, many high-risk people (18–100% of those with an event in JHS and 3–97% in MESA) were wrongly predicted as low risk by the older equations and were caught by the new equations. Hence, our goal is accurate risk estimation, not underestimation. We also

suggest that in further work, the equations should be recalibrated if necessary to local populations that are being treated, once electronic health record registries have sufficiently accurate adjudication to perform more localized inferences.

References
1. Vijan S, Sussman JB, Yudkin JS, Hayward RA. Effect of patients' risks and preferences on health gains with plasma glucose level lowering in type 2 diabetes mellitus. JAMA Intern Med 2014;174: 1227–1234
2. Raz I, Riddle MC, Rosenstock J, et al. Personalized management of hyperglycemia in type 2 diabetes: reflections from a Diabetes Care Editors' Expert Forum. Diabetes Care 2013;36:1779–1788
3. Lipska KJ, Krumholz HM. Is hemoglobin A1c the right outcome for studies of diabetes? JAMA 2017;317:1017–1018
4. Rodriguez-Gutierrez R, Gionfriddo MR, Ospina NS, et al. Shared decision making in endocrinology: present and future directions. Lancet Diabetes Endocrinol 2016;4:706–716
5. Basu S, Shankar V, Yudkin JS. Comparative effectiveness and cost-effectiveness of treat-to-target versus benefit-based tailored treatment of type 2 diabetes in low-income and middle-income countries: a modelling analysis. Lancet Diabetes Endocrinol 2016;4:922–932
6. Basu S, Vellakkal S, Agrawal S, Stuckler D, Popkin B, Ebrahim S. Averting obesity and type 2 diabetes in India through sugar-sweetened beverage taxation: an economic-epidemiologic modeling study. PLoS Med 2014;11:e1001582
7. Basu S, Sussman JB, Hayward RA. Detecting heterogeneous treatment effects to guide personalized blood pressure treatment: a modeling study of randomized clinical trials. Ann Intern Med 2017;166:354–360
8. Palmer AJ, Clarke P, Gray A, et al.; Mount Hood 5 Modeling Group. Computer modeling of diabetes and its complications: a report on the Fifth Mount Hood challenge meeting. Value Health 2013;16:670–685
9. Cook NR, Ridker PM. Calibration of the pooled cohort equations for atherosclerotic cardiovascular disease: an update. Ann Intern Med 2016;165: 786–794
10. Tao L, Wilson ECF, Griffin SJ, Simmons RK; ADDITION-Europe study team. Performance of the UKPDS outcomes model for prediction of myocardial infarction and stroke in the ADDITION-Europe trial cohort. Value Health 2013;16:1074–1080
11. Basu S, Sussman J, Berkowitz S, Hayward R, Yudkin J. Development and validation of Risk Equations for Complications of Type 2 Diabetes (RECODe) using individual participant data from randomized trials. Lancet Diabetes Endocrinol 2017;5:788–798
12. Gerstein HC, Miller ME, Byington RP, et al.; Action to Control Cardiovascular Risk in Diabetes Study Group. Effects of intensive glucose lowering in type 2 diabetes. N Engl J Med 2008;358:2545–2559
13. Ginsberg HN, Elam MB, Lovato LC, et al.; ACCORD Study Group. Effects of combination lipid therapy in type 2 diabetes mellitus. N Engl J Med 2010;362:1563–1574
14. Cushman WC, Evans GW, Byington RP, et al.; ACCORD Study Group. Effects of intensive blood-pressure control in type 2 diabetes mellitus. N Engl J Med 2010;362:1575–1585
15. Diabetes Prevention Program Research Group. Long-term effects of lifestyle intervention or metformin on diabetes development and microvascular complications over 15-year follow-up: the Diabetes Prevention Program Outcomes

Study. Lancet Diabetes Endocrinol 2015;3:866–875

16. Wing RR, Bolin P, Brancati FL, et al.; Look AHEAD Research Group. Cardiovascular effects of intensive lifestyle intervention in type 2 diabetes. N Engl J Med 2013;369:145–154

17. Hayes AJ, Leal J, Gray AM, Holman RR, Clarke PM. UKPDS outcomes model 2: a new version of a model to simulate lifetime health outcomes of patients with type 2 diabetes mellitus using data from the 30 year United Kingdom Prospective Diabetes Study: UKPDS 82. Diabetologia 2013;56:1925–1933

18. Goff DC Jr, Lloyd-Jones DM, Bennett G, et al.; American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation 2014;129(Suppl. 2):S49–S73

19. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. Ann Intern Med 2015;162:55–63

20. Bild DE, Bluemke DA, Burke GL, et al. Multi-Ethnic Study of Atherosclerosis: objectives and design. Am J Epidemiol 2002;156:871–881

21. Taylor HA, Wilson JG, Jones DW, et al. Toward resolution of cardiovascular health disparities in African Americans: design and methods of the Jackson Heart Study. Ethn Dis 2005;15(Suppl. 6):4–17

22. Levey AS, Stevens LA, Schmid CH, et al.; CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration). A new equation to estimate glomerular filtration rate. Ann Intern Med 2009;150:604–612

23. Diabetic retinopathy study. Report Number 6. Design, methods, and baseline results. Report Number 7. A modification of the Airlie House classification of diabetic retinopathy. Prepared by the Diabetic Retinopathy. Invest Ophthalmol Vis Sci 1981;21:1–226

24. Sobrin L, Penman A, Hoadley S, Wilson J, Taylor H, Chen C. Epidemiology of retinopathy in African Americans with impaired fasting glucose and type 2 diabetes in the Jackson Heart Study. Invest Ophthalmol Vis Sci 2013;54:2171

25. Kawasaki R, Cheung N, Islam FMA, et al.; Multi-Ethnic Study of Atherosclerosis. Is diabetic retinopathy related to subclinical cardiovascular disease? Ophthalmology 2011;118:860–865

26. American Diabetes Association. Classification and diagnosis of diabetes. Sec. 2. In *Standards of Medical Care in Diabetes—2017*. Diabetes Care 2017;40(Suppl. 1):S11–S24

27. van Buuren S, Groothuis-Oudshoorn, K. mice: multivariate imputation by chained equations in R. Journal of Statistical Software 2011;45:1–67

28. Demler OV, Paynter NP, Cook NR. Tests of calibration and goodness-of-fit in the survival setting. Stat Med 2015;34:1659–1680

29. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics 2000;56:337–344

30. Kerr KF, Wang Z, Janes H, McClelland RL, Psaty BM, Pepe MS. Net reclassification indices for evaluating risk prediction instruments: a critical review. Epidemiology 2014;25:114–121

31. Basu S, Sussman JB, Berkowitz SA, Hayward RA, Yudkin JS. Development and validation of Risk Equations for Complications of Type 2 Diabetes (RECODe) using individual participant data from randomised trials. Lancet Diabetes Endocrinol 2017;5:788–798

32. Basu S. Risk Equations for Complications of Type 2 Diabetes: RECODe [Internet]. Available from https://sanjaybasu.shinyapps.io/recode/. Accessed 24 March 2017

33. Cook NR, Ridker PM. Response to comment on the reports of over-estimation of ASCVD risk using the 2013 AHA/ACC risk equation. Circulation 2014;129:268–269

34. Ridker PM, Cook NR. Statins: new American guidelines for prevention of cardiovascular disease. Lancet 2013;382:1762–1765

35. American Diabetes Association. Pharmacologic approaches to glycemic treatment. Sec. 8. In *Standards of Medical Care in Diabetes—2017*. Diabetes Care 2017;40(Suppl. 1):S64–S74

36. Lipska KJ, Ross JS, Miao Y, Shah ND, Lee SJ, Steinman MA. Potential overtreatment of diabetes mellitus in older adults with tight glycemic control. JAMA Intern Med 2015;175:356–362

37. Maciejewski ML, Mi X, Sussman J, et al. Overtreatment and deintensification of diabetic

therapy among Medicare beneficiaries. J Gen Intern Med. 13 September 2017 [Epub ahead of print]. DOI:10.1007/s11606-017-4167-y

38. Basu S, Sussman JB, Rigdon J, Steimle L, Denton BT, Hayward RA. Benefit and harm of intensive blood pressure treatment: derivation and validation of risk models using data from the SPRINT and ACCORD trials. PLoS Med 2017;14:e1002410

39. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15:361–387

40. Kramer H, Jacobs DR, Bild D, et al. Urine albumin excretion and subclinical cardiovascular disease. The Multi-Ethnic Study of Atherosclerosis. Hypertension 2005;46:38–43

41. Jackson Heart Study Coordinating Center. *Jackson Heart Study Protocol. Manual 9. Central Laboratory and Specimen Repository. Specimen Collection and Processing*. Jackson, MS, University of Mississippi, 2001

42. Driver TH, Shlipak MG, Katz R, et al. Low serum bicarbonate and kidney function decline: the Multi-Ethnic Study of Atherosclerosis (MESA). Am J Kidney Dis 2014;64:534–541

43. Carpenter MA, Crow R, Steffes M, et al. Laboratory, reading center, and coordinating center data management methods in the Jackson Heart Study. Am J Med Sci 2004;328:131–144

44. Wang W, Young BA, Fülöp T, et al. Effects of serum creatinine calibration on estimated renal function in African Americans: the Jackson Heart Study. Am J Med Sci 2015;349:379–384

45. Penman A, Hoadley S, Wilson JG, Taylor HA, Chen CJ, Sobrin L. P-selectin plasma levels and genetic variant associated with diabetic retinopathy in African Americans. Am J Ophthalmol 2015;159:1152.e2–1160.e2

46. Keku E, Rosamond W, Taylor HA, et al. Cardiovascular disease event classification in the Jackson Heart Study: methods and procedures. Ethn Dis 2005;15(4 Suppl. 6):62–70

47. Bahrami H, Kronmal R, Bluemke DA, et al. Differences in the incidence of congestive heart failure by ethnicity: the Multi-Ethnic Study of Atherosclerosis. Arch Intern Med 2008;168:2138–2145