



HHS Public Access

Author manuscript

Lancet Diabetes Endocrinol. Author manuscript; available in PMC 2018 January 16.

Published in final edited form as:

Lancet Diabetes Endocrinol. 2017 October ; 5(10): 788–798. doi:10.1016/S2213-8587(17)30221-8.

Development and validation of Risk Equations for Complications Of type 2 Diabetes (RECODE) using individual participant data from randomised trials

Sanjay Basu, PhD,

Center for Population Health Sciences, Center for Primary Care and Outcomes Research, and Departments of Medicine and of Health Research and Policy, Stanford University, Palo Alto, CA, USA

Center for Primary Care, Harvard Medical School, Boston, MA, USA

Jeremy B Sussman, MD,

Division of General Medicine, University of Michigan, and Center for Clinical Management Research, Veterans Affairs Ann Arbor Healthcare, Ann Arbor, MI, USA

Seth A Berkowitz, MD,

Department of Medicine, Harvard Medical School, Boston, MA, USA

Division of General Internal Medicine and Diabetes Unit, Massachusetts General Hospital, Boston, MA, USA

Rodney A Hayward, MD, and

Division of General Medicine, University of Michigan, and Center for Clinical Management Research, Veterans Affairs Ann Arbor Healthcare, Ann Arbor, MI, USA

John S Yudkin, MD

Institute of Cardiovascular Science, Division of Medicine, University College London, London, UK

Summary

Background—In view of substantial mis-estimation of risks of diabetes complications using existing equations, we sought to develop updated Risk Equations for Complications Of type 2 Diabetes (RECODE).

Methods—To develop and validate these risk equations, we used data from the Action to Control Cardiovascular Risk in Diabetes study (ACCORD, n=9635; 2001–09) and validated the equations for microvascular events using data from the Diabetes Prevention Program Outcomes Study (DPPOS, n=1018; 1996–2001), and for cardiovascular events using data from the Action for

Correspondence to: Dr Sanjay Basu, Stanford University, Palo Alto, CA 94304, USA, basus@stanford.edu.

See Online for appendix

Contributors

SB and JSY conceived the study. SB did the primary analysis and drafted the manuscript. All authors interpreted the results and edited the manuscript.

Declaration of interests

We declare no competing interests.

Health in Diabetes (Look AHEAD, n=4760; 2001–12). Microvascular outcomes were nephropathy, retinopathy, and neuropathy. Cardiovascular outcomes were myocardial infarction, stroke, congestive heart failure, and cardiovascular mortality. We also included all-cause mortality as an outcome. We used a cross-validating machine learning method to select predictor variables from demographic characteristics, clinical variables, comorbidities, medications, and biomarkers into Cox proportional hazards models for each outcome. The new equations were compared to older risk equations by assessing model discrimination, calibration, and the net reclassification index.

Findings—All equations had moderate internal and external discrimination (C-statistics 0.55–0.84 internally, 0.57–0.79 externally) and high internal and external calibration (slopes 0.71–1.31 between observed and estimated risk). Our equations had better discrimination and calibration than the UK Prospective Diabetes Study Outcomes Model 2 (for microvascular and cardiovascular outcomes, C-statistics 0.54–0.62, slopes 0.06–1.12) and the American College of Cardiology/American Heart Association Pooled Cohort Equations (for fatal or non-fatal myocardial infarction or stroke, C-statistics 0.61–0.66, slopes 0.30–0.39).

Interpretation—RECODE might improve estimation of risk of complications for patients with type 2 diabetes.

Funding—National Institute for Diabetes and Digestive and Kidney Disease, National Heart, Lung and Blood Institute, and National Institute on Minority Health and Health Disparities, National Institutes of Health, and US Department of Veterans Affairs.

Introduction

Because of the risks and benefits of treatment, the care of patients with type 2 diabetes involves complex decision making.¹ In clinical practice, guidelines for risk factor reduction are now often based on total risk rather than on whether or not a patient has a single biomarker measurement higher or lower than a threshold value.² Hence, the accuracy of type 2 diabetes risk equations is central to clinical care of high-risk patients. Accurate equations are also needed for development and assessment of practice guidelines, and for comparative effectiveness research, which typically uses such equations for simulation modelling and to assess cost-effectiveness.^{3–5}

Although the falling rate of cardiovascular disease in high-income countries has been accounted for in some cardiovascular risk equations,^{6,7} this is not the case for many models of diabetes outcomes. As a result, risk of microvascular and cardiovascular complications of diabetes could be systematically misestimated in many populations.⁸ The policy of open data, encouraged by the US National Institutes of Health, offers an opportunity to develop updated risk equations. The release of individual participant data from large diabetes intervention trials enables development of updated equations for type 2 diabetes complications and mortality.

We sought to develop updated risk equations for microvascular and cardiovascular complications of type 2 diabetes using individual participant data from one large intervention study, to validate these risk equations using individual participant data from two

other large studies, and to compare the newly developed risk equations with older, widely used equations.^{7,9}

Methods

Data for development of risk equations

We derived risk equations from individual participant data from the Action to Control Cardiovascular Risk in Diabetes study (ACCORD; 2001–09), which had both microvascular and cardiovascular outcomes.^{10–12} In the ACCORD study (appendix), participants were aged 40–79 years with type 2 diabetes and had an HbA_{1c} of at least 7.5% (57 mmol/mol), and either history of cardiovascular disease or risk factors for cardiovascular disease (dyslipidaemia, hypertension, smoking, or obesity; table 1). Data from individual participants, across all study groups, were included for equation development (with dummy variables for study group to control for randomised treatment selection).¹³

Data for validation of risk equations

We validated the equations using individual participant data from the Diabetes Prevention Program Outcomes Study (DPPOS [1996–2001], for which only microvascular event data are available)¹⁴ and Action for Health in Diabetes (Look AHEAD [2001–12], for which only cardiovascular event data are available).¹⁵ We used data from people who developed type 2 diabetes before or during the DPPOS. DPPOS, used for validation of the microvascular risk equations, was the follow-up to the randomised, controlled Diabetes Prevention Program (DPP) trial of metformin, troglitazone, lifestyle, or placebo (with double-blinding of medication groups). The Look AHEAD trial, used for validation of the cardiovascular risk equations, was a randomised, controlled trial of intensive lifestyle modification versus diabetes support and education. Detailed characteristics of the patients are provided in the appendix.

Outcomes

We developed separate equations for each of several microvascular and cardiovascular outcomes, with unique equations for each alternative definition of each outcome. Microvascular outcomes were nephropathy, retinopathy, and neuropathy. Separate equations were developed for each ACCORD definition of nephropathy: development of microalbuminuria (albumin:creatinine ratio of 30 mg albumin per gram creatinine or greater in a random urine sample); development of macroalbuminuria (albumin:creatinine ratio of 300 mg albumin per gram or greater creatinine in random urine sample); renal failure, end-stage renal disease (dialysis), or serum creatinine greater than 3.3 mg/dL; doubling of serum creatinine or greater than 20 mL/min per 1.73 m² decrease in estimated glomerular filtration rate (based on the Modification of Diet in Renal Disease [MDRD] Study equation);¹⁶ a composite of any of macroalbuminuria, renal failure, end-stage renal disease, serum creatinine greater than 3.3 mg/dL, doubling of serum creatinine, or greater than 20 mL/min per 1.73 m² decrease in estimated glomerular filtration rate; and a composite of any of microalbuminuria, macroalbuminuria, renal failure, end-stage renal disease, or serum creatinine greater than 3.3 mg/dL. Separate equations were developed for each ACCORD definition of retinopathy: retinopathy requiring photocoagulation or vitrectomy; cataract

extraction; three-line reduction in visual acuity; severe vision loss (<20/200 visual acuity by Snellen chart); and a composite of photocoagulation, vitrectomy, or severe vision loss. Separate equations were also developed for each ACCORD definition of neuropathy: Michigan Neuropathy Screening Instrument score greater than 2.0;¹⁷ vibratory sensation loss; ankle jerk loss; and pressure sensation loss. Our use of several endpoints helped to capture people with early or intermediate outcomes rather than only end-stage microvascular complications. Cardiovascular outcomes were a composite of atherosclerotic cardiovascular disease (defined as first fatal or non-fatal myocardial infarction or stroke); fatal or non-fatal myocardial infarction; fatal or non-fatal stroke; congestive heart failure; or death from any cardiovascular cause. We modelled all-cause mortality as an additional outcome.

Candidate predictor variables assessed for inclusion in equations

Candidate predictor variables for microvascular and cardiovascular outcomes were taken from pre-randomisation eligibility screening or clinical examination data in ACCORD. Candidate predictors are listed in table 1, and included dummy variables for each study group, which controls for whether or not the participant was on intensive or standard glycaemic therapy, blood pressure-lowering treatment or lipid-lowering treatment. Squared and interaction terms were considered by including such terms for each predictor variable in the variable selection process detailed below.

Development of models

We developed Cox proportional hazards models for each outcome and used elastic net regularisation to select predictor variables. Elastic net regularisation is a machine learning approach designed to select models in the context of collinearity, which often leads to unstable estimates from traditional stepwise selection approaches;^{18,19} this approach fits a Cox model via penalised maximum likelihood, using 10-times internal cross-validation to minimise the risk of overfitting. We did not include time-varying covariates because the risk equations are intended for use in clinical settings to assist initial treatment decisions. Complete case analyses were done in the base case, without imputation, as only 616 (6%) of 9635 participants had missing values for any predictor variable. In sensitivity analyses we compared the results including people with missing values with those including people with missing values using multiple imputation with chained equations to impute missing covariates.²⁰

Assessment of internal and external model performance

We assessed model discrimination with the C-statistic (area under the receiver operating characteristic curves for time-to-event data, using an estimator that guarantees monotonicity based on a nearest neighbor estimator for the bivariate distribution function of risk score and survival time,²¹ and confidence intervals estimated using influence curves²²) using 10-times cross-validation. We assessed model calibration through the slope and intercept of the line between predicted and observed probabilities of each outcome by deciles of risk (Greenwood-D'Agostino-Nam [GND] test, a Cox model analogue to the Hosmer-Lemeshow test).²³ The GND test non-parametrically assesses the distance between predicted and observed Kaplan-Meier outcome rates, so that higher p values, indicating greater concordance between predicted and observed outcome rates, are desirable. Owing to

differences in study endpoint definitions, the microvascular outcomes that were possible to validate externally (ie, had a matching microvascular outcome definition in ACCORD and DPPOS) were the composite nephropathy outcome of development of microalbuminuria, macroalbuminuria, renal failure, or end-stage renal disease; the retinopathy outcome of retinopathy requiring photocoagulation or vitrectomy; and the neuropathy outcome of pressure sensation loss. All cardiovascular outcomes could be validated externally with Look AHEAD data.

We compared predictions from our model with those from the UK Prospective Diabetes Study Outcomes Model 2 (UKPDS OM2) for microvascular and macrovascular outcomes⁹ and the American College of Cardiology/American Heart Association (ACC/AHA) Pooled Cohort Equations (PCEs), which are an update of the Framingham risk equations for the composite cardiovascular outcome of first non-fatal or fatal myocardial infarction or stroke.⁷ The UKPDS OM2 had microvascular outcome definitions compatible with the ACCORD dataset (but not the DPPOS dataset), but only for nephropathy and retinopathy; cardiovascular outcome definitions were compatible with both ACCORD and Look AHEAD datasets for atherosclerotic cardiovascular disease, myocardial infarction, stroke, and congestive heart failure. RECODE were compared with UKPDS and ACC/AHA equations using the net reclassification index (NRI) to assess how well each set of equations distinguished high-risk from low-risk patients (10-year risk 10% vs <10% for nephropathy, retinopathy, atherosclerotic cardiovascular disease, and congestive heart failure; and 5% vs <5% for myocardial infarction and stroke). The NRI is calculated as the proportion of individuals with the outcome who were classified as low risk for the outcome using the UKPDS or ACC/AHA equations but classified as high risk by RECODE minus the proportion of individuals who did not have the outcome who were classified as high risk by the UKPDS or ACC/AHA equations but classified as low risk by RECODE (ie, high positive values would indicate improvement with RECODE in comparison with UKPDS and ACC/AHA equations). We note that the NRI is only one measure of model performance, and has been debated in terms of its value for decision making, although it is commonly used.²⁴ Analyses were done using *R* (version 3.2.3; *R* Foundation for Statistical Computing, Vienna) and was approved before study start by the Stanford University Institutional Review Board, e-Protocol ID#: 39274.

Role of the funding source

The sponsors had no role in the study design; collection, analysis, or interpretation of the data; or writing of the report. SB, JBS, and RAH had access to the raw data. The corresponding author had full access to all of the data and the final responsibility to submit for publication.

Results

The ACCORD sample for equation development included 9635 participants (table 1); 616 (6.0%) of 10 251 ACCORD participants were excluded due to missing candidate predictor variables. The mean age of participants was 62.8 years, 38.0% were women, and mean baseline HbA_{1c} was 8.3% (63 mmol/mol). Participants were followed up for a median of 4.7

years. Of 9635 patients, 292 (3.0%) developed renal failure or end-stage renal disease, 776 (8.1%) developed severe vision loss, 1201 (12.5%) lost pressure sensation in their feet, 880 (9.1%) had a myocardial infarction (fatal or nonfatal), 197 (2.0%) had a stroke (fatal or non-fatal), 454 (4.7%) developed congestive heart failure, 332 (3.4%) died from cardiovascular disease, and 719 (7.5%) died from any cause during the study follow-up period.

Tables 2–4 provide the RECODE coefficients and examples for calculating 10-year risk. The elastic net regularisation method revealed that the most important variables for predicting microvascular outcomes were HbA_{1c}, followed by age and serum creatinine; and for predicting cardiovascular outcomes, the most important variables were history of cardiovascular disease, followed by age and tobacco smoking.

In internal 10-times cross-validations of the micro-vascular outcomes (table 5), RECODE C-statistics for discrimination were a mean of 0.70 for nephropathy outcomes (range 0.60–0.84 across outcome definitions), 0.63 for retinopathy outcomes (0.55–0.68), and 0.61 for neuropathy outcomes (0.57–0.64). Calibration slopes between expected and observed outcome rates ranged from 0.78 to 1.28 (ideal=1) and calibration intercepts ranged from –0.009 to 0.073 (ideal=0; table 5, figure 1). Ten of the 15 microvascular equations passed the GND test for calibration, and three that did not pass were for three definitions of nephropathy (doubling of serum creatinine or >20 mL/min per 1.73 m² decrease in eGFR, macroalbuminuria, and renal failure or end-stage renal disease), and one definition of retinopathy (cataract extraction; table 2, table 5). The subset of equations that failed the GND test did so due to one subgroup with higher expected than observed outcome rates—that is, risk for one decile of the population was mis-estimated; figure 1); the degree of error was nevertheless lower than for the alternative UKPDS OM2 equations.

In internal 10-times cross-validations of the cardiovascular disease outcomes (table 5), RECODE C-statistics for discrimination ranged from a low of 0.69 for atherosclerotic cardiovascular disease and fatal or non-fatal myocardial infarction to a high of 0.75 for congestive heart failure, and the calibration of the RECODE was high (slopes 0.96 to 1.16, intercepts –0.005 to 0.001). All cardiovascular outcomes passed GND tests (table 5, figure 1). The all-cause mortality RECODE had a C-statistic of 0.70, a calibration slope of 1.03, a calibration intercept of –0.002, and passed the GND test.

Despite the relatively narrow inclusion and exclusion criteria in ACCORD, individual participant risks for both microvascular and cardiovascular risk varied dramatically, as shown in figure 1. Participants with high predicted microvascular risk were also more commonly those with higher predicted cardiovascular risk (Pearson correlation coefficients 0.32–0.90; appendix). In sensitivity analyses, in which we included people with missing covariates, results did not change to within rounding error (appendix).

The study sample for validation of RECODE included 1018 DPPOS trial participants who developed type 2 diabetes during the study (median follow-up 6.0 years) and 4760 Look AHEAD trial participants with type 2 diabetes (median follow-up 10.6 years; table 1). Of the 1018 DPPOS participants, 184 (18.1%) developed nephropathy (microalbuminuria, macroalbuminuria, renal failure, or end-stage renal disease), 115 (11.3%) developed

retinopathy (requiring photocoagulation or vitrectomy), and 120 (11.8%) developed neuropathy (pressure sensation loss) during follow-up. Of the 4760 Look AHEAD participants, 462 (9.7%) had an atherosclerotic cardiovascular disease event, 332 (7.0%) had a myocardial infarction (fatal or non-fatal), 157 (3.3%) had a stroke (fatal or non-fatal), 210 (4.4%) developed congestive heart failure, 106 (2.2%) died from cardiovascular disease, and 355 (7.5%) died from any cause during follow-up.

In external validation against DPPOS data, RECODE predicted microvascular outcomes with C-statistics varying from 0.57 (for retinopathy) to 0.69 (for neuropathy; table 5). Once again, calibration was high (all passed GND tests, table 5, figure 1). In external validation against Look AHEAD data, RECODE predicted cardiovascular outcomes with C-statistics ranging from 0.67 (for stroke) to 0.79 (for cardiovascular mortality; table 5), with calibration outcomes passing GND tests for myocardial infarction, stroke, and congestive heart failure (table 5, figure 1). The external validation of the all-cause mortality equation had a C-statistic of 0.71 and also passed the GND test (table 5). As with ACCORD participants, DPPOS participants' microvascular risks and Look AHEAD participants' cardiovascular risks varied quite widely (figure 1).

Compared with the new equations, the UKPDS OM2 equations had similar or slightly worse discrimination for both microvascular and cardiovascular outcomes (C-statistics ranging from 0.54 to 0.67; table 5, figure 1). However, the UKPDS OM2 had much poorer calibration across all outcomes (slopes ranging from 0.06 to 1.12, intercepts from -0.038 to 0.043, failing GND tests; table 5). The ACC/AHA PCEs also had worse discrimination for atherosclerotic cardiovascular disease risk than RECODE (C-statistics of 0.60 in ACCORD and 0.66 in Look AHEAD), and much worse calibration due to overestimating event rates (slopes of 0.28 in ACCORD and 0.39 in Look AHEAD, respectively; table 5).

We found a positive NRI for RECODE compared with either of the older sets of equations for every outcome (figure 2; appendix). The positive NRI results were particularly driven by the fact that older equations tended to overestimate risk of people who were actually low risk, whereas RECODE correctly identified people at low risk.

Discussion

We developed risk equations for microvascular and cardiovascular complications of type 2 diabetes using data commonly available in clinical practice. These new equations updated previously existing equations using newer, publicly available data through replicable open-source statistical code and an online risk calculator.²⁵ We developed the equations through rigorous cross-validation techniques to avoid overfitting and overestimation of discrimination statistics. We rigorously externally validated where possible, and successfully distinguished high-risk from low-risk patients more accurately than previous equations. Unlike previous studies, our external validation steps used individual patient data rather than aggregate study outcomes.⁸

Commonly used existing equations, such as those in the UKPDS OM2, were based on older cohorts (eg, the UKPDS began follow-up in 1977) than ACCORD, on which our equations

were based.⁹ Substantial changes in diabetes demographics (eg, younger age at diagnosis) and treatment (eg, earlier therapy and aggressive management according to cardiovascular disease risk) in the decades since UKPDS was started mean that estimates are difficult to apply in modern settings. The UKPDS equations, for instance, typically overestimated risk by 2–4 times in our assessment. The ACC/AHA equations do not include microvascular complications, omit important covariates not widely available in their development datasets (eg, statin treatment), and substantially overestimate risk, possibly due to older data used to derive the equations.^{26–28} RECODE are derived from more recent data so might mitigate such problems.

Wide, but consistently predictable, variation in risk was observed even within seemingly similar trial populations that were subject to strict inclusion and exclusion criteria. Given how different the entry criteria were across studies (high-risk in ACCORD, newly diagnosed in DPPOS, and low-risk in Look AHEAD), the consistent discrimination and calibration of the multivariable RECODE attest to high generalisability of the equations. The likelihood is low that we have developed equations with high internal discrimination but with poor generalisability.

We observed strong correlations between those at high microvascular risk and those at high cardiovascular risk. Furthermore, multiple risk factors contributed to high risk among microvascular and cardiovascular outcome, indicating the utility of multivariable risk scores over single biomarker values for guiding treatment.²⁹

The results of this study should nevertheless be interpreted in light of several limitations. Although our focus was on calibration (ensuring that risk predictions were in the correct range of observed outcome rates), we also calculated discrimination (distinguishing high risk from low risk) using C-statistics through a 10-times internal cross-validation method that will tend to have a lower C-statistic (usually by 0.05³⁰) than when calculated with the whole sample. Although the C-statistic is imperfect and might change over time,³¹ its value was lower than desired for some of our equations, and lower than some other equations for all-cause mortality tested among other populations.³²

Both the development and validation of the equations were done using data from randomised clinical trials, which enhances the internal validity of the equations given careful adjudication of clinical endpoints (unlike electronic medical record data in which diagnoses are wrong or missing), but might limit generalisability if risk factors for complications operate in a substantially different manner in populations that are not included in the trials. Although both trial and observational datasets have patients on intensive treatment regimens, observational datasets suffer from selection biases as to which patients are on intensive treatment; by contrast, the trial data used here allowed adjustment for random assignment to intensive treatment for glycaemic control, blood pressure lowering, and lipid lowering, thus avoiding selection bias. Furthermore, we did internal and external validation across studies with diverse participant characteristics, with notable variations in both microvascular and cardiovascular risk within and between the ACCORD, DPPOS, and Look AHEAD participants. Nevertheless, due to differences in study endpoint definitions, we could only validate a subset of possible microvascular endpoint definitions. Patient-relevant definitions

of clinical outcomes might not be fully concordant with clinical trial definitions; for example, the degree of pain associated with neuropathy could be more important for patients than neurological examination findings.^{33,34} Observer errors and laboratory measurement variability can lead to further uncertainty. The data were also limited to US populations so RECODE might generalise poorly to international contexts. For example, laboratory biomarkers might not be routinely measured in some settings; however, we did not derive equations that omitted laboratory biomarker data.

The data used in this study were collected before widespread use of some diabetes therapies, such as glucagon-like peptide-1 receptor agonists and sodium-glucose co-transporter-2 inhibitors. These newer agents have been suggested to reduce cardiovascular risk and their benefits and risks may not be fully captured by single biomarker values such as HbA_{1c}.^{35–38} Additional important omitted variables might further help explain microvascular and cardiovascular risk and could be considered for inclusion in RECODE once additional data become available; these include insulin resistance, cardiovascular measurements such as carotid intima-media thickness, and emerging biomarkers. Inclusion of such variables in the risk equations could, however, reduce their broad clinical utility.

The next logical step for research is to further validate RECODE in longitudinal cohort datasets. Unlike the trial data that were used for derivation and validation in this study, these data are not yet publicly available. Such further validation could also to test the equations over longer time periods than was possible using available trial data.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Research reported in this publication was supported by the National Institute for Diabetes and Digestive and Kidney Diseases (NIDDK); the National Heart, Lung, and Blood Institute (NHLBI); and the National Institute on Minority Health and Health Disparities of the National Institutes of Health under Award Numbers DP2MD010478 (SB), U54MD010724 (SB), K08HL121056 (SB), P60DK20572 (RH, JS), K23DK109200 (SAB); and by the US Department of Veterans Affairs under grants IIR11-088 (RAH, JBS) and CDA13-021 (RAH, JBS). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This manuscript was prepared using ACCORD, DPPOS and Look AHEAD research materials obtained from the NHLBI Biologic Specimen and Data Repository Information Coordinating Center and the NIDDK Central Database Repository. The manuscript does not necessary reflect the opinions or views of the ACCORD, DPPOS, Look AHEAD, NHLBI, or NIDDK. This manuscript was not prepared under the auspices of the DPP and DPPOS, or Look AHEAD and does not represent analyses or conclusions of the Look AHEAD Research Group, the DPP Research Group, the NIDDK Central Repositories, or the NIH.

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; 174:1227–34. [PubMed: 24979148]
2. Stone NJ, Robinson J, Lichtenstein AH, et al. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2013; 129(25 suppl 2):S1–45. [PubMed: 24222016]

3. 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; 4:922–32. [PubMed: 27717768]
4. 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. [PubMed: 24409102]
5. 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; 154:680–83.
6. Hippisley-Cox J, Coupland C, Vinogradova Y, et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. *BMJ.* 2008; 336:1475–82. [PubMed: 18573856]
7. Goff DC, Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk. *Circulation.* 2014; 129:S49–73. [PubMed: 24222018]
8. Palmer AJ. Computer modeling of diabetes and its complications: a report on the fifth mount hood challenge meeting. *Value Health.* 2013; 16:670–85. [PubMed: 23796302]
9. 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; 56:1925–33. [PubMed: 23793713]
10. Action to Control Cardiovascular Risk in Diabetes Study Group. Effects of intensive glucose lowering in type 2 diabetes. *N Engl J Med.* 2008; 358:2545–59. [PubMed: 18539917]
11. Action to Control Cardiovascular Risk in Diabetes Study Group. Effects of combination lipid therapy in type 2 diabetes mellitus. *N Engl J Med.* 2010; 362:1563–74. [PubMed: 20228404]
12. Action to Control Cardiovascular Risk in Diabetes Study Group. Effects of intensive blood-pressure control in type 2 diabetes mellitus. *N Engl J Med.* 2010; 362:1575–85. [PubMed: 20228401]
13. Burke JF, Hayward RA, Nelson JP, Kent DM. Using internally developed risk models to assess heterogeneity in treatment effects in clinical trials. *Circ Cardiovasc Qual Outcomes.* 2014; 7:163–69. [PubMed: 24425710]
14. 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; 3:866–75. [PubMed: 26377054]
15. The Look AHEAD Research Group. Cardiovascular effects of intensive lifestyle intervention in type 2 diabetes. *N Engl J Med.* 2013; 369:145–54. [PubMed: 23796131]
16. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med.* 2009; 150:604–12. [PubMed: 19414839]
17. Moghtaderi A, Bakhshpour A, Rashidi H. Validation of Michigan neuropathy screening instrument for diabetic peripheral neuropathy. *Clin Neurol Neurosurg.* 2006; 108:477–81. [PubMed: 16150538]
18. Tibshirani R, Bien J, Friedman J, et al. Strong rules for discarding predictors in lasso-type problems. *J R Stat Soc Ser B Stat Methodol.* 2012; 74:245–66.
19. Simon N, Friedman J, Hastie T, Tibshirani R. Regularization paths for Cox’s proportional hazards model via coordinate descent. *J Stat Softw.* 2011; 39:1.
20. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med.* 2011; 30:377–399. [PubMed: 21225900]
21. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics.* 2000; 56:337–44. [PubMed: 10877287]
22. van der Laan, MJ. Targeted learning: causal inference for observational and experimental data. New York: Springer; 2011.
23. Demler OV, Paynter NP, Cook NR. Tests of calibration and goodness-of-fit in the survival setting. *Stat Med.* 2015; 34:1659–80. [PubMed: 25684707]

24. Kerr KF, Wang Z, Janes H, McClelland RL, Psaty BM, Pepe MS. Net reclassification indices for evaluating risk prediction instruments: a critical review. *Epidemiol Camb Mass*. 2014; 25:114–21.
25. Basu, S. RECODE. Risk equations for complications of type 2 diabetes. <https://sanjaybasu.shinyapps.io/recode/> ((accessed March 24, 2017)
26. 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; 129:268–69. [PubMed: 24334112]
27. Ridker PM, Cook NR. Statins: new American guidelines for prevention of cardiovascular disease. *Lancet*. 2013; 382:1762–65. [PubMed: 24268611]
28. Cook NR, Ridker PM. Calibration of the pooled cohort equations for atherosclerotic cardiovascular disease: an update. *Ann Intern Med*. 2016; 165:786–94. [PubMed: 27723890]
29. 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; 4:922–32. [PubMed: 27717768]
30. Parker BJ, Günter S, Bedo J. Stratification bias in low signal microarray studies. *BMC Bioinformatics*. 2007; 8:326. [PubMed: 17764577]
31. Chambless LE, Diao G. Estimation of time-dependent area under the ROC curve for long-term risk prediction. *Stat Med*. 2006; 25:3474–86. [PubMed: 16220486]
32. Yang X, So WY, Tong PCY, et al. Development and validation of an all-cause mortality risk score in type 2 diabetes: the Hong Kong Diabetes Registry. *Arch Intern Med*. 2008; 168:451–57. [PubMed: 18332288]
33. Chan GCW, Tang SCW. Diabetic nephropathy: landmark clinical trials and tribulations. *Nephrol Dial Transplant*. 2016; 31:359–68. [PubMed: 25637638]
34. Rodriguez-Gutierrez R, Gionfriddo MR, Ospina NS, et al. Shared decision making in endocrinology: present and future directions. *Lancet Diabetes Endocrinol*. 2016; 4:706–16. [PubMed: 26915314]
35. Marso SP, Daniels GH, Brown-Frandsen K, et al. Liraglutide and cardiovascular outcomes in type 2 diabetes. *N Engl J Med*. 2016; 375:311–22. [PubMed: 27295427]
36. Marso SP, Bain SC, Consoli A, et al. Semaglutide and cardiovascular outcomes in patients with type 2 diabetes. *N Engl J Med*. 2016; 375:1834–44. [PubMed: 27633186]
37. Zinman B, Wanner C, Lachin JM, et al. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. *N Engl J Med*. 2015; 373:2117–28. [PubMed: 26378978]
38. Ingelfinger JR, Rosen CJ. Cardiac and renovascular complications in type 2 diabetes—is there hope? *N Engl J Med*. 2016; 375:380–82. [PubMed: 27331286]

Research in context

Evidence before this study

Population-based clinical care strategies promote outreach to individuals with a high risk of type 2 diabetes complications, to provide support and encourage treatment decisions that might lower risk. Tailoring therapy based on informed risk estimation might reduce complications of type 2 diabetes, complications induced by treatment itself, and overall costs of treatment and management of complications. However, commonly used risk equations systematically mis-estimate the risk of both microvascular and cardiovascular complications among diverse participants. The RECODE equations were derived and validated on all three National Institutes for Health-funded type 2 diabetes studies for which publicly available, de-identified individual participant data have been released, permitting a high level of transparency and replicability. Derivation and validation of risk equations from these data can enhance the internal validity of the equations, given careful adjudication of clinical endpoints. We compared the new equations with the UK Prospective Diabetes Study Outcomes Model 2 (UKPDS OM2) and American College of Cardiology/American Heart Association pooled cohort equations (ACC/AHA) risk scores because the latter two are the most commonly used risk equations presently for evaluating microvascular and cardiovascular risk among persons with type 2 diabetes.

Added value of this study

We developed Risk Equations for Complications Of type 2 Diabetes (RECODE) that use data commonly available in clinical practice to estimate risk of microvascular and cardiovascular disease outcomes. RECODE better estimated microvascular and cardiovascular outcomes when compared with older risk equations (UKPDS OM2/AHA pooled cohort equations) using clinical trial datasets different from those used to develop RECODE. We also produced an online risk calculator.

Implications of all the available evidence

Updated risk equations for complications of type 2 diabetes might be helpful for risk prediction to provide outreach to high-risk patients, and for comparative effectiveness research that relies on risk equations to simulate treatment guidelines and cost-effectiveness of interventions.

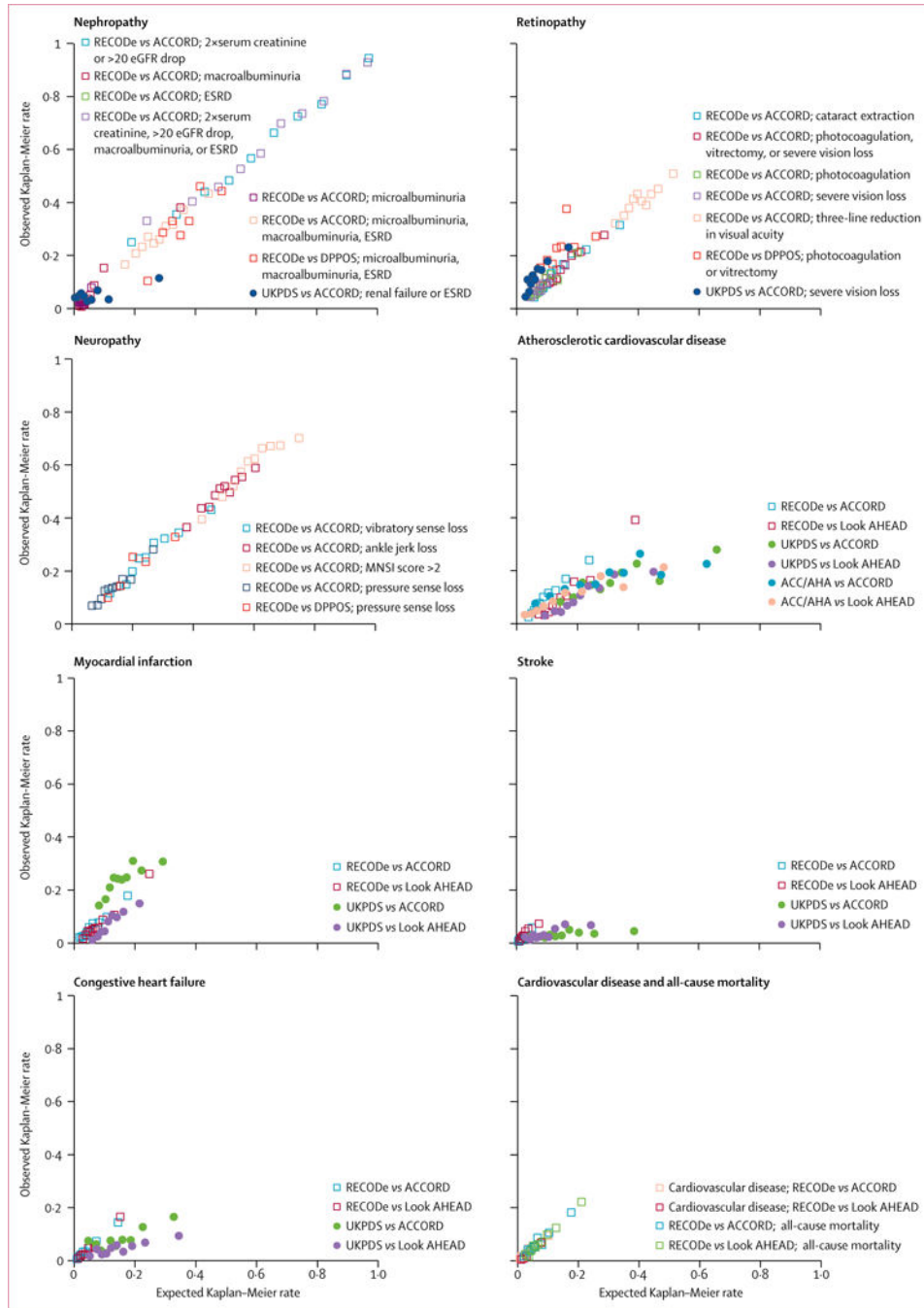


Figure 1. Calibration plots

Kaplan-Meier event rates over 10 years predicted by RECODE versus observed rates in the ACCORD study (n=9635, 2001–09), DPPOS (n=1018, 1996–2001), and Look AHEAD study (n=4760, 2001–12). Predictions using UKPDS OM2 and the ACC/AHA PCEs for atherosclerotic cardiovascular disease (non-fatal or fatal myocardial infarction or stroke) are presented if available. Points are displayed for deciles of predicted and observed Kaplan-Meier event rates, with fewer centiles than deciles used if fewer than 5 events were observed per group or to prevent unstable inferences per guidelines. ACCORD=Action to Control

Cardiovascular Risk in Diabetes study. ACC/AHA PCEs=American College of Cardiology/
American Heart Association Pooled Cohort Equations. DPPOS=Diabetes Prevention
Program Outcomes Study. Look AHEAD=Action for Health in Diabetes study. ESRD=end-
stage renal disease. MNSI=Michigan Neuropathy Screening Instrument. RECODE=Risk
Equations for Complications Of type 2 Diabetes. UKPDS OM2=United Kingdom
Prospective Diabetes Study Outcomes Model 2.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

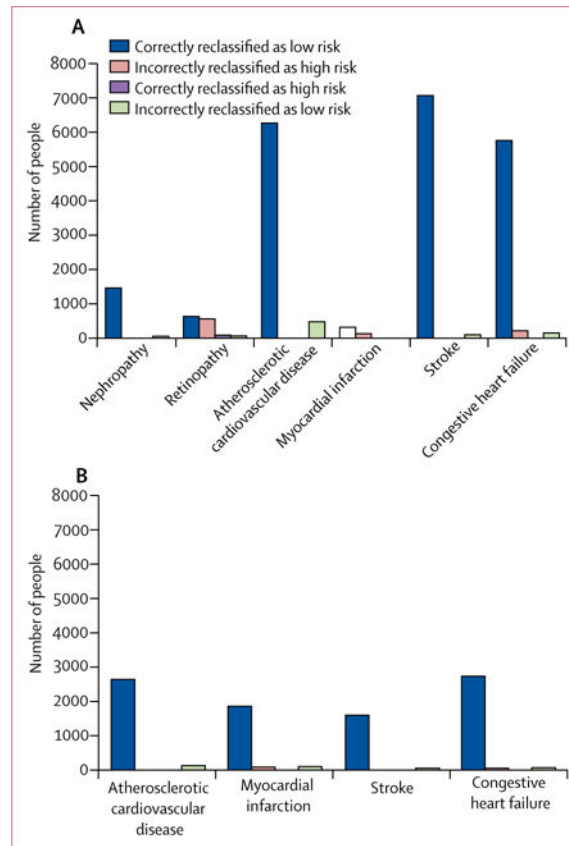


Figure 2. Outcome classification by RECODE versus UKPDS OM2
 Number of people correctly classified or misclassified as high risk or low risk using RECODE versus UKPDS OM2 in (A) ACCORD and (B) Look AHEAD. RECODE=Risk Equations for Complications Of type 2 Diabetes. ACCORD=Action to Control Cardiovascular Risk in Diabetes. Look AHEAD=Action for Health in Diabetes.

Table 1

Baseline characteristics

	ACCORD (n=9635)	DPPOS (n=1018)	Look AHEAD (n=4760)
Demographics			
Age, years	62.8 (6.7)	50.9 (8.0)	58.9 (6.7)
Aged 75 years or older	521 (5%)	0	31 (1%)
Sex			
Women	3662 (38%)	680 (67%)	2784 (59%)
Men	5973 (62%)	338 (33%)	1976 (42%)
Ethnicity			
Black	1834 (19%)	244 (24%)	776 (16%)
Hispanic or Latino	678 (7%)	175 (17%)	670 (14%)
Clinical features			
Tobacco smoking, current	1179 (12%)	52 (5%)	202 (4%)
BMI, kg/m ²	32.2 (5.4)	33.9 (5.9)	36.0 (5.9)
Blood pressure, seated			
Systolic, mm Hg	136.5 (17.1)	123.7 (14.0)	129.0 (17.1)
Diastolic, mm Hg	74.9 (10.7)	76.4 (8.8)	70.2 (9.5)
Pulse pressure, mm Hg	61.5 (14.6)	47.8 (12.1)	58.9 (14.3)
Heart rate, beats/min	72.7 (11.8)	N/A	71.4 (11.4)
History of cardiovascular disease	3437 (36%)	12 (1%)	665 (14%)
Drug use			
Blood pressure-lowering drugs	8109 (84%)	770 (76%)	3410 (72%)
Oral diabetes drugs (including metformin)	8024 (83%)	336 (33%)	3246 (68%)
Insulin treatment	3403 (35%)	N/A	724 (15%)
Statins	6148 (64%)	721 (71%)	2142 (45%)
Fibrates	601 (6%)	N/A	324 (7%)
Anticoagulant use	303 (3%)	N/A	N/A
Non-steroidal anti-inflammatory use	851 (9%)	N/A	N/A
Platelet aggregate inhibitor use	466 (5%)	N/A	N/A
Daily aspirin use	5274 (55%)	N/A	2140 (45%)
Biomarkers			
HbA _{1c} , %	8.3 (1.1)	6.1 (0.7)	7.3 (1.2)
HbA _{1c} , mmol/mol	67 (9)	43 (5)	56 (9)
Total cholesterol, mg/dL	183.2 (41.7)	196.0 (43.7)	191.4 (37.3)
HDL cholesterol, mg/dL	41.8 (11.6)	46.0 (12.3)	43.5 (11.9)
LDL cholesterol, mg/dL	104.7 (33.8)	99.5 (27.3)	112.7 (32.1)
Triglycerides, mg/dL	190.7 (145.8)	162.7 (256.8)	N/A

	ACCORD (n=9635)	DPPOS (n=1018)	Look AHEAD (n=4760)
Fasting plasma glucose, mg/dL	175.3 (55.8)	115.8 (22.7)	153.2 (45.6)
Alanine aminotransferase, IU/L	27.5 (16.0)	N/A	N/A
Creatine phosphokinase, IU/L	140.3 (130.2)	N/A	N/A
Serum potassium, mmol/L	4.5 (0.5)	N/A	N/A
Serum creatinine, mg/dL	0.9 (0.2)	0.8 (0.2)	0.8 (0.2)
Estimated glomerular filtration rate, mL/min/1.73 m ²	90.9 (27.3)	98.8 (15.7)	89.9 (16.1)
Urine albumin, mg/dL	10.7 (37.3)	10.8 (6.4)	4.8 (23.0)
Urine creatinine, mg/dL	127.3 (65.4)	123.3 (73.6)	121.0 (67.0)
Urine albumin:creatinine ratio, mg/g	99.2 (359.4)	N/A	43.1 (201.5)

Data are mean (SD) or n (%). N/A=not available in the dataset.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Coefficients for nephropathy and retinopathy outcomes using RECODE

	Nephropathy				Retinopathy						
	Micro-albuminuria (n=1551)	Macro-albuminuria (n=627)	Renal failure or end-stage renal disease (n=292)	Doubling of serum creatinine or >20 mL/min per 1.73 m ² decrease in eGFR (n=5910)	Macro-albuminuria, renal failure, end-stage renal disease, doubling of serum creatinine, or >20 mL/min per 1.73 m ² decrease in eGFR (n=6195)	Micro-albuminuria, macro-albuminuria, renal failure, or end-stage renal disease (n=2521)	Photo-coagulation or vitrectomy (n=901)	Cataract extraction (n=1476)	Three-line reduction in visual acuity (n=3559)	Severe vision loss (n=776)	Photo-coagulation or vitrectomy, or severe vision loss (n=1468)
Demographics											
Age, years	0.02114	0.00733	-0.01938	0.01222	0.01163	0.01979	-0.003326	0.07457	0.01452	0.02285	0.00969
Women	0.16956	0.27380	-0.01129	-0.60460	-0.56143	0.10390	0.17187	0.27510	0.07600	0.22640	0.22558
Ethnicity											
Black	-0.00804	-0.00556	0.08812	0.34610	0.29559	-0.02650	0.05655	-0.24960	-0.10650	-0.16770	-0.06502
Hispanic or Latino	0.17342	0.35630	0.23380	-0.13600	-0.09178	0.24012
Clinical features											
Tobacco smoking, current	0.28362	0.10010	0.14830	-0.08675	-0.06691	0.26523
Systolic blood pressure, mm Hg	0.00334	-0.00101	0.00303	0.00828	0.00888	0.00120	0.01228	0.00146	0.00138	0.00824	0.01090
Cardiovascular disease history	0.22312	0.25570	-0.02164	0.19560	0.019012	0.18020	0.28350	0.28350	-0.04092	0.1127	0.2186
Drug use											
Blood pressure- lowering drugs	0.28372	0.24180	-0.07952	0.14860	0.18583	0.29453	0.20691	0.10420	-0.00188	0.06393	0.18019
Oral diabetes drugs	0.06584	0.09015	-0.12560	0.11610	0.08134	0.04787	-0.40746	-0.16670	0.06002	-0.23490	-0.31715
Anticoagulants	0.42199	0.01091	0.03199	0.04788	0.01498	0.4226
Biomarkers											
HbA _{1c} , %	0.13847	0.09639	0.13690	0.10290	0.10364	0.13357	0.22339	0.09659	0.09346	0.1449	0.17249
Total cholesterol, mg/dL	0.00034	0.00009	-0.00111	0.00011	0.00064	0.00038	-0.00141	-0.00104	-0.00047	-0.00017	-0.00106
HDL cholesterol, mg/dL	-0.00970	-0.01135	0.00629	-0.00590	-0.00521	-0.00649	0.01181	0.00732	0.00015	0.00545	0.00845
Serum creatinine, mg/dL	0.67026	1.14900	0.86090	-3.41100	-2.89893	0.46067	0.81582	0.2835	0.06586	0.69470	0.84481

Nephropathy			Retinopathy							
Micro-albuminuria (n=1551)	Macro-albuminuria (n=627)	Renal failure or end-stage renal disease (n=292)	Doubling of serum creatinine or >20 mL/min per 1.73 m ² decrease in eGFR (n=5910)	Macro-albuminuria, renal failure, end-stage renal disease, doubling of serum creatinine, or >20 mL/min per 1.73 m ² decrease in eGFR (n=6195)	Micro-albuminuria, macro-albuminuria, renal failure, or end-stage renal disease (n=2321)	Photo-coagulation or vitrectomy (n=901)	Cataract extraction (n=1476)	Three-line reduction in visual acuity (n=3559)	Severe vision loss (n=776)	Photo-coagulation or vitrectomy, or severe vision loss (n=1468)
Urine albumin:creatinine ratio, mg/g	..	0.01335	0.00036	0.00021	0.00016	0.00016	0.00020	..

All risk factors listed in table 1 were considered for inclusion; the ones listed in tables 2 and 3 were selected for inclusion. The 10-year risk of an outcome can be computed as $1 - \lambda \cdot \exp(\hat{\Sigma}(\beta \times x))$, where β are the equation coefficients and x are the values for each covariate for an individual patient within the cohort under study. λ values for the most clinically relevant outcomes of renal failure or end-stage renal disease, and severe vision loss are: 0.973 for renal failure or end-stage renal disease, and 0.921 for vision loss, and corresponding values of $\text{mean}(\hat{\Sigma}(\beta \times x))$ were 0.23 for renal failure or end-stage renal disease, and 4.56 for vision loss. For example, a 60-year-old non-smoking white man with systolic blood pressure 140 mm Hg, without history of cardiovascular disease, not on any medications, and with HbA_{1c} of 8%, total cholesterol of 190 mg/dL, HDL of 50 mg/dL, serum creatinine 1.1 mg/dL, and urine microalbumin:creatinine ratio of 10 mg/g would have a risk of renal failure/end-stage renal disease of $1 - 0.973 \cdot \exp(-0.01938 \times 60 + 0.003027 \times 140 + 0.1369 \times 8 - 0.001112 \times 190 + 0.006289 \times 50 + 0.8609 \times 1.1 + 0.000362 \times 10 - 0.23) = 0.085$ or a 8.5% 10-year risk, where 0.23 is the $\text{mean}(\hat{\Sigma}(\beta \times x))$. People without a known covariate can have the associated term omitted from the equations to enable calculation of risk without the missing data. RECODE=Risk Equations for Complications of type 2 Diabetes.

Table 3

Coefficients for neuropathy outcomes using RECODe

	MNSI >2 (n=3221)	Vibratory sensation loss (n=2034)	Ankle jerk loss (n=3135)	Pressuresensation loss (n=1201)
Demographics				
Age, years	0.02434	0.04640	0.01832	0.03022
Women	-0.10030	-0.29080	-0.08671	-0.18680
Ethnicity				
Black	-0.22500	-0.17970	-0.1918	-0.09448
Hispanic or Latino
Clinical features				
Systolic blood pressure, mm Hg	-0.00414	0.00304	-0.00291	0.00456
History of cardiovascular disease	-0.06334	0.17080	-0.07515	0.26672
Drug use				
Blood pressure-lowering drugs	0.18210	0.12220	0.06106	0.18192
Oral diabetes drugs	-0.08004	-0.23430	-0.02732	-0.25747
Biomarkers				
HbA _{1c} , %	0.04517	0.07833	0.03741	0.18866
Total cholesterol, mg/dL	-0.00044	0.00031	-0.00016	0.00219
HDL cholesterol, mg/dL	-0.00807	-0.00513	-0.00603	-0.00539
Serum creatinine, mg/dL	0.07714	0.38360	0.01232	0.60442
Urine albumin:creatinine ratio, mg/g	0.00010	0.00013	0.00007	..

All risk factors listed in table 1 were considered for inclusion; the ones listed in tables 2 and 3 were selected for inclusion. The 10-year risk of an outcome can be computed as $1 - \lambda^{\exp(\Sigma(\beta \times x) - \text{mean}(\Sigma(\beta \times x)))}$, where β are the equation coefficients and x are the values for each covariate for an individual patient within the cohort under study. λ value for the most clinically relevant outcome of loss of pressure sensation was 0.870, and corresponding value of $\text{mean}(\Sigma(\beta \times x))$ was 4.75 for loss of pressure sensation. For example, a 60-year old white man with systolic blood pressure 140 mm Hg, without history of cardiovascular disease, not on any medications, and with HbA_{1c} of 8%, total cholesterol of 190 mg/dL, HDL of 50 mg/dL, and serum creatinine 1.1 mg/dL would have a risk of pressure sensation loss of $1 - 0.87^{\exp(3.022e - 02 \times 60 + 4.561e - 03 \times 140 + 1.887e - 01 \times 8 + 2.185e - 03 \times 190 - 5.389e - 03 \times 50 + 6.044e - 01 \times 1.1 - 4.75)} = 0.13$, or 13% 10-year risk, where 4.75 is the $\text{mean}(\Sigma(\beta \times x))$. People without a known covariate can have the associated term omitted from the equations to enable calculation of risk without the missing data. RECODe=Risk Equations for Complications of type 2 Diabetes. MNSI=Michigan Neuropathy Screening Instrument.

Table 4

Coefficients for outcomes of RECODE for macrovascular outcomes

	ASCVD (n=1053)	MI (fatal or non-fatal; n=880)	Stroke (fatal or non-fatal; n=197)	CHF (n=454)	Cardiovascular mortality (n=332)	All-cause mortality (n=719)
Demographics						
Age, years	0.03412	0.04363	0.02896	0.05268	0.05501	0.06703
Women	-0.16720	-0.20660	-0.00326	0.25290	-0.30560	-0.15290
Ethnicity						
Black	-0.11870	-0.11630	0.27160	-0.04969	0.07957	-0.02393
Clinical features						
Tobacco smoking, current	0.15100	0.23580	0.16650	0.29050	-0.05764	0.5399
Systolic blood pressure, mm Hg	0.00007	-0.00514	0.01659	0.00121	-0.00394	-0.00299
History of cardiovascular disease	0.77840	0.96180	0.41380	1.00700	1.01600	0.58880
Drug use						
Blood pressure-lowering drugs	0.05579	-0.12480	0.15980	0.63890	-0.15770	0.08776
Statins	-0.03361	0.04699	-0.18870	-0.11750	-0.20450	-0.26810
Anticoagulants	0.25240	0.54400	-0.13870	0.73650	0.69460	0.40360
Biomarkers						
HbA _{1c} , %	0.17160	0.21350	0.33650	0.20920	0.24540	0.16590
Total cholesterol, mg/dL	0.00193	0.00019	0.00171	-0.00136	-0.00127	-0.00095
HDL cholesterol, mg/dL	-0.00837	-0.01358	-0.00639	-0.01758	-0.01081	-0.00438
Serum creatinine, mg/dL	0.43550	0.08027	0.59550	0.82140	0.45440	0.35970
Urine albumin:creatinine ratio, mg/g	0.00033	0.00042	0.00030	0.00041	0.00047	0.00039

All risk factors listed in table 1 were considered for inclusion; the ones listed in table 4 were selected for inclusion. The 10-year risk of each outcome can be computed as $1 - \lambda \cdot \exp(\sum(\beta \times x)) - \text{mean}(\sum(\beta \times x))$, where β are the equation coefficients and x are the values for each covariate for an individual patient within the cohort under study. λ values were 0.85 for ASCVD, 0.93 for fatal or non-fatal MI, 0.98 for fatal or non-fatal stroke, 0.96 for CHF, 0.97 for cardiovascular mortality, and 0.93 for all-cause mortality, and $\text{mean}(\sum(\beta \times x))$ values were 3.65 for fatal or non-fatal MI, 6.96 for fatal or non-fatal stroke, 5.15 for CHF, 3.97 for cardiovascular mortality, and 4.66 for all-cause mortality in the validation study. For example, a 60-year-old white man with systolic blood pressure 140 mm Hg, without history of CVD, not on any medications, and with HbA_{1c} of 8%, total cholesterol of 190 mg/dL, HDL of 50 mg/dL, serum creatinine 1.1 mg/dL, and urine albumin:creatinine ratio of 10 mg/g, would have an all-cause mortality risk of $1 - 0.93 \cdot \exp(6.703e-02 \times 60 - 2.988e-03 \times 140 + 1.659e-01 \times 8 - 9.478e-04 \times 190 - 4.378e-03 \times 50 + 3.597e-01 \times 1.1 + 3.889e-04 \times 10 - 4.66) = 0.09$, or 9% 10-year risk, where 4.66 is the $\text{mean}(\sum(\beta \times x))$. People without a known covariate can have the associated term omitted from the equations to enable calculation of risk without the missing data. An online risk

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

calculator is available in both SI and US or conventional units.²⁵ RECOD_s=Risk Equations for Complications of type 2 Diabetes. ASCVD=atherosclerotic cardiovascular disease (non-fatal or fatal myocardial infarction or stroke). MI=myocardial infarction. CHF=congestive heart failure. CVD=cardiovascular disease.

Table 5

Internal and external validation statistics for microvascular and macrovascular RECODE

	Internal validation: RECODE			External validation: RECODE			Alternative risk equations: UKPDS OMI2			Alternative risk equations: ACC/AHA PCEs		
	Discrimination: C-statistic	Calibration: slope/intercept/ χ^2 , p value*	p value*	Discrimination: C-statistic	Calibration: slope/intercept/ χ^2 , p value*	p value*	Discrimination: C-statistic	Calibration: slope/intercept/ χ^2 , p value*	p value*	Discrimination: C-statistic	Calibration: slope/intercept/ χ^2 , p value*	p value*
Microvascular outcomes												
Nephropathy												
Microalbuminuria	0.62 (0.61–0.64)	0.94/0.015/5.7, 0.77
Macroalbuminuria	0.84 (0.82–0.86)	1.14/–0.009/79.4, <0.0001
Renal failure or end-stage renal disease	0.60 (0.56–0.64)	1.28/0.0003/30.8, <0.0001	0.54 (0.50, 0.59)	0.19/0.035/242.6, <0.0001
Doubling of serum creatinine or >20 mL/min per 1.73 m ² decrease in eGFR	0.76 (0.75–0.77)	0.91/0.053/42.9, <0.0001
Macroalbuminuria, renal failure, end-stage renal disease, doubling of serum creatinine, or >20 mL/min per 1.73 m ² decrease in eGFR	0.73 (0.72–0.74)	0.86/0.085/74.1, <0.0001
Macroalbuminuria, microalbuminuria, renal failure, or end-stage renal disease	0.61 (0.60–0.63)	0.96/0.011/4.6, 0.87	0.65 (0.61, 0.70)	1.31/–0.15/9.3, 0.16
Retinopathy												
Photocoagulation or vitrectomy	0.65 (0.63–0.67)	1.03/–0.003/15.7, 0.07	0.57 (0.51, 0.63)	0.72/0.12/13.9, 0.05
Cataract extraction	0.68 (0.66–0.69)	0.97/0.004/18.7, 0.03
Three-line improvement in visual acuity	0.55 (0.54–0.56)	0.78/0.089/9.2, 0.42
Severe vision loss	0.62 (0.60–0.64)	1.01/–0.001/6.9, 0.65	0.59 (0.57–0.62)	1.12/0.041/59.0, <0.0001
Photocoagulation or vitrectomy, or severe vision loss	0.63 (0.62–0.65)	0.97/0.004/11.5, 0.24
Neuropathy												
MNSI >2	0.60 (0.59–0.62)	1.01/–0.005/14.4, 0.11
Vibratory sensation loss	0.64 (0.63–0.66)	0.99/0.003/17.2, 0.05
Ankle jerk loss	0.57 (0.55–0.58)	0.96/0.019/5.0, 0.84
Pressure sensation loss	0.62 (0.61–0.64)	1.00/–0.0005/9.7, 0.37	0.69 (0.63, 0.74)	1.01/–0.002/1.0, 0.91
Macrovascular outcomes												

	Internal validation: RECODE			External validation: RECODE			Alternative risk equations: UKPDS OM2			Alternative risk equations: ACC/AHA PCEs		
	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*	Discrimination: C-statistic	Calibration: slope/intercept/ χ^2 , p value*	Discrimination: C-statistic	Calibration: slope/intercept/ χ^2 , p value*
Atherosclerotic cardiovascular disease (non-fatal or fatal myocardial infarction or stroke)	0.69 (0.67-0.71)	1.06/-0.005/13.7, 0.14	0.73 (0.71, 0.75)	1.13/-0.071/203.1, <0.0001	0.62 (0.60, 0.63) in ACCORD, 0.67 (0.64, 0.69) in Look AHEAD	0.56/0.043/602.6, <0.0001 in ACCORD, 0.53/-0.013/746.6, <0.0001 in Look AHEAD	0.61 (0.59, 0.63) in ACCORD, 0.66 (0.64, 0.69) in Look AHEAD	0.30/0.077/468.8, <0.0001 in ACCORD, 0.39/0.032/444.0, <0.0001 in Look AHEAD	0.61 (0.59, 0.63) in ACCORD, 0.66 (0.64, 0.69) in Look AHEAD	0.30/0.077/468.8, <0.0001 in ACCORD, 0.39/0.032/444.0, <0.0001 in Look AHEAD	0.61 (0.59, 0.63) in ACCORD, 0.66 (0.64, 0.69) in Look AHEAD	0.30/0.077/468.8, <0.0001 in ACCORD, 0.39/0.032/444.0, <0.0001 in Look AHEAD
Myocardial infarction (fatal or non-fatal)	0.69 (0.67-0.70)	1.00/0.0003/6.4, 0.70	0.71 (0.68, 0.74)	1.08/-0.016/17.0, 0.05	0.62 (0.59, 0.64) in ACCORD, 0.67 (0.65, 0.70) in Look AHEAD	0.80/0.106/47.6, <0.0001 in ACCORD, 0.94/-0.038/270.9, <0.0001 in Look AHEAD
Stroke (fatal or non-fatal)	0.70 (0.66-0.74)	1.16/-0.003/7.4, 0.38	0.67 (0.63, 0.71)	0.99/0.006/8.2, 0.22	0.61 (0.56, 0.66) in ACCORD, 0.63 (0.58, 0.68) in Look AHEAD	0.063/0.023/2275.6, <0.0001 in ACCORD, 0.279/0.007/659.5, <0.0001 in Look AHEAD
Congestive heart failure	0.75 (0.73-0.77)	1.01/-0.0004/3.1, 0.93	0.76 (0.73, 0.80)	1.13/-0.011/11.7, 0.07	0.61 (0.58, 0.65) in ACCORD, 0.61 (0.57, 0.65) in Look AHEAD	0.46/0.006/345.8, <0.0001 in ACCORD, 0.24/0.010/1246.5, <0.0001 in Look AHEAD
Cardiovascular mortality	0.74 (0.71-0.77)	0.96/0.001/7.8, 0.46	0.79 (0.74, 0.83)	1.00/-0.010/44.5, <0.0001
Other												
All-cause mortality	0.70 (0.68-0.72)	1.03/-0.002/14.7, 0.10	0.71 (0.68, 0.74)	1.10/-0.012/16.3, 0.06

UKPDS OM2=United Kingdom Prospective Diabetes Study Outcomes Model 2. ACC/AHA PCEs=American Heart Association Pooled Cohort Equations. MNSI=Michigan Neuropathy Screening Instrument. RECODE=Risk Equations for Complications Of type 2 Diabetes.

* p values <0.05 reflect larger difference between predicted and observed Kaplan-Meier event rates by the Greenwood-D'Agostino-Nam test (see figure 1 for calibration plots). 95% CIs are from 10-times cross-validations. Calibration slopes or intercepts are calculated between deciles of predicted and observed Kaplan-Meier event rates, with fewer centiles than deciles used if fewer than 5 events were observed per group to prevent unstable inferences per guidelines.²³