Predicting the Risk of Developing Type 2 diabetes in Chinese People who have Coronary Heart Disease and Impaired Glucose Tolerance

Shishi Xu[a,b], Charles A B Scott[a], Ruth L Coleman[a], Jaakko Tuomilehto[c,d], Rury R Holman[a]

[a] Diabetes Trials Unit, University of Oxford, OX3 7LJ, Oxford, UK

[b] Division of Endocrinology and Metabolism, West China Hospital, Sichuan University, 610041, Chengdu, China

[c] Department of Public Health Solutions, National Institute for Health and Welfare, 00271 Helsinki, Finland

[d] Department of Public Health, University of Helsinki, 00014 Helsinki, Finland

Corresponding author

Shishi Xu

Division of Endocrinology and Metabolism

West China Hospital, Sichuan University

Chengdu, China

Phone: +86 028 8542 2362

Email: shishixu@stu.scu.edu.cn

WORD COUNT: 2860 (max 3000)

Abstract: 200 (max 200)

Tables/Figures: (max 5)

References: 33 (max 50)

Original Article

Predicting the Risk of Developing Type 2 diabetes in Chinese People who have Coronary Heart Disease and Impaired Glucose Tolerance

Abstract

Aims: Robust diabetes risk estimates in Asian patients with impaired glucose tolerance (IGT) and coronary heart disease (CHD) are lacking. We developed a Chinese T2D risk calculator using Acarbose Cardiovascular Evaluation (ACE) trial data.

Methods: There were 3105 placebo-treated ACE participants with requisite data for model development. Clinically relevant variables, and those showing nominal univariate association with new-onset diabetes ($P<0.10$), were entered into BASIC (clinical variables only), EXTENDED (clinical variables plus routinely-available laboratory results) and FULL (all candidate variables) logistic regression models. External validation was performed using the Luzhou prospective cohort of 1088 Chinese patients with IGT.

Results: Over median 5.0 years, 493 (15.9%) ACE participants developed diabetes. Lower age, higher body mass index and use of corticosteroids or thiazide diuretics were associated with higher diabetes risk. C-statistics for the BASIC (using these variables), EXTENDED (adding male sex, fasting plasma glucose, 2-hour glucose and HbA_{1c}), and FULL models were 0.610, 0.757 and 0.761 respectively. The EXTENDED model predicted a lower 13.9% 5-year diabetes risk in the Luzhou cohort than observed (35.2%, 95% CI 31.3%–39.5%, C-statistic 0.643).

Conclusions: A risk prediction model utilizing routinely-available clinical variables can be used to estimate diabetes risk in Chinese people with CHD and IGT.

Keywords: Risk prediction, Type 2 diabetes mellitus, Impaired glucose tolerance, Coronary heart disease

1. Introduction

Primary prevention of type 2 diabetes (T2D) has become a major challenge, particularly for China which has the world's largest number of people with diabetes (116.4 million cases in 2019) equating to one in four of the world's adults with diabetes [1]. A major concern is that almost half of all adults in China – estimated to be 400 to 500 million people – have "pre-diabetes" [2, 3]. Identifying those individuals at highest risk of progressing to T2D to inform the need for early intervention is highly desirable, given that many trials have now shown that T2D can be delayed or prevented by lifestyle modification [4-6] or by pharmacological agents [7-10], with dietary intervention in the Da Qing study also reducing cardiovascular and all-cause mortality in the longer term [11].

To date, few diabetes risk prediction tools have been tailored specifically for people with impaired glucose tolerance (IGT). The two that were derived from large-scale IGT trials [12, 13] utilized predominantly Caucasian cohorts and have not been evaluated in Chinese populations which may well have different characteristics [14, 15]. Patients with established coronary heart disease (CHD) are of particular concern as they are at greater risk of developing T2D than the general population. Indeed, over half of all patients with CHD have been shown to have T2D or "pre-diabetes" in Western [16] and Chinese [17] cohorts attending cardiology outpatient clinics or admitted to hospital cardiovascular wards.

We have used data from the Acarbose Cardiovascular Evaluation (ACE) trial, which randomised Chinese patients with CHD and IGT to acarbose or placebo with prospective ascertainment of new-onset diabetes [10], to develop a T2D risk calculator for a Chinese population.

2. Methods

2.1 Population

The ACE trial design, baseline population characteristics and results have been published previously [10, 18, 19]. Briefly, ACE was a randomized, double-blind, placebo-controlled, event-driven, Phase IV superiority trial conducted in 176 outpatient clinics in tier 3 and tier 2 hospitals in China. Eligible participants were aged 50 years or older with established CHD (defined as previous myocardial infarction, previous unstable angina, or current stable angina), and IGT (confirmed by a 75g OGTT). Between March 2009 and October 2015, 6522 patients were enrolled and included in the intention-to-treat population, with 3272 assigned at random to acarbose and 3250 to placebo. As acarbose has been shown to reduce the risk of T2D [7, 10], only the placebo group were considered in this analysis.

2.2 Diagnosis of Diabetes

During the ACE trial, fasting plasma glucose (FPG) values were measured every 4 months and a 75g oral glucose tolerance test (OGTT) was performed annually. If either of these tests suggested diabetes, a confirmatory OGTT was done. Progression to diabetes was considered to have occurred if an elevated FPG (≥7.0 mmol/L) and/or 2-hour plasma glucose (2hPG) (≥11.1 mmol/L) value were recorded on two consecutive study visits, or if a diagnosis of diabetes was made by a non-trial physician and confirmed subsequently by an independent adjudication committee masked to study therapy allocation [10].

2.3 Risk factors for T2D

T2D risk factors considered included baseline demographic, clinical and laboratory variables, which were selected for their availability in routine clinical practice for patients with CHD, or were variables used commonly in previous diabetes risk scores. All values utilized were taken from the screening visit, (except for HbA_{1c} which was captured only at the randomization visit), as the ACE trial encouraged optimization of CHD risk reduction therapies during a

four-week run-in period prior to randomization [18].

2.4 Statistical analysis

ACE participants were categorized according to whether or not they had progressed to diabetes by their end of follow-up. Baseline characteristics for progressors and non-progressors were summarized by median and interquartile range for continuous variables, and numbers and percentages for categorical variables. Continuous variables were compared using Wilcoxon rank sum tests, and categorical variables were compared using Pearson chi-square tests. All statistical analyses were performed using SAS (version 9.4) [20] or R (version 3.4.3) [21].

2.4.1 Development of Five-Year T2D Risk Prediction Models

Three risk prediction models utilising ACE baseline data were developed using methodology aligned to the TRIPOD statement [22]: 1) a BASIC model, limited to demographic and readily-available clinical characteristics; 2) an EXTENDED model, which added routinely-available laboratory variables to the BASIC model; and 3) a FULL model, utilising all pre-specified ACE baseline variables.

The precise dates of diabetes onset are not available as ACE participants were only assessed every four months for possible development of diabetes. Accordingly, we used logistic regression rather than Cox models to build our risk calculator. Multivariable models were built with candidate variables included only if considered to be clinically relevant, or showing a nominally-significant ($P < 0.10$) univariate relationship with new-onset diabetes. Both forward and backward selection methods were used to eliminate non-significant variables. The principal criteria for selecting models were an improvement in the C-statistic (increment ≥ 0.005) and goodness of fit ($P > 0.05$ indicating good fit). Interactions between sex and other covariates were also tested.

2.4.2 Development of a Five-Year T2D Risk Score

For ease of use in everyday clinical practice a risk scoring algorithm was developed, based on the optimal prediction model with routinely available variables. Continuous variables were categorized by tertiles, quartiles or clinically-relevant cut-points. For example, body mass index (BMI) was categorized by the overweight (24 kg/m^2) and obesity (28 kg/m^2) threshold values recommended by the Chinese Guidelines [23]. All variables were then entered into a logistic regression analysis to compute β coefficients so that each could be assigned an estimated risk score. The sum of these scores was used to derive an estimated five-year risk of new-onset diabetes, which in turn could be classified as "Modest" (0-25%), "Moderate" (25-50%) or "High" (>50%), as described previously [12].

2.4.3 Internal validation

The performance of the three five-year diabetes risk prediction models and the five-year diabetes risk score in terms of discrimination (whether they distinguished between people who did/did not develop diabetes) and calibration (extent to which predicted probabilities agreed with the observed risk across groups of individuals) was evaluated in ACE placebo participants. The C-statistic was used to assess discrimination, which was classified as poor (0.5 to <0.6), acceptable (0.6 to 0.7), good (0.7 to <0.8), very good (0.8 to <0.9) or excellent (≥0.9). Calibration was estimated using the Hosmer-Lemeshow test, with a good fit indicated by a P value >0.05, and graphically by comparing the predicted probability against the observed probability across deciles of predicted risk

2.4.4 External validation

External validation of the diabetes prediction models and the risk score algorithm were

performed using data from the Luzhou survey, an independent prospective Chinese cohort followed between 2011 and 2016. This survey was part of the Risk Evaluation of cAncers in Chinese diabeTic Individuals: a lONgitudinal (REACTION) study, a multicenter prospective observational study investigating the association between diabetes and the risk of cancer among individuals with or without T2D in mainland China [24]. A total of 10,007 residents aged between 40 and 89 years were enrolled from five Luzhou communities using a multi-stage cluster random sampling method, of which 2,565 had IGT (with or without CHD). In the Luzhou survey cohort, incident diabetes was defined as any one, or a combination, of FPG ≥7.0 mmol/L, 2hPG ≥11.1 mmol/L or a self-report of a previous diagnosis of T2D made by a health care professional. In 2014 and 2016, 1,155 of these IGT participants were reviewed to determine if they had developed diabetes, of which 1,088 had complete data for the variables required to perform external validation of our final five-year diabetes risk prediction model and five-year diabetes risk score algorithm.

3. Results

3.1 Population

Of 3,250 ACE placebo-assigned participants, 3,105 (96%) had the requisite data for our analyses. Their baseline characteristics are listed in **Table 1**. Participants were predominantly Han (96.7%) and male (72.4%), with median age 63.0 years, BMI 25.4 kg/m^2, FPG 5.4 mmol/L, 2hPG 9.1 mmol/L and HbA_{1c} 5.9 % (41 mmol/mol). CHD history was categorized (not mutually exclusively) as myocardial infarction (44.0%), unstable angina (42.9%) or stable angina (22.1%), with most participants having a prior history of hypertension (65.5%), and taking statins (91.6%), beta-blockers (68.0%) and aspirin (94.4%).

Over median 5.0 years follow-up, 493 (15.9%) ACE participants progressed to diabetes. Compared with non-progressors, progressors were more likely to be younger, male, and current

smokers, with higher adiposity, glucose and lipid measures, and to have a prior history of hypertension or atrial fibrillation and to be taking thiazide diuretics or corticosteroids (**Table 1**).

3.2 Five-Year Diabetes Risk Prediction Model

The univariate associations with new-onset diabetes for 55 ACE candidate variables are summarized in **Supplemental Table S1**. Of these, the 33 that were clinically relevant or had nominally-significant associations with new-onset diabetes, were used for model development. Interactions between sex and other covariates did not achieve statistical significance (sex and age, P=0.055; sex and BMI, P=0.056) and were not included in the models.

In the BASIC model (**Figure 1A**), major risk factors for new-onset diabetes were lower age, higher BMI, and use of corticosteroids or thiazide diuretics. Modelling these variables yielded a C-statistic of 0.610 (**Table 2**) with a good fit (P=0.84) (**Supplemental Figure S1A**).

In the EXTENDED model (**Figure 1B**), the risk factors for new-onset diabetes added to the BASIC model were male sex and a higher baseline FPG, 2hPG and HbA_{1c}. Their inclusion increased the C-statistic from 0.610 to 0.757 (**Table 2**) and yielded a good fit (P=0.20) (**Supplemental Figure S1B**).

In the FULL model (**Figure 1C**), a new electrocardiographic abnormality was associated with a lower risk of diabetes incidence, but adding this variable only minimally improved the C-statistic from 0.757 to 0.761 (**Table 2**), also with a good fit (P=0.47) (**Supplemental Figure S1C**).

Accordingly, we elected to use the EXTENDED model as its performance was similar to the FULL model and because the variables required are more readily available.

3.3 Five-year T2D risk score algorithm

We constructed a risk scoring algorithm using the EXTENDED model equations which

produced risk scores ranging from 0 to 23 points (**Figure 2**). When the risk score was applied to ACE placebo participants it yielded a C-statistic of 0.754 with a good of fit (P=0.58) (**Supplemental Figure S1D**). The proportions of the ACE placebo population classified by the risk scoring algorithm as high, moderate and modest risk were 2.9%, 16.2% and 80.8% respectively, with 52.7%, 35.5% and 10.6% respectively developing diabetes. (**Figure 3**)

3.4 External validation

Baseline characteristics for the 1,088 Luzhou survey participants with IGT are listed in **Supplemental Table S2**. They were all of Han ethnicity, more often female (66.8%), with a median age of 60.0 years, and the majority (96.3%) had no history of CHD. Overall, 230 (21.1%, 95% CI 18.8% – 23.7%) participants progressed to T2D over a median of 3.0 years, with a higher incidence in those with a history of CHD (13 of 40, 32.5%). To obtain an estimated 5-year diabetes incidence for the Luzhou survey cohort we applied a linear extrapolation, multiplying the observed incidence by 5/3 to give 383 progressors (35.2%, 95% CI 31.3% – 39.5%).

The EXTENDED model predicted a 5-year diabetes incidence of 13.9% for the Luzhou survey cohort, substantially less than the projected proportion, with a C-statistic of 0.643 (0.602 – 0.685). The risk scoring algorithm predicted a 5-year diabetes incidence of 13.7%, with a C-statistic of 0.638 (0.595 – 0.680). The risk scoring algorithm classified 0.8%, 13.0% and 86.2% of Luzhou survey cohort participants as high, moderate and modest risk respectively, with 66.7%, 40.4% and 17.8% respectively progressing to diabetes. (**Figure 3**)

4. Discussion

We have developed a risk calculator, in compliance with the TRIPOD statement [22], that predicts the five-year likelihood that Chinese people with CHD and IGT will develop diabetes.

This calculator utilizes variables that are readily available in an outpatient setting, and can be used either in the form of equations or as a risk scoring algorithm.

The BASIC model, which just required information on age, BMI and whether or not a patient was taking corticosteroids or a thiazide diuretics, had only moderate discrimination. The EXTENDED model, that added sex, baseline FPG, 2hPG and HbA1c, showed good discrimination and a good fit. In the FULL model the only additional variable to be included was a new electrocardiographic abnormality but this did not materially alter the C-Statistic or goodness of fit. Accordingly, we recommend using the EXTENDED model equations, or the risk scoring algorithm derived from it which performed equally well.

Baseline variables shown here to be significantly associated with diabetes incidence are largely consistent with those in other diabetes prediction models. Higher BMI, male sex, use of corticosteroids, or thiazide diuretics and higher glucose levels are strongly associated with increased diabetes risk [25-31]. Among all variables, the three glucose measures (FPG, 2hPG and HbA_{1c}) were the most powerful predictors for future diabetes incidence. Our study found that lower age was associated with higher risk of diabetes, contrary to the trend in normal glucose tolerance populations, but consistent with previous IGT population findings [12]. Developing IGT at earlier age may indicate that this population carries a stronger genetic background related to diabetes, or more unfavourable environmental and lifestyle factors. We have no explanation for the finding that a new ECG abnormality (including clinically insignificant and significant abnormality) at baseline was significantly associated with a reduced risk of diabetes.

External validation of the EXTENDED model and the risk scoring algorithm in the Luzhou survey cohort of Chinese individuals with IGT [24] showed acceptable discrimination but substantially underestimated diabetes incidence in this population. This may be because only 3.7% of the Luzhou survey cohort had CHD as well as IGT. Also, compared with the ACE

study population, the Luzhou survey cohort population were younger, and had a lower proportion of males, current smokers, current alcohol users, prior hypertension, with very few takings statin, a lower BMI and a higher LDL-cholesterol. These differences between the two populations are largely consistent with the differences between people with and without CHD. Further external validation, and possibly calibration, will be required to maximise the model's predictive ability.

Previous European and Chinese surveys have shown a high prevalence of IGT among CHD patients, approximately 32% to 33% respectively [16, 17]. Those patients who progressed to diabetes have been reported to be at higher risk of adverse clinical outcomes with a greater mortality rate than those who do not develop diabetes [32, 33]. Using our T2D risk calculator could provide individualized risk estimates for Chinese patients with CHD and IGT, which in turn could enhance their awareness of T2D risk and prompt interventions that may prevent or delay the onset of T2D in this population.

Categorising estimated five-year diabetes risk as "Modest risk" (0%-25%), "Moderate" (>25%-50%) or "High" (>50%) provides a simple stratification that could help select an appropriate intensity of intervention. Possible recommendations might be routine lifestyle advice for those at modest risk, intensive lifestyle intervention for those at moderate risk and intensive lifestyle plus pharmacologic intervention for those at high risk of developing diabetes. Thus in clinical practice personalized diabetes risk estimates and risk classification could assist clinicians and patients when discussing the need to initiate primary prevention measures and the intensity required. External validation in Luzhou survey cohort showed that our model may have acceptable risk stratification ability in other IGT populations, which may help inform decision-making when considering primary prevention in this population.

Our study has several strengths. The ACE trial was a well-designed, large-scale, long-term secondary cardiovascular prevention study for which OGTT-confirmed new-onset diabetes was

a prespecified secondary outcome, making it an excellent resource to evaluate baseline predictors of the development of new onset diabetes. Also, as it was a multicenter study recruiting subjects from 176 sites across mainland China and Hongkong, participants were likely to be representative of the Chinese population as whole with CHD and IGT.

The study also has several limitations. Firstly, the ACE trial did not collect previously well-recognized predictors of diabetes such as a family history of diabetes, lifestyle and dietary factors (physical activity, fruit and vegetable consumption). Secondly, as the overwhelming majority of ACE population were taking the guideline recommend drugs for CHD management, it was not possible to test whether any of them individually had an impact on T2D risk, *e.g.* aspirin or statins. Thirdly, we could only validate our model at this time in a Chinese IGT population for whom most did not have CHD. Fourthly, ACE trial participants were recruited from top level Chinese hospitals (tier 2 and tier 3) meaning there was likely a selection bias to the population studied. Fifthly, all ACE trial subjects received appropriate lifestyle advice with respect to diet, exercise, and smoking and their cardiovascular therapy was optimized, which may have led to a lower diabetes incidence than might otherwise have been expected.

5. Conclusion

In Chinese people with CHD and IGT, lower age, male sex, obesity, use of corticosteroids, or thiazide diuretics as well as higher FPG, 2hPG and HbA_{1c} were major determinants of new-onset diabetes. A risk prediction model, utilizing routinely-available clinical variables and glycemic measures, can estimate T2D risk in Chinese people with CHD and IGT. With further calibration our simple risk calculator could inform decision-making when considering primary prevention T2D measures in this population.

Appendix A. Supplementary data

Supplementary material related to this article can be found in the online version.

Acknowledgements

The ACE trial was conducted by the University of Oxford Diabetes Trial Unit as an academically-independent study with funding from Bayer AG. R.R.H. is an Emeritus National Institute for Health Research Senior Investigator. We thank Qin Wan, for access to data from the Luzhou survey cohort.

Authors' contributions

S.X. performed the analysis and wrote the first draft of the manuscript.

C.S., R.L.C. and J.T. provided statistical advice.

R.R.H. designed the ACE trial, and reviewed the study design and results.

All authors reviewed the manuscript. R.R.H. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing interests

S.X., C.S. and R.L.C have no disclosures.

J.T. reports research grants and fees for consultancy, advisory board membership, or lectures from AstraZeneca, Bayer AG (related to the present study), Boehringer Ingelheim, Eli Lilly, Impeto Medical, Merck (Germany), Merck Sharp & Dohme, Sanofi Aventis, Novartis, Novo Nordisk, and Servier.

R.R.H. reports research support from AstraZeneca, Bayer and Merck Sharp & Dohme, and personal fees from Bayer, Intarcia, Merck Sharp & Dohme, Novartis and Novo Nordisk.

Ethics approval and consent to participate

The protocol of ACE trial was approved by the University of Oxford Tropical Research Ethics Committee, and by central or local ethics committees (as appropriate) at participating sites. All participants provided written informed consent.

REFERENCES

[1] International Diabetes Federation. *IDF Diabetes Atlas, 9th edn*. Brussels, Belgium: International Diabetes Federation, 2019. Available from http://www.diabetesatlas.org.

[2] Wang L, Gao P, Zhang M, *et al.* Prevalence and Ethnic Pattern of Diabetes and Prediabetes in China in 2013. *JAMA.* 2017; 317: 2515-2523.

[3] Xu Y, Wang L, He J, *et al.* Prevalence and control of diabetes in Chinese adults. *JAMA.* 2013; 310: 948-959.

[4] Pan XR, Li GW, Hu YH, *et al.* Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance. The Da Qing IGT and Diabetes Study. *Diabetes Care.* 1997; 20: 537-544.

[5] Knowler WC, Barrett-Connor E, Fowler SE, *et al.* Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med.* 2002; 346: 393-403.

[6] Tuomilehto J, Lindstrom J, Eriksson JG, *et al.* Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *N Engl J Med.* 2001; 344: 1343-1350.

[7] Chiasson JL, Josse RG, Gomis R, *et al.* Acarbose for prevention of type 2 diabetes mellitus: the STOP-NIDDM randomised trial. *Lancet.* 2002; 359: 2072-2077.

[8] Zinman B, Harris SB, Neuman J, *et al.* Low-dose combination therapy with rosiglitazone and metformin to prevent type 2 diabetes mellitus (CANOE trial): a double-blind randomised controlled study. *Lancet.* 2010; 376: 103-111.

[9] Armato JP, DeFronzo RA, Abdul-Ghani M, *et al.* Successful treatment of prediabetes in clinical practice using physiological assessment (STOP DIABETES). *Lancet Diabetes Endocrinol.* 2018; 6: 781-789.

[10] Holman RR, Coleman RL, Chan JCN, *et al.* Effects of acarbose on cardiovascular and diabetes outcomes in patients with coronary heart disease and impaired glucose tolerance

(ACE): a randomised, double-blind, placebo-controlled trial. *Lancet Diabetes Endocrinol.* 2017; 5: 877-886.

[11] Li G, Zhang P, Wang J, *et al.* Cardiovascular mortality, all-cause mortality, and diabetes incidence after lifestyle intervention for people with impaired glucose tolerance in the Da Qing Diabetes Prevention Study: a 23-year follow-up study. *Lancet Diabetes Endocrinol.* 2014; 2: 474-480.

[12] Bethel MA, Chacra AR, Deedwania P, *et al.* A novel risk classification paradigm for patients with impaired glucose tolerance and high cardiovascular risk. *Am J Cardiol.* 2013; 112: 231-237.

[13] Tuomilehto J, Lindstrom J, Hellmich M, *et al.* Development and validation of a risk-score model for subjects with impaired glucose tolerance for the assessment of the risk of type 2 diabetes mellitus-The STOP-NIDDM risk-score. *Diabetes Res Clin Pract.* 2010; 87: 267-274.

[14] Yoon KH, Lee JH, Kim JW, *et al.* Epidemic obesity and type 2 diabetes in Asia. *Lancet.* 2006; 368: 1681-1688.

[15] Ramachandran A, Ma RC, Snehalatha C. Diabetes in Asia. *Lancet.* 2010; 375: 408-418.

[16] Bartnik M, Ryden L, Ferrari R, *et al.* The prevalence of abnormal glucose regulation in patients with coronary artery disease across Europe. The Euro Heart Survey on diabetes and the heart. *Eur Heart J.* 2004; 25: 1880-1890.

[17] Hu DY, Pan CY, Yu JM, *et al.* The relationship between coronary artery disease and abnormal glucose regulation in China: the China Heart Survey. *Eur Heart J.* 2006; 27: 2573-2579.

[18] Holman RR, Bethel MA, Chan JC, *et al.* Rationale for and design of the Acarbose Cardiovascular Evaluation (ACE) trial. *Am Heart J.* 2014; 168: 23-29 e22.

[19] Theodorakis MJ, Coleman RL, Feng H, *et al.* Baseline characteristics and temporal

differences in Acarbose Cardiovascular Evaluation (ACE) trial participants. *Am Heart J.* 2018; 199: 170-175.

[20] Copyright © [2013] SAS Institute Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA.

[21] R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[22] Collins GS, Reitsma JB, Altman DG, *et al.* Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ.* 2015; 350: g7594.

[23] Chen C, Lu FC, Department of Disease Control Ministry of Health PRC. The guidelines for prevention and control of overweight and obesity in Chinese adults. *Biomed Environ Sci.* 2004; 17 Suppl: 1-36.

[24] Bi Y, Lu J, Wang W, *et al.* Cohort profile: risk evaluation of cancers in Chinese diabetic individuals: a longitudinal (REACTION) study. *J Diabetes.* 2014; 6: 147-157.

[25] Schmidt MI, Duncan BB, Bang H, *et al.* Identifying individuals at high risk for diabetes: The Atherosclerosis Risk in Communities study. *Diabetes Care.* 2005; 28: 2013-2018.

[26] Chen L, Magliano DJ, Balkau B, *et al.* AUSDRISK: an Australian Type 2 Diabetes Risk Assessment Tool based on demographic, lifestyle and simple anthropometric measures. *Med J Aust.* 2010; 192: 197-202.

[27] Rahman M, Simmons RK, Harding AH, *et al.* A simple risk score identifies individuals at high risk of developing Type 2 diabetes: a prospective cohort study. *Fam Pract.* 2008; 25: 191-196.

[28] Lindstrom J, Tuomilehto J. The diabetes risk score: a practical tool to predict type 2 diabetes risk. *Diabetes Care.* 2003; 26: 725-731.

[29] Wilson PW, Meigs JB, Sullivan L, *et al.* Prediction of incident diabetes mellitus in middle-aged adults: the Framingham Offspring Study. *Arch Intern Med.* 2007; 167: 1068-1074.

[30] Stern MP, Williams K, Haffner SM. Identification of persons at high risk for type 2 diabetes mellitus: do we need the oral glucose tolerance test? *Ann Intern Med.* 2002; 136: 575-581.

[31] Hippisley-Cox J, Coupland C. Development and validation of QDiabetes-2018 risk prediction algorithm to estimate future risk of type 2 diabetes: cohort study. *BMJ.* 2017; 359: j5019.

[32] Eitel I, Hintze S, de Waha S, *et al.* Prognostic impact of hyperglycemia in nondiabetic and diabetic patients with ST-elevation myocardial infarction: insights from contrast-enhanced magnetic resonance imaging. *Circ Cardiovasc Imaging.* 2012; 5: 708-718.

[33] Reinstadler SJ, Stiermaier T, Eitel C, *et al.* Relationship between diabetes and ischaemic injury among patients with revascularized ST-elevation myocardial infarction. *Diabetes Obes Metab.* 2017; 19: 1706-1713.

Figure Legends

Figure 1

Intercepts and ß coefficients for the BASIC (A), EXTENDED (B) and FULL (C) models, and the model equation (D) using these values to estimate the probability of developing diabetes within five years.

Figure 2

Risk scoring algorithm based on EXTENDED model equations for estimating five-year diabetes risk.

Figure 3

Diabetes Progression in three risk classes (using risk score algorithm) in ACE study and Luzhou survey cohort

Table 1. ACE participants baseline characteristics by progression or not to diabetes

	Overall	Non-Progressors	Progressors	P Value
	3105 (100%)	2612 (84.1%)	493 (15.9%)	…
Age (years)	63.00 (57.00, 70.00)	63.00 (58.00, 70.00)	61.00 (56.00, 68.00)	<0.0001
Sex (Male)	2249 (72.4%)	1872 (71.7%)	377 (76.5%)	0.033
Ethnicity (Han)	3001 (96.7%)	2523 (96.6%)	478 (97.0%)	0.78
Smoking status				0.034
Never Smoker	1364 (43.9%)	1167 (44.7%)	197 (40.0%)	…
Ex-smoker	1286 (41.4%)	1079 (41.3%)	207 (42.0%)	…
Current smoker	455 (14.7%)	366 (14.0%)	89 (18.1%)	…
Currently taking alcohol	368 (11.9%)	308 (11.8%)	60 (12.2%)	0.87
Weight (kg)	70.00 (63.00, 78.00)	70.00 (62.00, 77.00)	74.00 (65.00, 80.00)	<0.0001
Height (cm)	167.00 (160.00, 171.00)	167.00 (160.00, 170.00)	168.00 (161.00, 172.00)	0.0075
Waist circumference (cm)	91.00 (86.00, 97.00)	90.00 (85.00, 97.00)	93.00 (88.00, 100.00)	<0.0001
Hip circumference (cm)	100.00 (95.00, 105.00)	100.00 (94.00, 104.00)	100.00 (96.00, 106.00)	<0.0001
Body mass index (kg/m^2)	25.39 (23.44, 27.64)	25.20 (23.36, 27.41)	26.26 (24.44, 28.31)	<0.0001
Waist-to-hip ratio	0.92 (0.88, 0.95)	0.91 (0.88, 0.95)	0.92 (0.89, 0.96)	0.061
Waist to height ratio	0.55 (0.52, 0.59)	0.55 (0.51, 0.58)	0.56 (0.53, 0.59)	<0.0001
Systolic blood pressure (mmHg)	130.00 (120.00, 140.00)	130.00 (120.00, 140.00)	130.00 (120.00, 140.00)	0.12
Diastolic blood pressure (mmHg)	80.00 (70.00, 85.00)	80.00 (70.00, 85.00)	80.00 (70.00, 86.00)	0.060
Fasting plasma glucose (mmol/L)	5.42 (5.00, 5.90)	5.37 (4.97, 5.80)	5.82 (5.40, 6.30)	<0.0001
2-hour plasma glucose (mmol/L)	9.10 (8.36, 10.07)	9.00 (8.30, 9.92)	9.81 (8.86, 10.66)	<0.0001
HbA_{1c} (%)	5.90 (5.60, 6.30)	5.90 (5.56, 6.20)	6.20 (5.80, 6.50)	<0.0001
HbA_{1c} (mmol/mol)	41 (38, 45)	41 (37, 44)	44 (40, 48)	<0.0001
Total cholesterol (mmol/L)	4.05 (3.47, 4.79)	4.03 (3.47, 4.77)	4.16 (3.54, 4.86)	0.074
HDL-cholesterol (mmol/L)	1.12 (0.96, 1.33)	1.13 (0.96, 1.34)	1.08 (0.94, 1.26)	<0.0001
LDL-cholesterol (mmol/L)	2.23 (1.77, 2.85)	2.21 (1.76, 2.84)	2.30 (1.80, 2.90)	0.10
Triglycerides (mmol/L)	1.43 (1.04, 1.97)	1.39 (1.02, 1.93)	1.60 (1.16, 2.19)	<0.0001
Estimated glomerular filtration rate ($ml/min/1.73m^2$)	88.57 (74.79, 103.26)	88.35 (74.40, 103.47)	89.19 (77.21, 102.95)	0.19
Alanine amino transferase (U/L)	22.00 (16.42, 31.00)	22.00 (16.00, 31.00)	24.00 (18.00, 33.00)	<0.0001
Haemoglobin (g/dL)	14.20 (13.20, 15.10)	14.20 (13.10, 15.10)	14.50 (13.60, 15.30)	<0.0001
Mean red cell corpuscular volume (fL)	91.70 (89.00, 94.80)	91.80 (88.90, 94.90)	91.40 (89.00, 94.40)	0.11
White blood cell count ($x10^9/L$)	6.12 (5.20, 7.30)	6.10 (5.18, 7.24)	6.30 (5.37, 7.58)	0.0021
Platelets ($x10^9/L$)	195.00 (163.00, 232.00)	195.00 (163.00, 230.00)	197.00 (164.00, 238.00)	0.37

Cardiovascular medical history				
Hypertension	2034 (65.5%)	1691 (64.7%)	343 (69.6%)	0.043
Myocardial Infarction	1367 (44.0%)	1145 (43.8%)	222 (45.0%)	0.66
Unstable Angina	1333 (42.9%)	1126 (43.1%)	207 (42.0%)	0.68
Stable Angina	687 (22.1%)	579 (22.2%)	108 (21.9%)	0.95
Atrial fibrillation	122 (3.9%)	114 (4.4%)	8 (1.6%)	0.0060
Heart failure	116 (3.7%)	99 (3.8%)	17 (3.4%)	0.81
Revascularization	1469 (47.3%)	1239 (47.4%)	230 (46.7%)	0.79
Stroke or transient ischemic attack	212 (6.8%)	182 (7.0%)	30 (6.1%)	0.54
Concomitant medications				
Statin	2843 (91.6%)	2393 (91.6%)	450 (91.3%)	0.87
Any other lipid-lowering therapy	26 (0.8%)	22 (0.8%)	4 (0.8%)	1.00
Beta-blockers	2110 (68.0%)	1767 (67.6%)	343 (69.6%)	0.43
Angiotensin Receptor Blocker	803 (25.9%)	683 (26.1%)	120 (24.3%)	0.43
Angiotensin Converting Enzyme Inhibitors	1080 (34.8%)	912 (34.9%)	168 (34.1%)	0.76
Aldosterone antagonist	105 (3.4%)	86 (3.3%)	19 (3.9%)	0.62
Calcium channel blocker	963 (31.0%)	792 (30.3%)	171 (34.7%)	0.062
Thiazide diuretic	104 (3.3%)	79 (3.0%)	25 (5.1%)	0.029
Non-thiazide diuretic	83 (2.7%)	75 (2.9%)	8 (1.6%)	0.15
Any other antihypertensive therapy	80 (2.6%)	70 (2.7%)	10 (2.0%)	0.50
Aspirin	2930 (94.4%)	2461 (94.2%)	469 (95.1%)	0.48
Clopidogrel	1981 (63.8%)	1668 (63.9%)	313 (63.5%)	0.92
Other anti-platelet therapy	44 (1.4%)	41 (1.6%)	3 (0.6%)	0.15
Nitrates	1270 (40.9%)	1067 (40.8%)	203 (41.2%)	0.93
Other anti-anginal drugs	214 (6.9%)	179 (6.9%)	35 (7.1%)	0.92
Digitalis	57 (1.8%)	49 (1.9%)	8 (1.6%)	0.84
Antiarrhythmics	62 (2.0%)	56 (2.1%)	6 (1.2%)	0.24
Corticosteroids	30 (1.0%)	20 (0.8%)	10 (2.0%)	0.017
New electrocardiogram abnormality	1856 (59.8%)	1584 (60.6%)	272 (55.2%)	0.026

Data are shown as median and interquartile range or N (%)

Table 2. Five-year multivariable diabetes risk prediction models for Chinese people with CHD and IGT (N=3105, 493 events)

Model Variables	BASIC Model OR (95% CI)	P value	EXTENDED Model OR (95% CI)	P value	FULL Model OR (95% CI)	P value
Age (per 10 years increase)	0.80 (0.70-0.91)	0.00060	0.79 (0.69-0.91)	0.00096	0.80 (0.70-0.92)	0.0017
Male (vs female)	…	…	1.37 (1.08-1.76)	0.011	1.39 (1.09-1.78)	0.0085
Body mass index (per 1 kg/m^2)	1.10 (1.07-1.14)	<0.0001	1.07 (1.03-1.11)	<0.0001	1.07 (1.04-1.11)	<0.0001
Corticosteroid treatment (vs. none)	3.00 (1.32-6.41)	0.0057	2.95 (1.23-6.68)	0.012	2.82 (1.17-6.43)	0.016
Thiazide diuretic treatment (vs. none)	1.88 (1.15-2.95)	0.0084	1.71 (1.02-2.80)	0.036	1.71 (1.01-2.79)	0.039
Fasting plasma glucose (per 1 mmol/L)	…	…	2.09 (1.77-2.48)	<0.0001	2.10 (1.78-2.49)	<0.0001
2-hour plasma glucose (per 1 mmol/L)	…	…	1.62 (1.46-1.80)	<0.0001	1.63 (1.47-1.81)	<0.0001
HbA_{1c} (per 1 %)	…	…	1.81 (1.56-2.11)	<0.0001	1.85 (1.59-2.15)	<0.0001
Electrocardiogram abnormality (vs. normal)	…	…	…	…	0.70 (0.57-0.87)	0.00093
C-index (95% CI)	0.610 (0.583-0.637)		0.757 (0.735-0.780)		0.761 (0.738-0.784)	
Hosmer-Lemeshow goodness of fit P-value	0.84		0.20		0.47	

OR: Odds ratio, 95% CI: 95% confidence interval

Figure 1

(A)

BASIC Model	
β_n	x_n
-2.831	(Intercept)
-0.022	Age (years)
0.098	BMI (kg/m2)
1.095	Taking Corticosteroids (Yes=1, No=0)
0.617	Taking Thiazide diuretic (Yes=1, No=0)

(B)

EXTENDED Model	
β_n	x_n
-14.499	(Intercept)
-0.023	Age (years)
0.317	Sex (Male=1, Female =0)
0.068	BMI (kg/m2)
1.081	Taking Corticosteroids (Yes=1, No=0)
0.539	Taking Thiazide Diuretic (Yes=1, No=0)
0.739	FPG (mmol/L)
0.482	2hPG (mmol/L)
0.596	HbA1c (%)

(C)

FULL Model	
β_n	x_n
-14.619	(Intercept)
-0.022	Age (years)
0.328	Sex (Male=1, Female =0)
0.069	BMI (kg/m2)
1.036	Taking Corticosteroids (Yes=1, No=0)
0.534	Taking Thiazide diuretic (Yes=1, No=0)
0.744	FPG (mmol/L)
0.489	2hPG (mmol/L)
0.615	HbA1c (%)
-0.356	ECG abnormality (Yes=1, No=0)

(D)

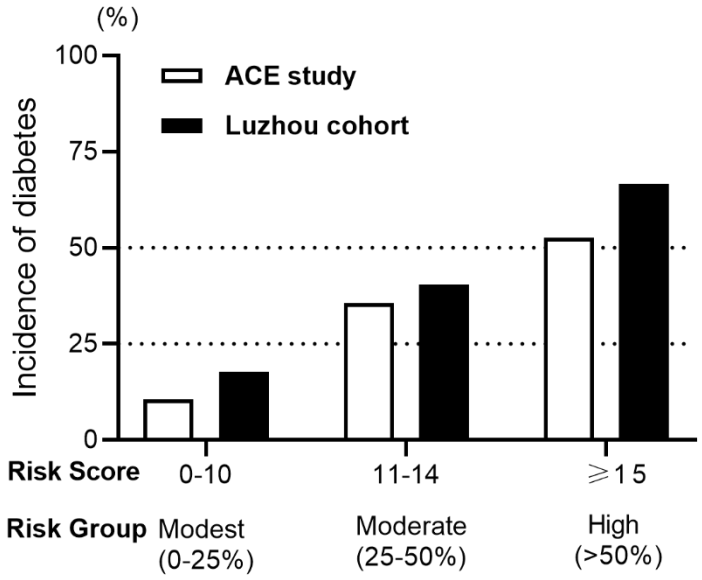
$$P = \frac{1}{1+e^{-\left(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n\right)}}$$

Figure 2

Step 1. Sum scores for each risk factor	
Risk Factor	**Points**
Age (years)	
≤60	2
60-80	1
≥80	0
Sex	
Female	0
Male	1
BMI (kg/m^2)	
<24	0
≥24 to <28	1
≥28	2
Taking corticosteroids	
No	0
Yes	4
Taking thiazide diuretics	
No	0
Yes	2
FPG (mmol/L)	
<5.6	0
≥5.6 to <6.1	2
≥6.1 to 7.0	4
2hPG (mmol/L)	
7.8 to <8.6	0
≥8.6 to <9.4	1
≥9.4 to <10.2	3
≥10.2 to 11.1	4
HbA_{1c} (%) (mmol/mol)	
<5.7 (<39)	0
≥5.7 to <6.5 (≥39 to <48)	2
≥6.5 (≥48)	4

Step 2. Read off 5-year diabetes risk corresponding to the total score	
Total Score	**5-year risk of diabetes**
0	1%
1	2%
2	3%
3	4%
4	5%
5	6%
6	8%
7	11%
8	14%
9	18%
10	23%
11	29%
12	35%
13	42%
14	50%
15	57%
16	65%
17	71%
18	77%
19	82%
20	86%
21	89%
22	92%
23	94%

Figure 3


(%)
100
75
50
25
0
Incidence of diabetes
ACE study
Luzhou cohort
Risk Score
0-10
11-14
⩾15
Risk Group
Modest
(0-25%)
Moderate
(25-50%)
High
(>50%)

Supplementary files

Supplemental Table S1. Univariate associations between 55 candidate variables and new-onset diabetes

	Unadjusted OR (95% CI)	P Value
BASIC Model		
Age (per 1 year) *	0.97(0.96-0.99)	<0.0001
Male (vs Female) *	1.28(1.03-1.61)	0.029
Han (vs Non-Han)	1.12(0.66-2.04)	0.68
Ex-smoker (vs Never smoker) *	1.14(0.92-1.41)	0.24
Current smoker (vs Never smoker) *	1.44(1.09-1.89)	0.0097
Current alcohol (vs non-Current alcohol) *	1.04(0.77-1.38)	0.81
Weight (per 1 cm) *	1.03(1.02-1.04)	<0.0001
Height (per 1 cm) *	1.01(1.00-1.03)	0.022
Body mass index (per 1 kg/m2) *	1.03(1.02-1.04)	<0.0001
Waist circumference (per 1 cm) *	1.03(1.02-1.04)	<0.0001
Hip circumference (per 1 cm) *	1.11(1.08-1.14)	<0.0001
Waist-to-hip ratio (per 1 cm/cm) *	3.88(0.75-19.86)	0.10
Waist to height ratio (per 1 cm/cm) *	32.11(5.39-190.62)	0.00014
SBP (per 1 mmHg) *	1.00(1.00-1.01)	0.13
DBP (per 1 mmHg) *	1.01(1.00-1.02)	0.017
History of hypertension (vs. none) *	1.25(1.01-1.54)	0.039
History of Myocardial Infarction (vs. none)	1.05(0.86-1.27)	0.62
History of Unstable Angina (vs. none)	0.96(0.79-1.16)	0.64
History of Stable Angina (vs. none)	0.98(0.78-1.24)	0.90
History of Atrial fibrillation (vs. none) *	0.36(0.16-0.70)	0.0058
History of previous heart failure (vs. none)	0.91(0.52-1.49)	0.71
History of revascularization(vs. none)	0.97(0.80-1.18)	0.75
History of stroke/transient ischemic attack (vs. none)	0.87(0.57-1.27)	0.48
Statin (vs. none) *	0.96(0.69-1.36)	0.80
Any other lipid-lowering therapy (vs. none)	0.96(0.28-2.53)	0.94
Beta-blockers (vs. none) *	1.09(0.89-1.35)	0.40
ARB (vs. none)	0.91(0.72-1.13)	0.40
ACEi (vs. none)	0.96(0.79-1.18)	0.72
Aldosterone antagonist (vs. none)	1.18(0.69-1.91)	0.53
CCB (vs. none) *	1.22(0.99-1.49)	0.055
Thiazide diuretic (vs. none) *	1.71(1.06-2.67)	0.022
Non-thiazide diuretic (vs. none)	0.56(0.25-1.09)	0.12
Any other antihypertensive therapy (vs. none)	0.75(0.36-1.40)	0.40
Aspirin (vs. none) *	1.20(0.79-1.91)	0.42
Clopidogrel(vs. none)	0.98(0.81-1.20)	0.88
Other anti-platelet therapy (vs. none)	0.38(0.09-1.06)	0.11
Digitalis (vs. none)	1.01(0.83-1.23)	0.89

Antiarrhythmics (vs. none)	1.04(0.70-1.49)	0.84
Nitrates(vs. none)	0.86(0.38-1.73)	0.70
Other anti-anginal drugs	0.56(0.22-1.21)	0.18
Steroids (vs. none) *	2.68(1.20-5.64)	0.011
EXTENDED Model additional variables		
Fasting plasma glucose (per 1 mmol/L) *	2.71(2.32-3.19)	<0.0001
2-hour plasma glucose (per 1 mmol/L) *	1.76(1.59-1.94)	<0.0001
HbA1c (per 1%) *	2.11(1.82-2.45)	<0.0001
Total-Cholesterol (per 1 mmol/L) *	1.07(1.01-1.15)	0.055
HDL-cholesterol (per 1 mmol/L) *	0.60(0.43-0.84)	0.0028
LDL-cholesterol (per 1 mmol/L) *	1.10(0.98-1.22)	0.097
Triglycerides (per 1 mmol/L) *	1.12(1.03-1.21)	0.0053
FULL Model additional variables		
eGFR (per 1 ml/min/1.73m2)	1.00(1.00-1.00)	0.99
Alanine amino transferase (per 1 U/L) *	1.01(1.00-1.01)	0.031
Haemoglobin (per 1 g/dL) *	1.14(1.06-1.21)	0.00012
Mean red cell corpuscular volume (per 1 fL)	0.99(0.97-1.00)	0.15
White blood cell count (per 1*10^9/L) *	1.08(1.02-1.14)	0.0052
Platelets (per 1*10^9/L)	1.00(1.00-1.00)	0.24
Electrocardiogram abnormality (vs. normality) *	0.80(0.66-0.97)	0.023

* 33 variables which based on clinical judgement were regarded as candidate predictors, or which showed nominal univariate association with new-onset diabetes (P<0.10)

Supplemental Table S2. Baseline characteristics of impaired glucose tolerance (IGT) participants in the Luzhou survey cohort and ACE population

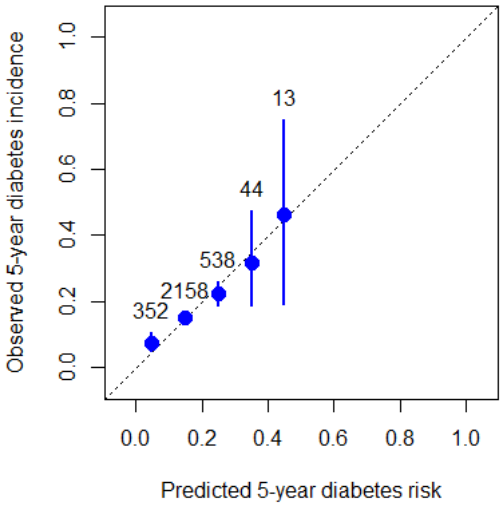
	ACE population (n=3105)	Luzhou survey cohort (n=1088)	P value
Age (years)	63.0 (57.0, 70.0)	60.0 (54.0, 66.0)	<0.0001
Sex			<0.0001
Female	856 (27.6%)	727 (66.8%)	
Male	2249 (72.4%)	361 (33.2%)	
Smoking status†			<0.0001
Never Smoker	1364 (43.9%)	895 (82.3%)	
Ex-smoker	1286 (41.4%)	52 (4.8%)	
Current smoker	455 (14.7%)	117 (10.8%)	
Currently taking alcohol†	368 (11.9%)	94 (8.6%)	0.024
Body mass index (kg/m^2)	25.4 (23.4, 27.6)	24.5 (22.5, 26.7)	<0.0001
Prior hypertension†	2034 (65.5%)	220 (20.2%)	<0.0001
Corticosteroids use	30 (1.0%)	0 (0%)	0.0023
Statin†	2843 (91.6%)	1 (0.1%)	<0.0001
Thiazide diuretic use	104 (3.3%)	9 (0.8%)	<0.0001
Fasting plasma glucose (mmol/L)	5.42 (5.00, 5.90)	5.51 (5.21, 5.87)	<0.0001
2-hour plasma glucose (mmol/L)	9.10 (8.36, 10.1)	8.89 (8.28, 9.75)	<0.0001
HbA_{1c} (%)	5.90 (5.60, 6.30)	6.00 (5.70, 6.20)	0.0059
HbA_{1c} (mmol/mol)	41 (38, 45)	42 (39, 44)	0.0059
Total cholesterol (mmol/L)	4.05 (3.47, 4.79)	4.64 (3.86, 5.35)	<0.0001
HDL-cholesterol (mmol/L)	1.12 (0.96, 1.33)	1.18 (0.98, 1.40)	<0.0001
LDL-cholesterol (mmol/L)	2.23 (1.77, 2.85)	2.56 (2.01, 3.13)	<0.0001
Triglycerides (mmol/L)	1.43 (1.04, 1.97)	1.43 (1.01, 2.13)	0.539
Prior cardiovascular disease†	3105 (100%)	48 (4.4%)	<0.0001
Prior coronary heart disease†	3105 (100%)	40 (3.7%)	<0.0001
Type 2 diabetes progressors	493 (15.9%)	230 (21.1%)	<0.0001

Data are shown as median and interquartile range or N (%)

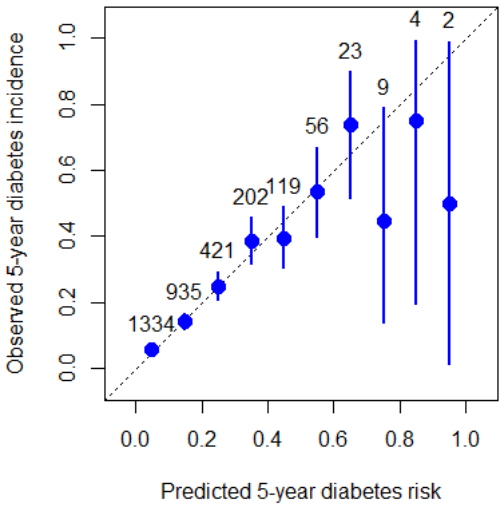
†These variables had missing values in Luzhou survey: smoking status (2.2%), currently taking alcohol (6.3%), prior hypertension (12.3%), statin (0.2%), cardiovascular disease (15.3%) and coronary heart disease (15.6%). Their analyses were based on complete data cases.

Supplemental Figure S1. Calibration plots of the observed 5-year diabetes incidence by deciles of the predicted 5-year risk of diabetes for the BASIC model (A), the EXTENDED model (B), the FULL model (C), and the diabetes risk score (D).

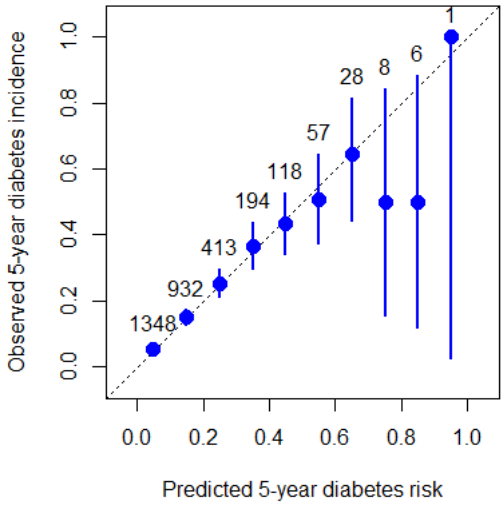
(A)



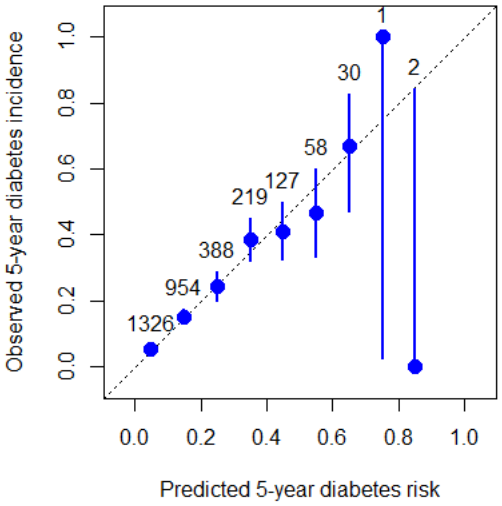
(B)



(C)



(D)



Declaration of Competing interests

S.X., C.S. and R.L.C have no disclosures.

J.T. reports research grants and fees for consultancy, advisory board membership, or lectures from AstraZeneca, Bayer AG (related to the present study), Boehringer Ingelheim, Eli Lilly, Impeto Medical, Merck (Germany), Merck Sharp & Dohme, Sanofi Aventis, Novartis, Novo Nordisk, and Servier.

R.R.H. reports research support from AstraZeneca, Bayer and Merck Sharp & Dohme, and personal fees from Bayer, Intarcia, Merck Sharp & Dohme, Novartis and Novo Nordisk.