A simple risk score identifies individuals at high risk of developing Type 2 diabetes: a prospective cohort study

Mushtaqur Rahmana, Rebecca K Simmonsb, Anne-Helen Hardings, Nicholas J Warehamb and Simon J Griffina


Background. Randomized trials have demonstrated that Type 2 diabetes is preventable among high-risk individuals. To date, such individuals have been identified through population screening using the oral glucose tolerance test.

Objective. To assess whether a risk score comprising only routinely collected non-biochemical parameters was effective in identifying those at risk of developing Type 2 diabetes.

Methods. Population-based prospective cohort (European Prospective Investigation of Cancer-Norfolk). Participants aged 40–79 recruited from UK general practices attended a health check between 1993 and 1998 (n = 25 639) and were followed for a mean of 5 years for diabetes incidence. The Cambridge Diabetes Risk Score was computed for 24 495 individuals with baseline data on age, sex, prescription of steroids and anti-hypertensive medication, family history of diabetes, body mass index and smoking status. We examined the incidence of diabetes across quintiles of the risk score and plotted a receiver operating characteristic (ROC) curve to assess discrimination.

Results. There were 323 new cases of diabetes, a cumulative incidence of 2.76/1000 person-years. Those in the top quintile of risk were 22 times more likely to develop diabetes than those in the bottom quintile (odds ratio 22.3; 95% CI: 11.0–45.4). In all, 54% of all clinically incident cases occurred in individuals in the top quintile of risk (risk score > 0.37). The area under the ROC was 74.5%.

Conclusion. The risk score is a simple, effective tool for the identification of those at risk of developing Type 2 diabetes. Such methods may be more feasible than mass population screening with biochemical tests in defining target populations for prevention programmes.

Keywords. Diabetes, general practice, incidence, risk score, screening.

Introduction

Type 2 diabetes is a growing public health problem and gives rise to significant morbidity, mortality and long-term financial costs. There is now strong evidence that diabetes is preventable through changes in key behaviours such as diet and physical activity. However, to date, the major diabetes primary prevention trials have targeted people with impaired glucose tolerance (IGT), which can only be identified through an oral glucose tolerance test. Mass population screening by such means is unlikely to be an acceptable and cost-effective way of identifying people who might benefit from health promotion interventions aimed at reducing the burden of disease associated with hyperglycaemia. Finding simpler, more pragmatic methods for identifying individuals at high risk of future progression to diabetes and who might benefit from targeted prevention is an important goal.

Various strategies have been developed to identify those with prevalent but undiagnosed diabetes, and some of these have also been assessed for their ability to predict the future development of diabetes. All have required that people in the target population...
complete a questionnaire or provide a blood sample. In contrast, the Cambridge Diabetes Risk Score only includes information on non-biochemical risk factors that are increasingly becoming routinely available to the family doctor. The risk score was independently developed and validated in population-based cross-sectional studies and demonstrated a sensitivity of 77% and a specificity of 72% for the detection of prevalent undiagnosed diabetes compared to the oral glucose tolerance test in a primary care population. In this report, we aimed to evaluate how well the score would predict the future clinical incidence of diabetes in a large, population-based cohort from the European Prospective Investigation of Cancer (EPIC)-Norfolk study.

Methods

Study design

EPIC-Norfolk is a prospective cohort study in which men and women aged 40–79 were recruited from general practices in the Norfolk region. Full details of the population are reported elsewhere. In brief, between 1993 and 1998, 77 630 individuals were invited to take part and 25 639 (33%) attended a baseline health examination. This included anthropometric and blood pressure measurements and completion of a general health questionnaire, with questions on personal and family history of disease, medication and lifestyle factors, including smoking habits. Participants were also asked whether a doctor had ever told them that they had any of the conditions contained in a list that included diabetes, heart attack and stroke. Non-fasting blood samples were taken, and starting in 1995 when funding became available, glycosylated haemoglobin (HbA1c) was measured on fresh ethylenediaminetetra-acetic acid blood samples using high-performance liquid chromatography in a single laboratory with a Diabetic Control and Complications Trial-defined normal range (HbA1c of 7% has the maximum specificity and sensitivity for detecting diabetes; this value has been shown to predict nephropathy and retinopathy as well as fasting blood glucose and 2-hour post-challenge blood glucose.

Results

Individuals with known diabetes at baseline (n = 845) and those with missing data for the variables in the risk score (n = 299) were excluded, leaving 24 495
participants for analysis. Baseline characteristics for the study population are shown in Table 1. Men and women were aged 58 years on average. Women had significantly lower baseline values for BMI and systolic and diastolic blood pressure than men. Women were also significantly more likely to have a family history of diabetes and to have never smoked.

This study accumulated a total of 117,027 person-years of follow-up, with a mean duration of 4.8 ± 1.3 (standard deviation) years. During this time, 323 people were diagnosed with diabetes, a cumulative incidence rate of 2.76 per 1000 person-years. Table 2 shows the association between individual risk score variables and the clinical incidence of diabetes. Age and BMI were significantly related to diabetes at follow-up. Men, those with a positive smoking history or family history of diabetes and those prescribed anti-hypertensive medication were also at increased risk of developing diabetes.

In total, 23% of individuals had a risk score > 0.37 (top quintile). These individuals were 22 times more likely to develop diabetes than those in the bottom quintile [odds ratio (OR) 22.3; 95% CI: 11.0–45.4] (Fig. 2). Over half (54%) of those with incident diabetes had a risk score in the top quintile (>0.37). The cumulative incidence of diabetes for this group was 7.5 per 1000 person-years. A threshold of 0.38 produced a likelihood ratio of a positive test of 2.73 (95% CI: 2.46–3.03) (Table 3). The area under the ROC was 74.5% (Fig. 3).

Sensitivity analyses in the subgroup of participants with HbA1c data demonstrated that the association between the top quintile of the Cambridge Risk Score and incident diabetes was attenuated but remained highly statistically significant (OR 10.4; 95% CI: 3.1–34.3).

Discussion

The Cambridge Diabetes Risk Score performs moderately well at predicting clinically incident Type 2 diabetes in the EPIC-Norfolk cohort, with an area under the ROC of 75%. Those in the top quintile of risk were 22 times more likely to develop diabetes than those in the bottom quintile. The Cambridge Risk Score appears to be a simple and effective tool for identification of those at risk of developing Type 2 diabetes using routinely collected non-biochemical information from patient records.

The Cambridge Risk Score has been shown to predict undiagnosed prevalent diabetes with an area under the ROC of 80% in a cross-sectional study, as well as undiagnosed hyperglycaemia metabolic syndrome and all-cause mortality. It also works well without information on family history and smoking, factors which used to be poorly recorded in general practice. In this study, we have shown that it can also identify individuals at risk of developing incident Type 2 diabetes in a large prospective cohort. Furthermore, we have demonstrated that the Cambridge Risk Score has good predictive value when tested in a population distinct from the one in which it was originally developed.

Methods for identifying individuals at future risk of diabetes have included biochemical tests, questionnaires and surveys on dietary habits and physical activity. However, such methods either require laboratory testing or the collection of new information, which will incur cost and may be associated with
false reassurance or anxiety. In addition, the use of arbitrary thresholds in single tests may not be as beneficial as the use of multiple risk factors in risk assessment. A previous diabetes risk score has shown the effectiveness of simple clinical parameters in the identification of those at risk of diabetes, but only drug-treated diabetes was used as the outcome measure in this study. We used patient records as well as self-report to identify individuals with diabetes at follow-up. Similarly, a German Diabetes Risk Score demonstrated an area under the ROC of 0.82–0.84 in different populations using information on age, waist circumference, height, history of hypertension, physical activity, smoking, and consumption of red meat, whole-grain bread, coffee, and alcohol, but this score required the collection of detailed information on physical activity and dietary behaviour. The Cambridge Diabetes Risk Score, which uses information routinely available in primary care records, could be used to help identify individuals or subgroups of the population who might benefit from further investigation, e.g. with an oral glucose tolerance test (OGTT) and more comprehensive risk assessment, or even direct preventive action.

The key issues for identification of a high-risk group for targeted prevention are the ease with which the group can be identified, the total number of individuals identified as being at high risk and the level of absolute risk in that group, which when combined with the efficacy of the preventive intervention impacts on the number needed to treat to prevent one incident case. Defining a high-risk group on the basis of biochemical results such as IGT clearly identifies individuals at high absolute risk of progression to diabetes, but the major problem is one of feasibility. By contrast, the risk score that we have used in this study is easy to employ and does not require the collection of new information. Using the approach described here, 20% of individuals could be defined as being at high risk of progression to diabetes and 54% of incident cases would come from that group. Those with scores > 0.37 (top quintile) were five times more likely to be diagnosed with diabetes than those with scores < 0.37 (bottom four quintiles). If an intervention similar to that used in the Diabetes Prevention Program were targeted at this group, then 48 individuals with a high-risk score would require a lifestyle intervention to prevent one case of diabetes. The relative merits of risk prediction and direct preventive action compared to the approach of labelling individuals as having a biochemical disorder with subsequent behavioural and pharmacological therapy need to be assessed. Similarly, a preventive strategy involving a risk score as part of systematic stepwise risk assessment would need formal evaluation, e.g. in a randomized trial.

**Strengths and limitations**

EPIC-Norfolk is a large population-based cohort, so the performance of the risk score was unlikely to be affected by spectrum bias, e.g. if an instrument is found to be effective in identifying a condition in one setting, it may not be as effective in another clinical situation where the spectrum of disease is different. Indeed, the Cambridge Risk Score has been found to be effective in predicting a range of related
conditions in different populations.\textsuperscript{17,18,25} High ascertainment of diabetes incidence was achieved\textsuperscript{14} and overall incidence was similar to figures reported in other UK studies looking at comparable populations.\textsuperscript{26}

EPIC-Norfolk is a predominantly Caucasian cohort and the sensitivity and specificity of the original Cambridge Risk Score have been shown to be reduced in predicting undiagnosed hyperglycaemia in Caribbean and South Asian populations living in the UK.\textsuperscript{25} Diabetes risk scores do not typically perform as well in populations in which they were not developed and the Cambridge Risk Score will need to be validated in other prospective cohorts. The accuracy of any risk score depends on the level of association of different risk factors with diabetes and the prevalence of risk factors, which will change over time and in different populations. As such, there is a need to reform, develop and update existing risk scores (as associations and beta coefficients change) to improve prediction of risk. Finally, the Cambridge Risk Score is based on information that might not be readily available from patient records in less-developed health care settings and may need to be modified accordingly. Similarly, the risk score might perform less well in routine practice due to less precise assessment of risk factors than occurred in this study.

The length of follow-up in the EPIC-Norfolk cohort was relatively short. Thus, while the risk score may be good at identifying individuals rapidly progressing to diabetes, it may miss those with a slower onset. Alternatively, longer follow-up might increase the ORs for the association between the risk score and the incidence of diabetes if more of those in the highest risk score category go on to develop diabetes.

In terms of our definition of incident diabetes, some selection bias may be present as people attending GP surgeries may have been more likely to be tested and, consequently, diagnosed with diabetes, e.g. if they were visibly overweight, or reported an unhealthy diet and/or low levels of physical activity. Thus, those individuals with diabetes who had low-risk scores may have been missed through a lack of testing. However, as we achieved 99% case ascertainment for all EPIC-Norfolk participants, whether or not they returned for the second health check, few clinically incident individuals with diabetes are likely to have been missed. In addition, incident diabetes was also defined biochemically using HbA\textsubscript{1c} > 7% in our sensitivity analyses, and while the association was attenuated due to the exclusion of those with prevalent undiagnosed diabetes at baseline, it remained significant, indicating that the effect of selection bias is unlikely to be large.

**Conclusion**

The Cambridge Diabetes Risk Score is a simple, effective tool for the identification of those at increased risk of future incident diabetes, as well as prevalent undiagnosed diabetes\textsuperscript{17,25,17} and mortality.\textsuperscript{19} As it comprises information routinely available in primary care, it may have a role in defining individuals and populations for programmes of testing, treatment and prevention.

**Acknowledgements**

We gratefully acknowledge the contribution of EPIC-Norfolk participants.
Declaration

Funding: British Heart Foundation; Cancer Research UK; Department of Health; Food Standards Agency; European Union (Europe Against Cancer Programme); Ministry of Agriculture, Fisheries and Food; Medical Research Council; Stroke Association; Wellcome Trust; World Health Organization; NHS Executive Eastern (MR).

Ethical approval: The Norwich District Ethics Committee approved this study and participants gave written consent prior to the investigations.

Conflicts of interest: None declared.

References