The Evaluation of Screening and Early Detection Strategies for Type 2 Diabetes and Impaired Glucose Tolerance (DETECT-2) update of the Finnish diabetes risk score for prediction of incident type 2 diabetes

M. Alssema · D. Vistisen · M. W. Heymans · G. Nijpels ·
C. Glümer · P. Z. Zimmet · J. E. Shaw · M. Eliasson ·
C. D. A. Stehouwer · A. G. Tabák · S. Colagiuri ·
K. Borch-Johnsen · J. M. Dekker ·
for the DETECT-2 collaboration

Abstract

Aims/hypothesis The Finnish diabetes risk questionnaire is a widely used, simple tool for identification of those at risk for drug-treated type 2 diabetes. We updated the risk questionnaire by using clinically diagnosed and screen-detected type 2 diabetes instead of drug-treated diabetes as an endpoint and by considering additional predictors.

Methods Data from 18,301 participants in studies of the Evaluation of Screening and Early Detection Strategies for Type 2 Diabetes and Impaired Glucose Tolerance (DETECT-2) project with baseline and follow-up information on oral glucose tolerance status were included. Incidence of type 2 diabetes within 5 years was used as the outcome variable. Improvement in discrimination and classification of the

M. Alssema (✉) · J. M. Dekker
Department of Epidemiology and Biostatistics and the EMGO Institute
for Health and Care Research, VU University Medical Center,
Amsterdam, the Netherlands
e-mail: m.alssema@vumc.nl

D. Vistisen · K. Borch-Johnsen
Steno Diabetes Center A/S,
Gentofte, Denmark

M. W. Heymans
Department of Methodology and Applied Biostatistics,
Institute of Health Sciences, VU University,
Amsterdam, the Netherlands

G. Nijpels
Department of General Practice and the EMGO Institute for Health
and Care Research, VU University Medical Center,
Amsterdam, the Netherlands

C. Glümer
Research Centre for Prevention and Health,
The Capital Region of Denmark,
Glostrup, Denmark

P. Z. Zimmet · J. E. Shaw
Baker IDI Heart and Diabetes Institute,
Melbourne, VIC, Australia

M. Eliasson
Department of Public Health and Clinical Medicine,
Umeå University, Umeå and Sundyby Hospital,
Luleå, Sweden

C. D. A. Stehouwer
Department of Internal Medicine and Cardiovascular Research
Institute Maastricht (CARIM),
Maastricht University Medical Centre,
Maastricht, the Netherlands

A. G. Tabák
Department of Epidemiology and Public Health,
University College London,
London, UK

A. G. Tabák
1st Department of Internal Medicine,
Semmelweis University Faculty of Medicine,
Budapest, Hungary

S. Colagiuri
Institute of Obesity, Nutrition and Exercise, University of Sydney,
Sydney, VIC, Australia

K. Borch-Johnsen
Faculty of Health Science, University of Aarhus,
Aarhus, Denmark

logistic regression model was assessed by the area under the receiver-operating characteristic (ROC) curve and by the net reclassification improvement. Internal validation was by bootstrapping techniques.

Results Of the 18,301 participants, 844 developed type 2 diabetes in a period of 5 years (4.6%). The Finnish risk score had an area under the ROC curve of 0.742 (95% CI 0.726–0.758). Re-estimation of the regression coefficients improved the area under the ROC curve to 0.766 (95% CI 0.750–0.783). Additional items such as male sex, smoking and family history of diabetes (parent, sibling or both) improved the area under the ROC curve and net reclassification. Bootstrapping showed good internal validity.

Conclusions/interpretation The predictive value of the original Finnish risk questionnaire could be improved by adding information on sex, smoking and family history of diabetes. The DETECT-2 update of the Finnish diabetes risk questionnaire is an adequate and robust predictor for future screen-detected and clinically diagnosed type 2 diabetes in Europid populations.

Keywords Epidemiology · Incidence · Prediction model · Primary prevention · Type 2 diabetes

Abbreviations

DETECT-2	Evaluation of Screening and Early Detection Strategies for Type 2 Diabetes and Impaired Glucose Tolerance
ROC	Receiver-operating characteristic

Introduction

Lifestyle and drug intervention in individuals at high-risk for type 2 diabetes mellitus have been shown to be able to prevent or delay the occurrence of type 2 diabetes [1, 2]. More specifically, lifestyle intervention was most effective among individuals with a high score on the Finnish diabetes risk tool that predicts the occurrence of type 2 diabetes [3]. Therefore, detection of high-risk individuals using a risk score may be relevant to identify those who may benefit most from interventions.

Next to existing risk scores to detect undiagnosed diabetes [4–6], several investigators have developed risk scores that predict incident diabetes [7–15]. Some of these risk scores include genetic information [7], laboratory measurements [8, 9] or systolic blood pressure [10] and are therefore applicable in clinical practice only. Others contain readily available information that is suitable for self-report questionnaires [11–15]. The questionnaire approach has been used as a first step in risk assessment or screening programmes because it does not require blood testing and is therefore simple, short and inexpensive [16]. The Finnish diabetes risk score is a widely applied risk score in Europid populations. The Finnish risk questionnaire has been shown to predict drug-treated type 2 diabetes in Finland [11] and type 2 diabetes as diagnosed by an OGTT in a Dutch population [17].

However, the Finnish risk questionnaire was developed to predict drug-treated type 2 diabetes without screening with an OGTT. It therefore seems appropriate to update it using a combined endpoint of clinically diagnosed (not only drug-treated) and screen-detected type 2 diabetes [18]. It is possible that the current risk questionnaire poorly identifies people whose diabetes would not be diagnosed and treated in clinical practice. For example, this may concern those with isolated elevation of post-load glucose levels, because an OGTT is usually not applied in general practice. It has been shown that there is limited overlap between fasting and post-load glucose levels; however, post-load glucose levels are strongly associated with cardiovascular risk [19]. Furthermore, the Finnish investigators had no information on family history of type 2 diabetes, hip circumference or smoking, which are potentially important predictors [20–22]. Updating an existing prediction tool instead of developing a new model may be a good alternative with the advantage of combining prior information on predictors with new data [23].

The aim of the present study was to update the Finnish diabetes risk questionnaire for the prediction of future clinically diagnosed and screen-detected type 2 diabetes and to consider additional predictors in an international population.

Methods

Study population The Evaluation of Screening and Early Detection Strategies for Type 2 Diabetes and Impaired Glucose Tolerance (DETECT-2) project is an international data-pooling collaboration addressing issues related to screening for type 2 diabetes [24]. Minimum requirements for DETECT-2 participation were (1) population-based surveys of large cohorts (*n*>500 people); (2) with all participants without known diabetes having a 75 g OGTT to classify type 2 diabetes according to the WHO 2006 criteria [18].

For the present analysis, all centres in the DETECT-2 collaboration with published incidence rates of type 2 diabetes were asked for baseline and follow-up information on oral glucose tolerance status. A total of six centres located in Europe (Hoorn Study [25], Inter99 [26], The Northern Sweden MONICA Study [27] and Whitehall-II [28]), Australia (AusDiab [29]) and Africa (Mauritius [30]), were included in the present analysis. Participation rates at the follow-up examination ranged from 50% in NSW-MONICA to 75% in Inter99.

Of the 23,226 persons with information on glucose tolerance status at follow up, we excluded 1,334 patients with known diabetes at baseline, 985 patients with screen-detected diabetes at baseline and 343 persons with missing or incomplete information on glucose status at baseline, leaving 20,564 participants for the present analysis. Participants provided informed consent and the studies have been approved by the relevant local ethics committees in accordance with the Declaration of Helsinki (revision 2000).

Incidence of type 2 diabetes This was defined as self-reported clinically diagnosed type 2 diabetes and/or screen-detected type 2 diabetes at the follow-up examination as defined by standard 75 g OGTT according to the WHO 2006 criteria [18]. To reach comparable follow-up periods for each centre, the follow-up period for each individual was truncated at 5 years after baseline and the assumption was made that the date of type 2 diabetes diagnosis was in the middle of the follow-up period. Hence, incident cases are all persons who developed type 2 diabetes within 5 years after the baseline examination.

The Finnish diabetes risk score in DETECT-2 The Finnish diabetes risk score predicts the incidence of drug-treated type 2 diabetes within 10 years. The variables of the Finnish diabetes risk score are: age (in categories <45; 45–54; 55–64 and ≥65 years), BMI (in categories ≤25; 25–29.99 and ≥30 kg/m^2 and higher), waist circumference (in categories for men: <94; 94–101.99 and ≥102 cm and for women: <80; 80–87.99 and ≥88 cm), use of blood pressure medication and history of high blood glucose [11]. Family history of diabetes was proposed as an item of the original Finnish risk score but the scores for parent or sibling diabetes were based on literature reports, and no data were available to verify the attributed score [11].

In the present analysis, if information on the use of blood pressure medication was missing then we set this variable to 'no'. Information on history of high blood glucose was limited to the history of gestational diabetes. Multiple imputations by chained equations were applied to eliminate missing values on age (0.03%), BMI (0.4%), waist circumference (0.6%) and history of gestational diabetes (missing in four of the six centres=65% of the study population), resulting in five imputed datasets. Regression coefficients and standard errors of the predictors of type 2 diabetes in the five imputed datasets were comparable to each other and to the pooled estimates after applying Rubin's rules, which means that imputation variation was low. Therefore, we used one randomly chosen imputed dataset for further analysis.

Updating The Finnish diabetes risk scores were re-estimated by multiple logistic regression analysis with all variables of the Finnish model.

Thereafter, relevant predictors (identified by a literature search) were added one by one to the multiple logistic regression model. These included sex, smoking, hip circumference (in sex-specific tertiles), use of lipid-lowering medication and family history of diabetes (parent or sibling or both). Finally, in additional analyses, we assessed an item of the Rose questionnaire [31]: reported 'shortness of breath when walking with people of the same age' as a potential predictor that was available in two centres. Except for potential interactions with study region, interactions were not considered because we aimed to develop a risk score that was internationally applicable and adding an interaction term would hamper the practical feasibility of the model. We investigated whether these variables improved the discriminative value of the model. The discriminative value was assessed using the area under the receiver-operating characteristic (ROC) curves, plotting sensitivity against 1 − specificity. The area under the ROC curve is a global summary statistic of the discriminative value of a model, describing the probability that the score is higher in an individual developing type 2 diabetes than in an individual not developing type 2 diabetes. The difference between models in area under the ROC curve was tested by a method described by Hanley and McNeil [32]. To test model goodness-of-fit we performed the Hosmer and Lemeshow test [33]. We also calculated the net reclassification improvement of the models that included additional predictors compared with the model without the additional predictors according to the method of Pencina et al. [34]. For that procedure, three risk categories were chosen a priori: low risk (<10%), intermediate risk (10–20%) and high risk (>20%).

Sensitivity analyses In sensitivity analyses, we assessed the ability of the model to detect undiagnosed diabetes and to predict future clinically diagnosed (i.e. without screen-detected cases) diabetes.

Validation and comparison with existing risk scores It is well-known that the apparent performance of a model is better in the derivation dataset than in another dataset, which is called overfitting. To estimate the amount of overfitting in the regression coefficients and the area under the ROC curves, bootstrapping techniques were used [35]. These bootstrap-adjusted measures represent the values that can be expected when the model is applied to future similar populations. Therefore, 200 bootstrap samples of equal size to the original dataset were drawn from the original dataset. The amount of overfitting in regression coefficients can be estimated by the slope index. The slope is calculated with the logistic regression model and the linear predictor scores as the only determinant. The linear predictor scores in each bootstrap and original sample are calculated by multiplying the model's regression coefficients by the predictor values for each patient and summing them (including the intercept).

Ideally this slope is 1, which means that the model's predicted and observed probabilities agree over the whole range of predictions. The difference in slope between the bootstrap and original sample can then be used as a shrinkage factor by multiplying the slope with each regression coefficient in the original model [23]. Thereafter, these bootstrap-adjusted regression coefficients were transformed into scores, i.e. multiplied by 4 and rounded off for simplicity to make the score values comparable to those of the Finnish risk score. For each individual, the total risk score can then be calculated by adding up the scores of each separate item. The performance of the updated model was compared with existing risk models for prevalent and incident diabetes for which (almost complete) information was available. Analyses were performed in SPSS 15.0.0 (SPSS, Chicago, IL, USA) and bootstrapping was performed in R 2.9.0 for Windows.

Results

Characteristics of the study population for each centre are given in Table 1. Because of interactions between almost all risk predictors and the study region Mauritius, we excluded

this study cohort from further analyses. In total, the number of persons who developed type 2 diabetes during the 5 year follow up was 844 of 18,301 (4.6%). Incidence rate varied from 4.6 to 20.6 per 1,000 person-years and was highest among the older Hoorn study participants.

The Finnish diabetes risk score had an area under the ROC curve of 0.742 (95% CI 0.726–0.758) in the whole study population, ranging from 0.63 to 0.78 among centres. Re-estimating the coefficients of the original Finnish score showed that all variables of the Finnish risk model were independent and significant predictors in the multiple regression model (Table 2). Also, re-estimating the regression coefficients resulted in an improvement in the area under the ROC curve (0.766 [95% CI 0.750–0.783]) compared with the original Finnish model ($p<0.05$).

Information on sex, smoking and hip circumference was added one by one to the model. Male sex and smoking were both significant predictors in the multiple model, contributing significantly to the area under the ROC curve ($p<0.05$), compared with the re-estimated Finnish model. However, hip circumference did not improve the area under the ROC curve significantly.

Of those variables that were available in fewer than five centres, a family history of diabetes was a significant

Table 1 Characteristics of the study populations

Characteristic	AusDiab	Inter99	Hoorn	NSW-MONICA	Whitehall II
Baseline (year)	1999–2000	1999–2001	1989–1991	1986–1999	1991–1993
N	5,660	4,510	1,353	1,544	5,234
Age (years)	51.2 (12.7)	46.3 (7.8)	60.3 (6.9)	47.6 (12.4)	49.1 (5.9)
Sex (% male)	44.4	49.9	45.6	47.7	71.6
BMI (kg/m^2)	26.6 (4.6)	26.0 (4.3)	26.2 (3.1)	25.4 (3.8)	25.2 (3.6)
Waist circumference (cm)	89.6 (13.3)	85.9 (12.7)	89.6 (10.2)	86.7 (11.6)	83.7 (11.3)
Hip circumference (cm)	104.3 (9.4)	100.5 (9.3)	101.6 (6.3)	99.0 (7.3)	96.9 (7.1)
Former smoking (%)	29.1	26.7	34.4	26.4	35.4
Current smoking (%)	11.4	34.6	30.1	17.8	11.4
Use of anti-hypertensives (%)	12.8	6.1	15.8	8.3	6.5
Use of lipid-lowering treatment (%)	6.9	1.0	1.4	0.6	0.8
History of gestational diabetes (%)	1.8	2.2 [a]	0.9	1.4[a]	0.5 [a]
Parent only with diabetes (%)	17.1	12.7	12.2	–	–
Sibling only with diabetes (%)	–	2.5	4.4	–	–
Parent and sibling with diabetes (%)	–	1.6	2.5	–	–
Shortness of breath when walking (%)	–	1.6	5.6	–	–
Follow-up					
Period at risk (years)	4.9 (0.6)	4.9 (0.6)	4.8 (0.6)	5.0 (0.3)	5.0 (0.3)
5-year incidence, n/N[b] (%)	335/5,660 (5.9)	207/4,510 (4.6)	134/1,353 (9.9)	48/1,544 (3.1)	120/5,234 (2.3)
Incidence rate per 1,000 person-years[c]	12.1	9.4	20.6	6.2	4.6

[a] By multiple imputation

[b] N is the number of participants in the study, n is the number of incident cases

[c] Estimated incidence rate in cases per 1,000 person-years was calculated by assumption that the date of type 2 diabetes diagnosis was in the middle of the follow-up period

Table 2 Multiple logistic regression models predicting 5 year incidence of type 2 diabetes in the DETECT-2 cohort ($n=18{,}301$)

Variable	Re-estimated Finnish risk model		Extended model		Extended model including family history[a]		Range in coefficients among centres[b]	Score points[c]
	Regression coefficient	OR (95% CI)	Regression coefficient	OR (95% CI)	Regression coefficient	OR (95% CI)		
Age (years)								
<45	0	1	0	1	0	1	–	0
45–55	0.248	1.28 (1.05–1.57)	0.267	1.31 (1.06–1.61)	0.280	1.32 (1.05–1.67)	–0.130 to 0.513	1
55–65	0.578	1.78 (1.44–2.20)	0.628	1.87 (1.51–2.33)	0.572	1.77 (1.39–2.27)	0.304–1.177	2
≥65	0.860	2.36 (1.82–3.07)	0.953	2.59 (1.99–3.39)	0.835	2.30 (1.73–3.07)	0.472–1.808	4
BMI (kg/m^2)								
<25	0	1	0	1	0	1	–	0
25–29.99	0.329	1.39 (1.12–1.73)	0.261	1.30 (1.04–1.63)	0.180	1.20 (0.93–1.55)	–0.015 to 0.703	1
≥30	0.508	1.66 (1.25–2.22)	0.485	1.62 (1.21–2.18)	0.354	1.43 (1.02–1.99)	–0.461 to 1.106	2
Waist circumference (cm)								
Female <80; male <94	0	1	0	1	0	1	–	0
Female 80–87.99; male 94–101.99	0.355	1.43 (1.14–1.79)	0.397	1.49 (1.18–1.88)	0.201	1.22 (0.94–1.60)	0.098–0.552	2
Female ≥88; male ≥102	1.050	2.86 (2.23–3.66)	1.121	3.07 (2.37–3.98)	0.958	2.61 (1.94–3.50)	0.556–1.628	4
Use of anti-hypertensives	0.698	2.01 (1.67–2.43)	0.727	2.07 (1.71–2.50)	0.749	2.12 (1.71–2.61)	0.368–0.936	3
History of gestational diabetes	3.101	22.23 (16.82–29.37)	3.140	23.10 (17.30–30.83)	2.835	17.02 (12.32–23.54)	–0.101 to 5.491	12
Sex (male vs female)	–	–	0.259	1.30 (1.11–1.51)	0.357	1.43 (1.20–1.70)	0.033–0.705	1
Smoking								
Former smoking	–	–	0.149	1.16 (0.98–1.38)	0.180	1.20 (0.98–1.46)	–0.027 to 0.159	1
Current smoking	–	–	0.563	1.76 (1.45–2.12)	0.521	1.68 (1.37–2.08)	–0.285 to 0.575	2
Family history of diabetes[d]								
Parent only with diabetes	–	–	–	–	0.648	1.91 (1.56–2.33)	0.612–0.874	3 [e]
Sibling only with diabetes	–	–	–	–	0.759	2.14 (1.28–3.57)	0.211–1.042	3 [e]
Parent and sibling with diabetes	–	–	–	–	0.926	2.53 (1.38–4.62)	0.764–1.061	4 [e]

[a] This extended model is estimated in the three centres ($n=11{,}523$) that had information on family history of diabetes available

[b] Five study centres for all variables except for family history of diabetes, for which the range among three centres that had this information available is shown

[c] Score points are calculated from the extended model as: regression coefficient × slope index (=0.974) × 4, and rounded off for simplicity

[d] The score for this item is 4 at maximum, someone can have no family members or a parent only or a sibling only or both a parent and a sibling with diabetes

[e] The score points for family history only were estimated in three centres that had this information available

predictor in the multiple model (Table 2) and contributed to the area under the ROC curve on top of the re-estimated model including sex and smoking ($p<0.05$). In contrast, information on the use of lipid-lowering medication did not statistically significantly change the area under the ROC curve. The full, extended model included three additional variables to the Finnish risk model; sex, smoking and family history of diabetes. In the three centres that had information on all variables, the area under the ROC curve improved from 0.754 (95% CI 0.735–0.773) to 0.764 (95% CI 0.746–0.783, $p<0.05$) ranging from 0.70 to 0.78 among centres (Fig. 1). Some heterogeneity between study region and risk predictors was found, as observed by the range in coefficients among centres, but this could not be attributed to one specific study centre (Table 2). As indicated by non-

overlapping confidence intervals, the regression coefficient for age (category 45–55 years) was lower in AUSDIAB, the coefficient for BMI (≥30 kg/m^2) was lower in Hoorn and the coefficient for history of gestational diabetes was higher in AUSDIAB. The Hosmer and Lemeshow test results showed good calibration ($\chi^2=10.0$, $p=0.27$, Fig. 2).

Reclassification improved by 7.3% among persons who developed diabetes and worsened by 1.0% among persons who did not develop diabetes, so net reclassification was 6.3% ($p<0.001$) compared with the re-estimated Finnish model.

Sensitivity analyses Because the present risk model was developed as a risk questionnaire, which is often used as a first step in risk assessment, prevalent undiagnosed type 2
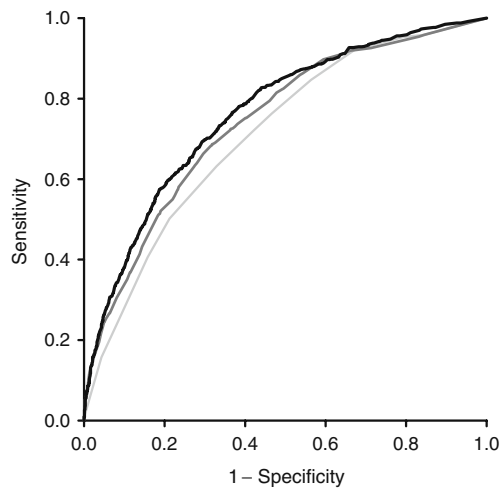
Fig. 1 ROC curve of the extended model (AUC 0.764 [95% CI 0.746–0.783]) compared with the re-estimated model (AUC 0.742 [95% CI 0.723–0.762]) and the original Finnish model (AUC 0.712 [95% CI 0.692–0.731]) in three centres (*n*=11,523) with information on the extended model including family history. Light grey line, original Finish risk score; medium grey line, re-estimated score, black line, extended score

diabetes may remain undiagnosed. Therefore, we assessed to what extent the present developed model is able to detect prevalent, screen-detected type 2 diabetes. Prevalent undiagnosed type 2 diabetes was equally well predicted as incident type 2 diabetes by the present model (area under the ROC curve 0.765 [95% CI 0.746–0.783] as compared with 0.764 [95% CI 0.746–0.783] for incident type 2 diabetes).

The performance of the updated risk model to predict future clinically diagnosed cases only was higher than if clinically diagnosed and screen-detected patients were predicted together (area under the ROC curve 0.787 [95% CI 0.760–0.814] for clinically diagnosed only compared with 0.764 [95% CI 0.746–0.783]).
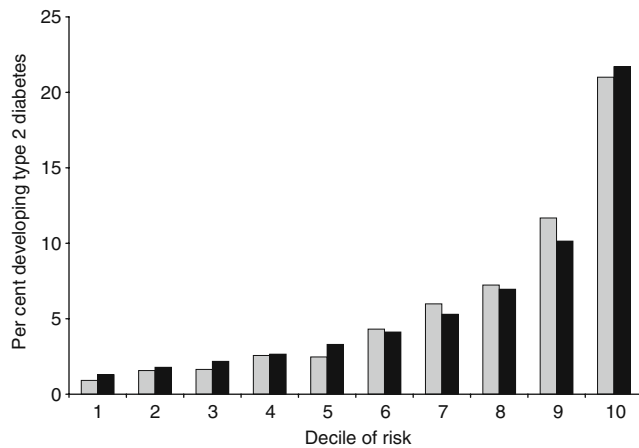
Validation and comparison with existing risk scores Comparison of bootstrap-adjusted areas and the original areas under the ROC curve showed only a marginal difference. The area under the ROC curve of the extended model decreased from 0.764 (95% CI 0.746–0.783) to 0.759 (95% CI 0.739–0.779). Bootstrap-adjusted regression coefficients were also marginally smaller than the original coefficients (slope index 0.974, Table 2).

The original Finnish risk score and the updated score performed better than existing risk scores for prevalent or incident type 2 diabetes. The area under the ROC curve of the Framingham Offspring Study [13] was 0.658 (95% CI 0.637–0.679), of the Aekplakorn et al. risk score [14] 0.654 (95% CI 0.633–0.675), of Cambridge risk score [4] without information on use of steroids 0.669 (95% CI 0.647–0.691) and of the Danish diabetes risk score [5] without information on physical activity 0.665 (95% CI 0.643–0.687).

Cut-off points Test characteristics for a range of cut-off values are presented in Table 3. At a score of 7 or higher, the sum of sensitivity and specificity was maximised. Of the total population, 40% had a score of 7 or higher. A score of 7 or higher captures 76% of the cases who will develop type 2 diabetes (sensitivity). Furthermore, 63% of the persons who do not develop type 2 diabetes had a score lower than 7 (specificity).

Additional analyses Information on self-reported shortness of breath when walking was evaluated as a risk predictor in two centres (Inter99 and Hoorn). This item was a significant predictor in the multiple model but adding information on this item to the extended model did not improve discrimination (*p*>0.05) or net reclassification (*p*>0.05).

Fig. 2 Observed and expected risk for developing type 2 diabetes (%) in deciles of expected risk according to the extended model in three centres (*n*=11,523) with information on the extended model including family history; grey bars, observed risk; black bars, expected risk

Table 3 Sensitivity, specificity and positive predictive value for various cut-off values in a subsample of three centres with information on the extended model including family history

Score	Percentage of the population	Sensitivity (%)	Specificity (%)	Positive predictive value (%)
4 or higher	68	91	33	8
5 or higher	58	88	44	9
6 or higher	49	84	54	10
7 or higher	40	76	63	11
8 or higher	30	66	72	13
9 or higher	22	57	80	15
10 or higher	16	48	86	17
11 or higher	11	37	90	19
12 or higher	8	31	94	23
13 or higher	5	23	96	26

Discussion

The present dataset is derived from a large data pooling collaboration containing five cohort studies with information on glucose tolerance status based on an OGTT. This provides an excellent opportunity to assess screen-detected next to clinically diagnosed incident diabetes and as such to validate and update the widely used Finnish diabetes risk score.

We showed that the Finnish score was already a good predictor for clinically diagnosed and screen-detected type 2 diabetes. Also, adding easily accessible information on sex, smoking and family history of diabetes improved the test's predictive power.

The present study differs from the Finnish study with respect to both population characteristics and the way that type 2 diabetes was diagnosed [11], for this reason we cannot simply ascribe the observed differences to the inclusion of screen-detected incident type 2 diabetes in our endpoint. Nevertheless, the present definition including screen-detected and clinically diagnosed type 2 diabetes involves more patients than the drug-treated patients that were predicted by the original Finnish score. Drug-treated patients probably have more advanced diabetes than screen-detected patients or patients treated by lifestyle only. This may at least in part, explain the lower discrimination of the present risk score as compared with the original Finnish risk score in the Finnish study in which it was developed. Indeed, the present risk score also predicted clinically diagnosed patients better than screen-detected patients. We think however that use of the present definition of type 2 diabetes is an improvement because it will also detect future diabetes patients who are less likely to be detected by general practice, for example because of isolated elevation of post-load glucose levels. This will lead to the identification of a population that may benefit from preventive interventions given their high cardiovascular risk [19, 36].

One advantage of updating an existing model instead of developing a new one is that it combines earlier information with new information. Updating methods vary from re-calibration to extensive revision: re-estimation of regression coefficients and consideration of new predictors [23]. As we also used a different definition of the endpoint, and our dataset had sufficient power, we chose to extensively update this model.

Recent assessment, especially of new biomarkers for cardiovascular disease, have shown that measures of discrimination, calibration and reclassification are, when used in isolation, not sufficient to decide whether one model is superior to another [34]. Therefore, we used more than one of these measures to decide whether or not to include additional risk factors.

In the present analysis, we excluded patients with prevalent known or screen-detected type 2 diabetes.

However, because the questionnaire usually serves as a first step in risk assessment, prevalent type 2 diabetes may remain undiagnosed. Therefore, the predictive value for prevalent undiagnosed type 2 diabetes was also evaluated, which showed equally good performance of the model for prevalent undiagnosed and incident type 2 diabetes.

We evaluated the concise Finnish risk model instead of the full risk model, which also includes physical activity and daily consumption of vegetables, fruit or berries [11]. The main reason for this is the lack of information in the present dataset. Furthermore, it is uncertain whether asking one isolated question about diet or physical activity is predictive for type 2 diabetes, taking into account that for the development of the original risk model, these items were obtained from an extensive questionnaire. Finally, even when obtained from an extensive questionnaire, the items on physical activity and diet only minimally improved prediction [11].

We previously reported that the proposed score points associated with the presence of one family member with diabetes were relatively high (5 compared with the present 3 for parent or sibling) [17]. We also found that someone with a sibling with diabetes is at similar risk to someone with a parent with diabetes. Data on this are scarce, but this finding is in line with a study among older Australians who reported no consistent difference in the risk between those with parental and sibling history of type 2 diabetes [37] but contradicts a study among ethnic Chinese people [38]. The difference between the studies may relate to age of onset of diabetes in the family members. Among younger people, a sibling with diabetes is likely to have developed diabetes at a younger age than a parent with diabetes, and hence may reflect a greater risk. However, among older adults, there is less likely to be a difference in age of onset between siblings and parents with diabetes. Furthermore, we found that the presence of both a parent and a sibling with diabetes greatly increases the risk for diabetes. This graded effect of family history on risk for diabetes was pointed out by the NHANES investigators [39].

Our study obviously has some limitations. First, although especially the case for smaller datasets, external validation is needed before implementing a prediction model. It is well-known that the performance of a prediction model is generally higher in the development dataset than in another dataset, especially when the population differs (in time or place) from the original population. However, to some extent, external validation can be studied by testing the performance of the model in a non-random selection of the population, which was done here by separate assessments of the performance in each study centre. This showed good performance, which is promising for external application of the model. Second, although we showed good model performance in Europids,

extrapolation to other populations should be done with caution. Indeed, we found different associations between risk predictors and incident type 2 diabetes in the Mauritius cohort and therefore excluded the latter from the present analysis. Recalibration of the score in other populations would probably lead to more precise risk estimates and is therefore indicated if data are available. Ethnicity within England and Wales has indeed been shown to be an important predictor for future drug-treated type 2 diabetes [15]. In the same study, socioeconomic differences were also found to be predictive. Studies are needed to assess whether ethnic and socioeconomic measures would improve existing models in the developing world. Third, the characteristics of participants of current population-based cohort studies may differ from the characteristics of people participating in real-life step-wise screening programmes because attendance rate for screening programmes is generally low [40]. Selective participation of those at lower risk for example will reduce the positive predictive value of a risk score. Selective participation was not observed in the ADDITION-Denmark study, using a mail-distributed risk chart as a first screening step [40]. Instead, overall participation rate was low, which indicated that one of the major challenges in implementing primary prevention programmes is to reach appropriate attendance rates. Fourth, although test characteristics of the present score show good results, follow-up measurements will be indicated for a substantial part of the population (40%), which has important implications for healthcare. However, of this high-risk population, 11% will indeed develop diabetes within the relatively short timeframe of 5 years whereas lifetime diabetes risk will be higher. Further, diabetes risk scores have also been shown to predict cardiovascular disease [41, 42]. Thus, the 11% of the high-risk population that will develop diabetes within 5 years is the minimum that will benefit from timely intervention. Fifth, the WHO definition of type 2 diabetes requires two abnormal glucose values for the diagnosis of type 2 diabetes without symptoms [18]. Our diagnosis of type 2 diabetes is based on one OGTT only. To our knowledge, no observational studies have information on the incidence of type 2 diabetes measured by duplicate OGTT, so the present data are the best estimate of OGTT-based incidence of type 2 diabetes currently available.

The present risk questionnaire may serve as a first assessment tool identifying those who may need further blood testing, probably not limited to glucose testing alone. Integrated prevention strategies and targeted interventions should also involve blood pressure and lipid levels because these risk factors are highly inter-related. Future research may aim to develop questionnaires that predict overall risk, including cardiovascular disease risk.

To conclude, the predictive value of the original Finnish risk questionnaire for clinically diagnosed and screen-detected type 2 diabetes could be improved by re-estimation of the score points and by additional information on sex, smoking and family history of diabetes. The DETECT-2 update of the Finnish diabetes risk questionnaire is an adequate and robust risk score to predict future type 2 diabetes in Europid populations.

References

1. Lindstrom J, Ilanne-Parikka P, Peltonen M et al (2006) Sustained reduction in the incidence of type 2 diabetes by lifestyle intervention: follow-up of the Finnish Diabetes Prevention Study. Lancet 368:1673–1679
2. Knowler WC, Barrett-Connor E, Fowler SE et al (2002) Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med 346:393–403
3. Lindstrom J, Peltonen M, Eriksson JG et al (2008) Determinants for the effectiveness of lifestyle intervention in the Finnish Diabetes Prevention Study. Diabetes Care 31:857–862
4. Griffin SJ, Little PS, Hales CN, Kinmonth AL, Wareham NJ (2000) Diabetes risk score: towards earlier detection of type 2 diabetes in general practice. Diabetes Metab Res Rev 16:164–171
5. Glumer C, Carstensen B, Sandbaek A, Lauritzen T, Jorgensen T, Borch-Johnsen K (2004) A Danish diabetes risk score for targeted screening: the Inter99 Study. Diabetes Care 27:727–733
6. Ruige JB, de Neeling JN, Kostense PJ, Bouter LM, Heine RJ (1997) Performance of an NIDDM screening questionnaire based on symptoms and risk factors. Diabetes Care 20:491–496
7. Meigs JB, Shrader P, Sullivan LM et al (2008) Genotype score in addition to common risk factors for prediction of type 2 diabetes. N Engl J Med 359:2208–2219
8. Chien K, Cai T, Hsu H et al (2009) A prediction model for type 2 diabetes risk among Chinese people. Diabetologia 52:443–450
9. Stern MP, Williams K, Haffner SM (2002) Identification of persons at high risk for type 2 diabetes mellitus: do we need the oral glucose tolerance test? Ann Intern Med 136:575–581
10. Schmidt MI, Duncan BB, Bang H et al (2005) Identifying individuals at high risk for diabetes: the Atherosclerosis Risk in Communities Study. Diabetes Care 28:2013–2018
11. Lindstrom J, Tuomilehto J (2003) The diabetes risk score: a practical tool to predict type 2 diabetes risk. Diabetes Care 26:725–731
12. Schulze MB, Hoffmann K, Boeing H et al (2007) An accurate risk score based on anthropometric, dietary, and lifestyle factors to predict the development of type 2 diabetes. Diabetes Care 30:510–515
13. Wilson PWF, Meigs JB, Sullivan L, Fox CS, Nathan DM, D'Agostino RBS (2007) Prediction of incident diabetes mellitus in middle-aged adults: the Framingham Offspring Study. Arch Intern Med 167:1068–1074
14. Aekplakorn W, Bunnag P, Woodward M et al (2006) A risk score for predicting incident diabetes in the Thai population. Diabetes Care 29:1872–1877

15. Hippisley-Cox J, Coupland C, Robson J, Sheikh A, Brindle P (2009) Predicting risk of type 2 diabetes in England and Wales: prospective derivation and validation of QDScore. BMJ 338:b880

16. Sandbaek A, Griffin SJ, Rutten G et al (2008) Stepwise screening for diabetes identifies people with high but modifiable coronary heart disease risk. The ADDITION Study. Diabetologia 51: 1127–1134

17. Alssema M, Feskens EJ, Bakker SJ et al (2008) Finnish questionnaire reasonably good predictor of the incidence of diabetes in the Netherlands. Ned Tijdschr Geneeskd 152: 2418–2424

18. World Health Organization (2006) Definition and diagnosis of diabetes mellitus and intermediate hyperglycemia—report of a WHO/IDF consultation. WHO, Geneva

19. DECODE Study Group et al (2001) Glucose tolerance and cardiovascular mortality: comparison of fasting and 2-hour diagnostic criteria. Arch Intern Med 161:397–405

20. Mykkanen L, Laakso M, Uusitupa M, Pyorala K (1990) Prevalence of diabetes and impaired glucose tolerance in elderly subjects and their association with obesity and family history of diabetes. Diabetes Care 13:1099–1105

21. Snijder MB, Zimmet PZ, Visser M, Dekker JM, Seidell JC, Shaw JE (2004) Independent and opposite associations of waist and hip circumferences with diabetes, hypertension and dyslipidemia: the AusDiab Study. Int J Obes Relat Metab Disord 28:402–409

22. Patja K, Jousilahti P, Hu G, Valle T, Qiao Q, Tuomilehto J (2005) Effects of smoking, obesity and physical activity on the risk of type 2 diabetes in middle-aged Finnish men and women. J Intern Med 258:356–362

23. Steyerberg EW, Borsboom GJ, van Houwelingen HC, Eijkemans MJ, Habbema JD (2004) Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. Stat Med 23:2567–2586

24. Colagiuri S, Borch-Johnsen K (2003) DETECT-2: early detection of type 2 diabetes and IGT. Diabetes Voice 48:11–13

25. de Vegt F, Dekker JM, Jager A et al (2001) Relation of impaired fasting and postload glucose with incident type 2 diabetes in a Dutch population: the Hoorn Study. JAMA 285:2109–2113

26. Glumer C, Jorgensen T, Borch-Johnsen K (2003) Prevalences of diabetes and impaired glucose regulation in a Danish population: the Inter99 Study. Diabetes Care 26:2335–2340

27. Stegmayr B, Lundberg V, Asplund K (2003) The events registration and survey procedures in the Northern Sweden MONICA Project. Scand J Public Health Suppl 61:9–17

28. Marmot M, Brunner E (2005) Cohort Profile: the Whitehall II Study. Int J Epidemiol 34:251–256

29. Dunstan DW, Zimmet PZ, Welborn TA et al (2002) The Australian Diabetes, Obesity and Lifestyle Study (AusDiab)—methods and response rates. Diabetes Res Clin Pract 57:119–129

30. Soderberg S, Zimmet P, Tuomilehto J et al (2004) High incidence of type 2 diabetes and increasing conversion rates from impaired fasting glucose and impaired glucose tolerance to diabetes in Mauritius. J Intern Med 256:37–47

31. Rose G, Blackburn H (1968) Cardiovascular survey methods. WHO, Geneva

32. Hanley JA, McNeil BJ (1983) A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 148:839–843

33. Hosmer DW, Lemeshow S (2000) Applied logistic regression, 2nd edn. Wiley, New York

34. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS (2008) Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 27:157–172

35. Efron B, Tibshirani R (1993) An introduction to the bootstrap. Monographs on statistics and applied probability. Chapman & Hall, New York

36. de Vegt F, Dekker JM, Ruhe HG et al (1999) Hyperglycaemia is associated with all-cause and cardiovascular mortality in the Hoorn population: the Hoorn Study. Diabetologia 42:926–931

37. Cugati S, Wang JJ, Rochtchina E, Mitchell P (2007) Ten-year incidence of diabetes in older Australians: the Blue Mountains Eye Study. Med J Aust 186:131–135

38. Chien KL, Hsu HC, Su TC et al (2008) Sibling and parental history in type 2 diabetes risk among ethnic Chinese: the Chin-Shan Community Cardiovascular Cohort Study. Eur J Cardiovasc Prev Rehabil 15:657–662

39. Valdez R, Yoon PW, Liu T, Khoury MJ (2007) Family history and prevalence of diabetes in the U.S. population: the 6-year results from the National Health and Nutrition Examination Survey (1999–2004). Diabetes Care 30:2517–2522

40. Christensen JO, Sandbaek A, Lauritzen T, Borch-Johnsen K (2004) Population-based stepwise screening for unrecognised type 2 diabetes is ineffective in general practice despite reliable algorithms. Diabetologia 47:1566–1573

41. Silventoinen K, Pankow J, Lindstrom J, Jousilahti P, Hu G, Tuomilehto J (2005) The validity of the Finnish Diabetes Risk Score for the prediction of the incidence of coronary heart disease and stroke, and total mortality. Eur J Cardiovasc Prev Rehabil 12:451–458

42. Chamnan P, Simmons RK, Hori H et al (2010) A simple risk score using routine data for predicting cardiovascular disease in primary care. Br J Gen Pract 60:327–334