Diabetologia

Population-based stepwise screening for unrecognised Type 2 diabetes is ineffective in general practice despite reliable algorithms

J. O. Christensen[1] · A. Sandbæk[2] · T. Lauritzen[2] · K. Borch-Johnsen[1, 3]

[1] Steno Diabetes Center, Gentofte, Denmark
[2] The Department of General Medical Practice, University of Aarhus, Aarhus C, Denmark
[3] The Faculty of Health Sciences, University of Aarhus, Aarhus C, Denmark

Abstract

Aims/hypothesis. The yield of screening programmes for Type 2 diabetes in the existing healthcare setting might be lower than anticipated from tests of screening algorithms in data from epidemiological surveys. Our aims were to evaluate the reliability of the algorithms and the effectiveness of a proposed stepwise screening programme for Type 2 diabetes in general practice.
Methods. The screening programme had four steps: (i) mail-distributed self-administered risk-chart; (ii) screening tests: random blood glucose (RBG) and HbA_1c; (iii) diagnostic procedure 1 for fasting blood glucose (FBG) (if RBG ≥5.5 mmol/l or HbA_1c ≥6.1%); and (iv) OGTT as diagnostic procedure 2 (if 5.6≤FBG<6.1 mmol/l or HbA_1c ≥6.1%). Abnormalities of glucose metabolism were classified according to the WHO 1999 criteria, based on capillary whole blood. The subjects were all patients between 40 and 69 years of age (*n*=60,926) who were registered in 88 general practices and had not been previously diagnosed with diabetes.

Results. A total of 11,263 individuals had a high-risk risk-score and attended the screening consultation (step 1 test-positive). Of these, 30.1% needed diagnostic tests (step 2 test-positive) and 27.2% of these needed an OGTT (step 3 test-positive). The test-positive proportions were equal to the proportions obtained in data from a population-based survey from Step 2 onwards, and the algorithms were thus reliable. The identification rate was only 19% of all prevalent undiagnosed diabetes according to a recently published prevalence estimate. This was due to a large dropout rate among high-risk individuals prior to entry into the programme.
Conclusions/interpretation. Population-based mail-distributed stepwise screening for Type 2 diabetes in general practice is ineffective, despite reliable screening algorithms, primarily because many high-risk individuals fail to participate.

Keywords Effective · General practice · Reliable · Screening · Type 2 diabetes

J. O. Christensen (✉)
Steno Diabetes Center, Niels Steensens Vej 2,
2820 Gentofte, Denmark
E-mail: joc@dadlnet.dk
Tel.: +45-44438792, Fax: +45-44430706

Abbreviations: FBG, fasting blood glucose · RBG, random blood glucose · WHO, World Health Organization

Introduction

The increasing prevalence of Type 2 diabetes and the high proportion of undiagnosed Type 2 diabetes [1, 2, 3] has raised the need for developing simple and sensitive screening strategies. Recently, the Ministry of Interior and Health in Denmark launched a diabetes action plan recommending intensified and optimised case finding among high-risk individuals in general practice [4]. This is possible only with tools or programmes that ensure systematic tracing of high-risk individuals. Several tools designed to identify high-risk individuals have already been developed [5, 6, 7, 8, 9, 10, 11, 12]. These tools were not only developed

but also tested in datasets from population-based surveys. However, the performance and the yield of such screening tools, when evaluated in such datasets, do not necessarily reflect the performance and the yield when applied as part of a programme in an existing healthcare setting. In a healthcare setting the yield of a screening programme is affected by the interest or willingness of people to participate in the programme, and by the ability of the healthcare professionals to adhere to the programme's guidelines. Only few studies have evaluated what can realistically be achieved by launching such programmes in an existing healthcare setting [13, 14].

The aims of this study were therefore to test a proposed stepwise screening programme for previously unrecognised Type 2 diabetes in general practice under the conditions given by the healthcare system. This was done by evaluating: (i) the reliability of the algorithms laid down in the programme (i.e., do the algorithms perform in the healthcare setting as in a population-based study?); (ii) the effectiveness of the programme (i.e., which proportion of the prevalent cases with unidentified diabetes is actually identified?); and (iii) the degree of dropout.

Subjects and methods

The screening programme. Based on theoretical considerations, we developed a stepwise screening programme to be applied in general practice under the conditions given by the existing healthcare system in Denmark. The programme consisted of a screening algorithm and a diagnostic algorithm, each with two steps (Fig. 1). Each of the steps divided the participants into two groups: (i) a high-risk group (the test-positive individuals), who would subsequently go on to the next step in the programme; and (ii) a low-risk group (the test-negative individuals) who would not be examined further. The purpose of the screening algorithm was to minimise the number of subsequent diagnostic tests, and that of the diagnostic algorithm was to identify as many persons as possible with previously unknown diabetes mellitus in the high-risk group, and at the same time to minimise the number of OGTTs.

To avoid referring large numbers of patients to central laboratories to have gold-standard plasma glucose measured, all practices measured blood glucose in capillary whole blood, using HemoCue.

The algorithms. The first step of the screening algorithm consisted of a mail-distributed, self-administered questionnaire in the form of a risk-chart. The risk-chart used was an early version of the "Danish Risk Score" developed in the Inter99 Study [6]. It included the same questions, but with a preliminary and simplified parameterisation. Table 1 lists the questions laid

Table 1. The seven questions in the risk-chart (Step 1)

Question		Answer	Points
1	Sex	Male	1
		Female	0
2	How old are you?	40–44	0
		45–49	1
		50–54	2
		55–59	3
		60–69	4
3	Have you ever had diabetes that went away again? [a]	Yes	2
		No	0
4	Do or did your parents or siblings have diabetes?	Neither parents nor siblings have/had diabetes.	0
		One of my siblings or one of my parents have/had diabetes.	1
		Two or more of my siblings or parents have/had diabetes.	2
		I do not know.	0
5	Have you ever been told that you have/had high blood pressure or hypertension?	Yes	2
		No	0
		I do not know.	0
6	Modified BMI chart dividing patients into three BMI classes	BMI <25	0
		BMI 25–30	1
		BMI >30	2
7	Which of the following groups do your recreational activities fit into?	I usually sit down and read, watch TV, go to the movies and spend my spare time with sedentary activities.	1
		I walk or cycle or conduct physical activities at least 4 hours a week (construction, housework, table tennis, bowling).	1
		I am active with sports at least three times a week or regularly perform heavy house or garden work.	0
		I participate in sports competitions or go in for long-distance running several times a week.	0
Add your points together.			0–14

The questions are translated from the original Danish. [a] The question about previous diabetes (gestational diabetes) was not included in the questionnaires in the Inter99 study and thus was not a result of the regression analysis [6]. It was added to the risk-chart because these patients have a high risk of developing Type 2 diabetes later in life [18]
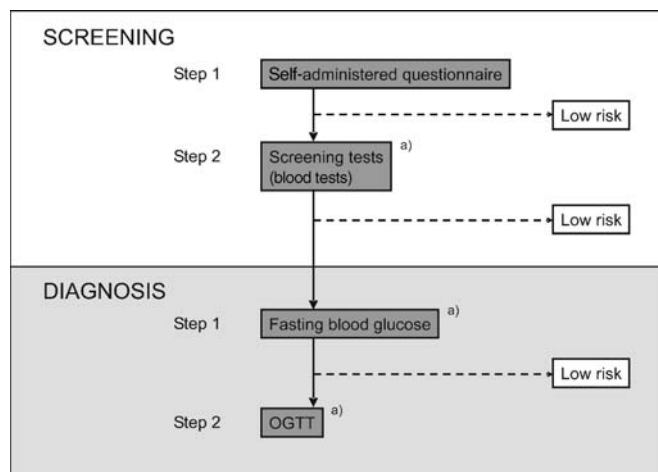
Fig. 1. Overall outline of the stepwise screening programme. a = if fasting blood glucose or 2-h blood glucose were diabetic, diagnosis was confirmed with a second diagnostic test on a separate day
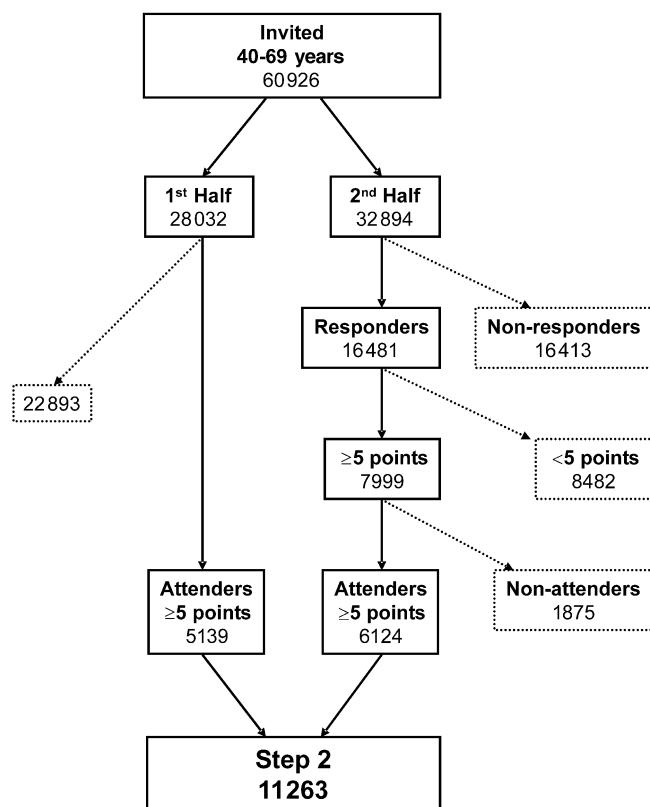

Fig. 2. Screening algorithm: Step 1

down in the risk-chart (translated from Danish to English) and the parameterisation. The numerical score associated with the answer to each of the questions could be summarised to an overall risk-score between 0 and 14 points after completion of the questionnaire.

The second step took place in general practice and included the biochemical screening tests for random blood glucose (RBG) and HbA_1c. The cut-off points (RBG ≥ 5.5 mmol/l or $HbA_1c \geq 6.1\%$) defined a high-risk group that subsequently underwent diagnostic procedures.

In the diagnostic algorithm everyone started with FBG. Based on the criteria below, only high-risk individuals went on for an OGTT.

In the first step of the diagnostic algorithm the combined cut-off points (FBG<5.6 mmol/l and HbA_1c <6.1% [unless diabetic in the screening test: RBG$\geq$11.1 mmol/l]) defined a group with low risk of having unidentified diabetes mellitus. This group did not have further examinations.

Following the FBG, individuals with an HbA_1c of $\geq 6.1\%$ or $5.6 \leq FBG < 6.1$ mmol/l underwent an OGTT.

Individuals with one diabetic blood glucose value underwent confirmatory procedures.

The exact outlines of the algorithms are depicted in Figures 2 and 3.

Subjects. We included 88 general practices participating in the Danish arm of the ADDITION study [15]. Practice size varied from 234 to 2618 patients of 40 to 69 years of age and from 1 to 4 general practitioners. Invitations to participate went to all patients who were registered with these practices and aged 40 to 69 years on January 1st, 2001 and also did not have pre-existing diabetes mellitus or severe mental, psychiatric, or somatic diseases that would hinder participation in the study (male : female ratio: 0.96). All eligible persons were randomly divided into 12 batches. The batches of invitations were mailed in regular intervals from the 26th of March 2001 to the 18th of June 2002. All invitations were prepared by the study secretariat at Aarhus University, but were sent with the patients' own general practitioner as the sender.

Strictly speaking, the patients selected to take part in this study are not a real population-based sample, but very close to being so. Because patients registered in general practice account for 97% or more of the whole population in Denmark and because these patients are the only ones entitled to free healthcare in Denmark, the study can, for practical purposes, be regarded as population-based.

Methods. The letter of invitation contained a short description of the epidemiology of Type 2 diabetes in Denmark, the risks associated with untreated Type 2 diabetes and an invitation to participate in a screening programme for unrecognised Type 2 diabetes. The letter further contained the risk-chart described above (Table 1) to be filled-in at home. Individuals with a risk-score of 5 points or more were asked to make a non-fasting appointment for screening with his/her general practitioner. To be able to estimate the response and attendance rates among people with a risk-score of 5 points or more, we changed the invitation after the sixth batch and asked the individuals in the remaining six batches to respond either by making an appointment for screening or by returning their filled-in risk-chart to the study secretariat, regardless of their risk-score. The invitations thus fell into two groups: a first group (*n*=28,032) whose data could be used for analysis of yield and reliability of algorithms, and a second group (*n*=32,894) whose data could also be used to assess questions about non-response and non-attendance. All information about the participants was collected by the general practitioners or their staff and reported in case record forms, which were mailed to the study secretariat in Aarhus.

At the first visit to the practice the participants gave informed consent to participate in the study. During this visit, venous blood for HbA_1c and a non-fasting RBG was taken and measured and the risk-chart was collected.

At the second visit FBG was measured and subsequently a standard 75-g OGTT was carried out if needed according to the algorithm.
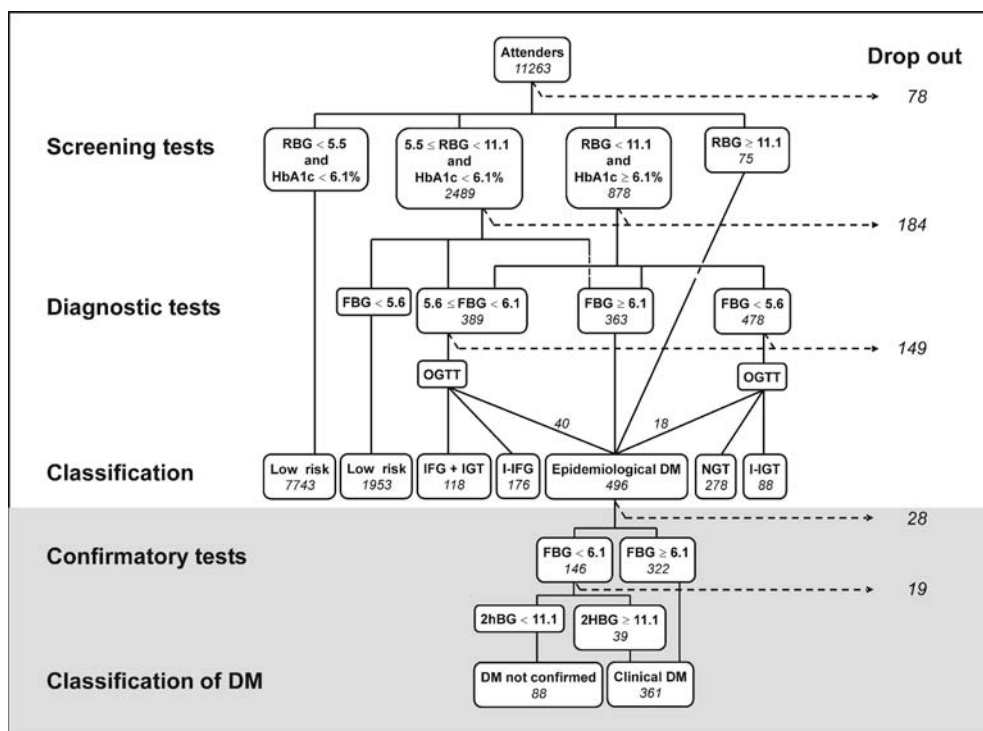
Fig. 3. Screening algorithm: Step 2, diagnostic algorithm and classification. Each box gives a criterion and the number of patients fulfilling this criterion. Dotted lines = patients that could not be classified due to missing examinations (classified as dropouts at the level indicated in the figure). The classification is based on the WHO 1999 criteria using capillary whole-blood cut-off values: NGT: fasting blood glucose (FBG) <5.6 mmol/l and 2-h BG <7.8 mmol/l; 5.6≤FBG<6.1 mmol/l and 2-h BG <7.8 mmol/l; IGT: FBG <5.6 mmol/l and 7.85 2-h BG <11.1 mmol/l; IFG+IGT: 5.6≤FBG<6.1 mmol/l and 7.8≤2-h BG<11.1 mmol/l; epidemiological diabetes mellitus: RBG ≥11.1 mmol/l or FBG ≥6.1 mmol/l or 2-h BG ≥11.1 mmol/l

To confirm the clinical diagnosis of diabetes mellitus a third visit was planned for participants who had their first diabetic blood glucose value at the second visit.

Measures of blood glucose. Blood glucose concentrations were analysed in finger capillary blood samples. This was done instantly using a HemoCue B-glucose analyser. The analysis method used by the HemoCue analyser is photometric and based on a glucose dehydrogenase reaction in whole blood. The result is displayed as a whole-blood capillary value. Coefficients of variation were 4.0%, 3.5%, 2.2% and 1.0% at glucose concentrations of 3.9, 5.4, 8.7 and 17.7 mmol/l respectively [16]. All practices had the same version of the HemoCue B-glucose analyser and all machines were initially calibrated against a standardised chemical glucose dehydrogenase method. All general practitioners and their staff involved in blood sampling received standardised and individualised training by an expert from the HemoCue company on how to handle blood sampling and analysis. During the study daily checks of the calibration were done in the clinics and the supervisor from the HemoCue Company made regular control visits to all the clinics. For random capillary blood glucose (screening test) one sample was taken, but for all diagnostic tests (FBG and 2-h blood glucose after OGTT) two samples were taken and the average of the two results was used to minimise the measurement error.

HbA$_1$c was analysed in venous blood sampled in EDTA tubes. All samples were mailed directly from the general practitioner to a central laboratory (Aarhus County Hospital, Aarhus C, Denmark) and analysed on the day of arrival in the laboratory using ion-exchange high-performance liquid chromatography on a Tosoh machine (normal range 4.2% to 6.3%). The result was reported back to the general practitioner.

Classification. The diagnostic classification was based on the 1999 WHO criteria for capillary whole blood [17]. The clinical diagnosis of diabetes mellitus was made only if an initial diagnosis was confirmed by a second diabetic blood glucose value on another day. Patients with one diabetic blood glucose value were labelled "epidemiological diabetes mellitus" (Fig. 3).

Pilot study. Data from a pilot study that was conducted in advance of the study itself was included. This was done to enable us to judge the performance of the algorithms and to estimate the sensitivity of the steps included in the algorithms.

The pilot study was a population-based survey, in which all patients of 40 to 69 years of age who were registered with five general practices (n=2051) were invited to participate. Of those invited, 1028 agreed to participate in the pilot study. The five practices and their patients did not participate in the main study. All participants had RBG and HbA$_1$c measured and filled in the risk-chart described in Table 1. On a subsequent day, everyone with RBG ≥4.5 mmol/l or HbA$_1$c ≥5.9% underwent an FBG or OGTT. People with RBG <4.5 mmol and HbA$_1$c <5.9% were considered at such low risk of having diabetes mellitus that no further examinations were carried out. These patients were classified as non-diabetic. The rest of the patients were classified according to the 1999 WHO criteria [17], based on capillary whole-blood glucose values.

Data analysis. All data handling and statistical analysis was done using the statistical computer software release 8.2 (SAS Institute, Cary, N.C., USA).

People with a missing risk-score or a risk-score of less than 5 points were not included in the analysis of the screening programme (*n*=807). Data on individuals dropping out of the programme were included in the analysis until the point of dropout (*n*=458). Results of examinations not defined in the protocol were not included in the analysis.

Means and standard deviations were used to describe normally distributed continuous variables and Student's *t* tests used for comparison. For continuous non-normally distributed data, median and interquartile ranges were used to describe the data and Wilcoxon rank-sum non-parametric tests were used for comparisons. For categorical variables, frequencies and proportions were calculated. The chi square test and 95% confidence intervals ($p \pm 1.96 \sqrt{p \cdot (1-p)/n}$) were used for comparisons.

The performance (sensitivity, specificity and test-positive fraction) of all the steps in the programme was tested in the preliminary pilot study using the cut-off values described in the algorithms section above.

The reliability of the algorithms was tested by comparing the proportion of individuals continuing to each step in this programme with the proportion expected on the basis of the test-positive fractions obtained in the pilot study. The effectiveness of the programme was evaluated by comparing the observed outcome with that expected on the basis of the prevalence of unidentified diabetes mellitus observed in the Inter99 study [3].

Ethics. The study was approved by the local scientific ethics committees in Aarhus and Copenhagen Counties, Denmark, and conducted in accordance with the principles of the Helsinki Declaration revised in 1996.

Results

Figure 2 shows the selection of participants and the dropout prior to entry in the programme. Overall 18.5% (*n*=11,263) of the invited patients made an appointment for screening and had a risk-score of 5 points or more. There was no difference between the first and second group, either in attendance (18.3% [*n*=5139] and 18.6% [*n*=6124] respectively, *p*=0.3), or in age or sex distributions (data not shown) and so they were fully comparable. However, from the data collected in the second group we looked into the response and attendance rates and found that 50.1% (*n*=16 481) responded by returning the filled-in risk-chart or by making an appointment for screening. Of these, 48.5% (*n*=7999) had a risk-score of 5 points or more. The attendance rate among individuals with a risk-score of 5 points or more was 76.6% (*n*=6124).

Figure 3 shows the rest of the procedures in the programme. The second step of the screening algorithm (RBG and HbA$_1$c) resulted in 3367 individuals (test-positive fraction [95% CI]: 30.1% [29.3 to 30.9%]) in whom diagnostic procedures were required. In the diagnostic algorithm 867 individuals (test-positive fraction [95% CI]: 27.2% [25.7 to 28.8%]) needed an OGTT for diagnostic purposes. The programme thus reduced the number of fasting tests by approximately 70% and the number of OGTT's by approximately 92% compared to a one-step approach with only a risk-chart. The programme identified 361 individuals with clinical diabetes mellitus (0.6% of those invited) out of 496 individuals with one diabetic blood glucose value (epidemiological diabetes mellitus). Clinical diabetes mellitus was not confirmed in 20% (*n*=88) of all individuals who had confirmatory procedures performed (*n*=88+361). Clinical diabetes mellitus thus constitutes 80% of epidemiologically defined diabetes mellitus when identified by stepwise screening.

The dropout prior to entry into the study was analysed in the second group of persons invited to participate (*n*=32,894) (Fig. 2). The dropout fell into two groups: non-responders (*n*=16,413) and non-attendees with a risk-score of 5 points or more (*n*=1875). Responders were more likely to be women and to be older than non-responders (Table 2). Attendees were

Table 2. Non-responder/non-attendee analysis

	Non-responders	Responders	*p* value	Non-attenders with ≥5 points	Attenders with ≥5 points	*p* value
n	16,413	16,481		1875	6124	
Sex			<0.001[a]			<0.001[a]
women	47.2	54.6		40.0	48.0	
men	52.8	45.4		60.0	52.0	
Age (years)			<0.001[b]			<0.001[b]
Median age	52.0	52.9		57.5	59.5	
Risk score						<0.001[a]
5				47.4	27.5	
6				25.2	27.2	
7				14.3	21.7	
8				6.9	13.5	
9				4.9	7.2	
10				1.2	2.2	
11					0.5	
12				0.1	0.1	
13					0.0	

Results are percentages, except for age, *n* and *p* values. [a] Chi square test; [b] Wilcoxon rank-sum test

more likely to be women; they were older and had a higher risk-score than non-attendees. The dropout within the programme was 458 (4.1%), of which 78 did not have a full set of screening tests, 184 did not have the required fasting consultation, 149 did not have the requested OGTT and 47 with one diabetic blood glucose value did not have confirmatory tests performed (Fig. 3). No significant differences in age, sex-distribution and risk-score distribution were found between people dropping out and people completing the programme (data not shown).

The performance of the stepwise programme when tested in the population-based pilot study was: Step 1 (risk-chart): sensitivity 77.8%, specificity 62.1%, test-positive fraction 38.9% (95% CI 35.9–41.9). Step 2 (RBG + HbA$_1$c): sensitivity 85.7%, specificity 75.7%, test-positive fraction 27.7% (95% CI 23.2–32.0). Step 3 (FBG): sensitivity 100%, specificity 67.4%, test-positive fraction 30.8% (95% CI 21.9–39.7). The overall sensitivity of the combined procedures was calculated by multiplying the sensitivity of each step in the screening programme: $(0.778 \cdot 0.857 \cdot 1.0 \cdot 100\%)=66.7\%$ on the assumption that only test-positive individuals proceeded to the next step in the programme. The test-positive rates resulting from the procedures in Step 2 (RBG and HbA$_1$c) and onwards were not significantly different from the rates found in this study, but the rate resulting from Step 1 (risk-chart) was significantly lower (18.5% vs 38.9%).

Finally the yield of the programme (proportion of invited participants found to have diabetes mellitus) was compared with a recent estimate of the prevalence of unidentified epidemiological diabetes mellitus in the Danish population (Inter99 study [3]: 4.2%, 95% CI 3.7–4.7%). The 496 individuals (0.8%) identified by the programme as having epidemiological diabetes mellitus corresponded to a 19% detection rate.

Discussion

This study shows that in individuals defined as high-risk on the basis of a questionnaire and attending their general practitioner for further screening and diagnostic testing, the stepwise screening programme performed as expected when compared with the performance in a population-based study (the pilot study). On the other hand, our study also clearly demonstrates that the final outcome of the application of these screening tools in a healthcare setting differs from what would be expected on the basis of epidemiological population-based surveys.

In this paper, test-positive proportions in each of the steps in a stepwise screening programme applied in the daily clinic under existing healthcare conditions were compared with those obtained by the same programme in a study (the pilot study) that mimics a population-based survey. Reliability thus refers to the comparison of the above-named test-positive proportions. This definition was chosen because data collected in a study of performance in the daily clinic do not allow comparison of sensitivity, specificity or positive and negative predictive values. The only parameters applicable for an evaluation of the performance of the algorithms were the test-positive fractions for each step in the programme. We considered the algorithms to be reliable if the test-positive fractions as applied in everyday clinical practice did not differ from those obtained in an ideal world (population-based survey). If these fractions were equal, we assumed that the performance of the algorithms in terms of sensitivity, specificity and positive and negative predictive value would also be equal. However, test-positive fractions of a test depend on the population, whereas the test characteristics (sensitivity, specificity, positive and negative predictive values) depend on the test. Our conclusion is thus based on the underlying assumption that the risk-factor distribution in the two populations was the same, which seemed to be the case, since among responders with a risk-score of 5 points or more there were no differences in the risk factor distribution between the two studies.

Based on the most recently published data for prevalence of undiagnosed epidemiological diabetes mellitus (4.2%, 95% CI 3.7–4.7%) in Denmark [3], 2559 of the 60 926 persons invited to participate in this study should have had unidentified epidemiological diabetes mellitus. Of these 1707 could have been identified by the programme, if all individuals with a risk-score of 5 points or more had participated in the programme (66.7% sensitivity of the whole programme). In fact, the screening programme only identified 496 individuals with previously undiagnosed diabetes mellitus, leaving 1211 undetected cases of diabetes mellitus. There are several explanations for this low detection rate, but the most likely is the large dropout, and primarily the dropout prior to entry into the programme.

There are two ways of estimating the dropout among persons with a risk-score of 5 points or more prior to entry into the study, but both ways are based on assumptions. Firstly, if we assume that risk-score distributions are equal among responders and non-responders, the estimated dropout would be 61%: 50.1% response and 77% attendance of people with a risk-score ≥5 points $(1-[0.501\times0.77]=0.61)$. Secondly, if we assume that the risk-score distributions are equal among responders and non-responders, both in the pilot study and among all persons invited to participate in this study, the estimated dropout would be 52%, calculated on the basis of 38.9% test-positive subjects in the pilot study and 18.5% of all invited participants who attended and had a risk-score ≥5 points in this study $(1-[18.5/38.9]=0.52)$. The dropout among individuals with a risk-score of 5 points or more prior to entry into the study is therefore between 52% and 61%, and thus accounts for between 888

(73%) and 1041 (86%) of the undetected patients with unknown diabetes in the population. The dropout within the study only accounts for 36 undetected individuals, if we assume the same prevalence of diabetes mellitus in the groups dropping out as in the groups completing the programme.

The low yield might also be explained by selective non-response. The above calculations of how dropout might explain much of the low yield assume that the risk distribution is the same among non-responders as among responders and that the prevalence of diabetes mellitus is the same among non-attendees as among the attendees. Whether these assumptions are true is not known. The only data we have for a responder/ non-responder comparison are age and sex and there are clearly differences. One parameter points to higher risk (more men) and the other points to lower risk (younger) among the non-responders. Among attendees/non-attendees we have a little more information. Again non-attendees are younger, but more likely to be male, but their risk profile is more favourable than that of attendees. Thus, there are no indications that the low yield of the programme is due to selective non-response/non-attendance, but rather that the yield is low despite selective attendance.

Another explanation for the low yield could be that the prevalence of unknown Type 2 diabetes is lower in this study population than that reported by the Inter99 study [3]. Bias might have increased the prevalence of diabetes mellitus in the Inter99 study [3]. Another recent Danish survey [2] reports a slightly lower prevalence of unidentified diabetes mellitus among 60-year-old Danish subjects than the Inter99 study (4.9% vs 7.5%), but this difference does not explain the lower detection rate in our study.

The true explanation is probably a combination of all these factors, but a major factor is certainly the large dropout prior to entry into the study.

An important question is therefore whether the dropout can be minimised?

Changing the algorithms would probably only affect the internal dropout and our calculations show that only little can be gained. Thus if there had been no internal dropout, only 36 additional individuals could have been detected. The data from this study do not allow us to estimate the effects of changing the algorithms in a way that might enhance the sensitivity of the programme. Theoretically such enhancement inevitably leads to more low-risk individuals being classified as high-risk, which in turn might increase the risk of dropout.

One way to increase the response rate could be reminder procedures. The Hoorn screening study [13] used reminder procedures and had an overall response rate of 78%, but the effect of sending out reminders has not been reported. In Denmark, the population-based screening study Inter99 also used reminders, but still only had an overall response rate of 53% [3]. The

Hoorn Study also had a very high attendance rate among high-risk individuals (87%). But the high-risk individuals were invited directly after central evaluation of their risk, which might have increased the attendance rate, compared to a self-administered approach. However, a study where high-risk individuals were systematically identified in clinical practice from information in clinical databases and thereafter selectively invited to be screened for unidentified diabetes mellitus had a much lower attendance rate (44%) [14] than in our study and in the Hoorn Study.

The current Danish Health Care System does not allow procedures such as central evaluation of mail-distributed risk-charts or direct identification of high-risk individuals in clinical databases, nor are reminder procedures feasible. But changing the mail-distributed population-based approach into systematic tracing of high-risk individuals inside general practice could be feasible and more effective. It could be done, for example, by presenting the risk-chart to people attending the practice for other reasons (opportunistic screening); the same algorithms could be used.

The dropout within the study was comparable to that found (i) in a population-based high-risk stepwise approach (the Hoorn screening study [13]) and (ii) in a pure population-based approach (Inter99 [3]), but it was lower than in the high-risk approach [14]. As in [13, 14] the dropout was particularly high among people who underwent an OGTT. Our findings thus emphasise the difficulties associated with the use of OGTTs in population-based screening, even when required only in high-risk individuals.

All calculations and considerations above refer to the epidemiological definition of diabetes mellitus (at least one diabetic blood glucose value), because data for comparison utilising the clinical definition are not available. Screening should, however, not be for epidemiological diabetes mellitus, but for clinical diabetes mellitus. Our study showed that only 70% of the individuals classified with diabetes mellitus on the basis of the epidemiological definition ended up with the clinical diagnosis (10% dropout and 20% not confirmable). Expectations of the yield of future screening programmes must therefore be adjusted downwards by a further 30%, when based on data from epidemiological surveys. The 10% dropout rate among patients with the highest possible risk (people with one diabetic blood glucose value) emphasises the need for action to improve people's awareness of Type 2 diabetes as a severe disease if unattended.

This study was conducted and analysed very much in line with the way procedures would have been performed if the programme had been launched nationwide as part of the Danish Health Care System. There was no follow-up on non-response or non-attendance, the data analyses did not take into account that some of the dropouts had follow-up examinations at a later stage and that 807 people with a low risk-score also

attended the screening. The results of this study therefore do not overestimate, but rather represent an accurate estimation of what can be achieved by such a public screening programme.

We therefore conclude that the screening algorithms included in a stepwise screening programme for Type 2 diabetes in general practice were reliable because they performed as expected when applied in clinical practice. However, a population-based approach using mail-distributed invitations was ineffective primarily because only a minor fraction of the high-risk individuals in the population participated.

References

1. King H, Rewers M (1993) Global estimates for prevalence of diabetes mellitus and impaired glucose tolerance in adults. WHO Ad Hoc Diabetes Reporting Group. Diabetes Care 16:157–177
2. Drivsholm T, Ibsen H, Schroll M, Davidsen M, Borch-Johnsen K (2001) Increasing prevalence of diabetes mellitus and impaired glucose tolerance among 60-year-old Danes. Diabet Med 18:126–132
3. Glümer C, Jørgensen T, Borch-Johnsen K, Inter99 Study (2003) Prevalences of diabetes and impaired glucose regulation in a Danish population: the Inter99 study. Diabetes Care 26:2335–2340
4. Indenrigs—og Sundhedsministeriet. Handlingsplan om diabetes. Available from http://www.im.dk/publikationer/diabetesbeh/handlingsplan.pdf, accessed 6 February 2004. (Danish language only)
5. 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
6. Glümer C, Carstensen B, Sandbæk A, Lauritzen T, Borch-Johnsen K (2004) A Danish diabetes risk score for targeted screening. Inter99(6). Diabetes Care 27:727–733
7. Baan CA, Ruige JB, Stolk RP et al. (1999) Performance of a predictive model to identify undiagnosed diabetes in a health care setting. Diabetes Care 22:213–219
8. 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
9. Herman WH, Smith PJ, Thompson TJ, Engelgau MM, Aubert RE (1995) A new and simple questionnaire to identify people at increased risk for undiagnosed diabetes Diabetes Care 18:382–387
10. Ruige JB, Neeling JN de, Kostense PJ, Bouter LM, Heine RJ (1997) Performance of an NIDDM screening questionnaire based on symptoms and risk factors. Diabetes Care 20:491–496
11. Baehring TU, Schulze H, Bornstein SR, Scherbaum WA (1997) Using the World Wide Web—a new approach to risk identification of diabetes mellitus. Int J Med Inf 46:31–39
12. Herman WH, Engelgau MM, Zhang Y, Brown MB (2000) Use of GHb (HbA(1c)) to screen for undiagnosed diabetes in the U.S. population. Diabetes Care 23:1207–1208
13. Spijkerman AM, Adriaanse MC, Dekker JM et al. (2002) Diabetic patients detected by population-based stepwise screening already have a diabetic cardiovascular risk profile. Diabetes Care 25:1784–1789
14. O'Connor PJ, Rush WA, Cherney LM, Pronk NP (2001) Screening for diabetes mellitus in high-risk patients: cost, yield, and acceptability. Effective Clinical Practice 4:271–277
15. Lauritzen T, Griffin S, Borch-Johnsen K, Wareham NJ, Wolffenbuttel BHR, Rutten G (2000) The ADDITION study: proposed trial of the cost-effectiveness of an intensive multifactorial intervention on morbidity and mortality among people with Type 2 diabetes detected by screening. Int J Obes Relat Metab Disord 24:S6–S11
16. Voss EM, Cembrowski GS (1993) Performance characteristics of the HemoCue B-Glucose analyzer using whole-blood samples. Arch Pathol Lab Med 117:711–713
17. Alberti KGMM, Zimmet PZ (1998) Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: Diagnosis and classification of diabetes mellitus. Provisional report of a WHO consultation. Diabet Med 15:539–553
18. Henry OA, Beischer NA (1991) Long-term implications of gestational diabetes for the mother. Baillieres Clin Obstet Gynaecol 5:461–483