Application of diabetes risk scores in health checkups

A comparison of the German Diabetes Risk Score (GDRS) and FINDRISK test

Iris Trefflich, Carmen Jahn, Franziska Jannasch, Susanne Jäger, Matthias B. Schulze, Kristin Mühlenbruch

Abstract

In the study “Praxistest Diabetes” (PraDi) we investigated the application of the German Diabetes Risk Score (GDRS) and FINDRISK in health checkups in terms of the predictive ability to identify undiagnosed prediabetes and diabetes. 403 men and women aged 35–70 years participated in the PraDi study. The predictive ability of prediabetes/diabetes (fasting plasma glucose FPG ≥ 5.6 mmol/L) or diabetes (FPG ≥ 7.0 mmol/L) was evaluated with discrimination (area under the receiver-operating-characteristic curve [ROC-AUC]).

With regard to blood glucose tests, 93 participants had prediabetes/diabetes and 7 participants had undiagnosed diabetes. The ROC-AUC for the GDRS for identification of prediabetes/diabetes was 0.78 (95% confidence interval [95%-CI]: 0.73–0.94) which was higher than observed for FINDRISK with 0.73 (95%-CI: 0.68–0.79; p = 0.029). The ROC-AUC for identification of diabetes was 0.83 (95%-CI: 0.73–0.94) for the GDRS and 0.80 (95%-CI: 0.61–0.99) for FINDRISK (p = 0.67). In the context of health checkups, the GDRS showed a more precise identification of patients with prediabetes/diabetes or diabetes compared to FINDRISK.

Keywords: type 2 diabetes mellitus, risk score, predictive accuracy of diabetes risk tests, German Diabetes Risk Score, FINDRISK

Introduction

In the past years, a large amount of prediction models was published enabling the prediction of future development of type 2 diabetes based on known risk factors [1]. Many of these prediction models are based on solely non-invasively assessable risk factors [2] and therefore, seem to be well suitable for application in clinical practice. However, until now implementation of such prediction models in screening was limited [3]. For early detection of type 2 diabetes in Germany, the German Diabetes Association (DDG) already recommends the application of non-invasive tests [4], but there might still be some barriers for the use in clinical practice.

For Germany, however, two tests were developed which are also increasingly applied by the general population. This is on the one hand the German Diabetes Risk Score (GDRS) which was developed at the German Institute of Human Nutrition (DIfE) and was validated in various external populations [5], and the FINDRISK (“find your risk”), which is an adaptation of the Finnish risk score (FINDRISK, Finnish Diabetes Risk Score) [7]. For both tests simple questionnaire versions were developed [6, 8] including modifiable lifestyle factors such as physical activity, diet, smoking behavior or waist circumference as well as non-modifiable risk factors such as age, family history of diabetes or prevalent hypertension. Detailed information regarding the development, calculation and validation can be found in respective previous studies [5–8].

With regard to application of the GDRS [6] and FINDRISK [7], previous cross-sectional studies [9–12] showed that both tests seem to be valid for the identification of type 2 diabetes.

So far, a comparison of the both prediction models when applied in the context of health checkups is lacking. In addition, both tests were developed for the identification of people at high risk for developing type 2 diabetes and not for the application in the screening context.

Citation:
This article is available online: DOI: 10.4455/eru.2018.041
Besides, a high risk can already mean that people are suffering from prediabetes or diabetes. Therefore, the aim of this study was to evaluate the application of the GDRS and FINDRISK in the context of health checkups in terms of the identification of undiagnosed prediabetes and diabetes.

Methods
Study population
For the underlying investigation, data from the “Praxistest Diabetes” (PradI) study was used. The PradI study is a cross-sectional study which was performed in six general practices in different districts of Potsdam from June 2012 until July 2013. The target population consisted of participants of a voluntary health examination at their general practitioner (GP) in the framework of statutory primary prevention (checkup-35). Overall, 700 questionnaires were distributed to the general practices and 508 (72%) were completed and could be used for analyses. For this investigation, participants outside the age range of 35–70 years (n = 23), who had prevalent known diabetes (n = 3) or with missing information on fasting blood glucose (FBG) or anthropometric data (n = 56) or relevant data for the calculation of the GDRS or FINDRISK (n = 22) were excluded. Finally, 403 participants remained for analysis.

Participants were recruited by the medical personnel in the general practice while waiting for the health checkup. Participants gave written informed consent and filled in the GDRS and an additional questionnaire including components for the calculation of the FINDRISK score points. For both scores the currently available versions were used which means that for the GDRS the originally published version from 2007 was applied [5]. The GDRS is based on the risk factors age, physical activity, hypertension, intake of wholegrain-bread, red meat consumption and coffee consumption, smoking behavior, alcohol consumption, body height and waist circumference. FINDRISK includes the risk factors age, diabetes in the family, waist circumference, physical activity, intake of fruits, vegetables or wholegrain-bread, drug-treated hypertension, ever increased blood glucose and body mass index (BMI). Each risk factor is presented with response categories and allocated points which are summed up at the end of the test. Based on the sum of score points an assignment into risk categories is made. Body weight and waist circumference were measured in the practice and assessed by the GPs in a separate questionnaire. The GPs further confirmed a potential prevalence of diagnosed diabetes or of ever diagnosed hypertension.

The FBG values were based on laboratory results and also filled in by the GP in the same questionnaire. For data entry and data management the Research Electronic Data Capture (RedCap) tool was used [13].

The classification into prediabetes and diabetes was based on FBG cut-offs from the health checkup. With regard to the definition proposed by the DDG [14], participants were classified into prediabetes or diabetes (henceforth denoted as prediabetes/diabetes) with a FBG ≥ 5.6 mmol/L and into diabetes with a FBG ≥ 7.0 mmol/L. The definition based on HbA1c or the 2h-value of the oral glucose tolerance test was not investigated as these measurements are non-components of the checkup-35.

Statistical Analysis
Characteristics of the study population were presented with descriptive statistics. For normally distributed variables the arithmetic mean and corresponding standard deviation (SD) were calculated and for non-normal variables median and interquartile range (IQR). Categorical variables were described using relative frequencies. The relation between score points of the GDRS and FINDRISK with FBG measures was evaluated by Spearman correlation coefficients.

The predictive accuracy for the GDRS and FINDRISK was evaluated in terms of discrimination by use of the receiver-operating-characteristic curve (ROC) and compared between the tests. This statistical performance measure describes how well the respective test distinguishes between diseased and healthy individuals [13, 16]. The area under the ROC curve (ROC-AUC) was determined for each test and for both prediabetes/diabetes and diabetes as outcomes. For the comparison of these ROC-AUCs the method proposed by Hanley and McNeil was applied [17] and the respective p-value for this test was reported. We performed a power calculation for these comparisons [17] and statistical significance level was defined as 0.05.

For existing risk categorizations of the two tests, we determined sensitivity, specificity, positive and negative predictive values for at least a high risk which means ≥ 40 score points for the GDRS and ≥ 12 score points for FINDRISK. Additionally, the optimal cut-off for each test and each outcome was determined by calculating the Youden Index (YI) [18] for which sensitivity and specificity reach maximum values at maximum YI.


Results
In PradI 236 women and 167 men of middle adult age participated (• Table 1). The average waist cir-
The circumference was 85.5 cm (SD = 18) for women and 95.0 cm (SD = 14) for men. Women were on average normal weight with a mean BMI of 24.2 kg/m² while men were on average slightly overweight with a mean BMI of 25.5 kg/m². FBG was on average within the normal range for both women (5.10 mmol/L; IQR = 0.76) and men (5.12 mmol/L; IQR = 0.69). Furthermore, participants were more likely to be physically active, not suffering from hypertension and to be non-smokers. Diabetes in the family was prevalent for 44.1% of women and 34.7% of men.

This picture was confirmed by classification of PraDi participants into pre-defined risk categories (Table 2). With the GDRS 63.3% of the population was allocated into the two lower risk categories ("low" or "still low") while it was 76.7% (up to 11 points) with FINDRISK. Within the middle risk group 13.9% of the participants were observed using FINDRISK (12–14 points) while it was 21.8% using the GDRS ("increased") and in the high or very high risk area with the GDRS 14.9% and with FINDRISK 9.5% of the participants. An agreement for the risk classification using the two risk tests was observed for 47.5% of the study population using the GDRS and for 17.2% using FINDRISK.

For the distribution of FBG across risk categories of the two scores, we observed an increasing trend (Figure 1). This was monotone for the GDRS while for FINDRISK a sharp increase in the highest risk group could be observed. Score points of both tests were positively correlated with FBG (Spearman correlation: \( r_{GDRS} = 0.41, r_{FINDRISK} = 0.32 \)).

Based on the DDG classification, 310 of the participants had normal FBG (< 5.6 mmol/L), 93 had a

### Table 1: Characteristics of the study population “Praxistest Diabetes” (PraDi)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Women</th>
<th>Men</th>
</tr>
</thead>
<tbody>
<tr>
<td>n (%)</td>
<td>236 (59)</td>
<td>167 (41)</td>
</tr>
<tr>
<td>age [Years] (SD)</td>
<td>49.0 (15.0)</td>
<td>48.0 (15.0)</td>
</tr>
<tr>
<td>height [cm] (SD)</td>
<td>167 (9.0)</td>
<td>179 (9.0)</td>
</tr>
<tr>
<td>weight [kg] (SD)</td>
<td>67.3 (17.8)</td>
<td>83.0 (18.0)</td>
</tr>
<tr>
<td>waist circumference [cm] (SD)</td>
<td>85.5 (18.0)</td>
<td>95.0 (14.0)</td>
</tr>
<tr>
<td>BMI [kg/m²] (SD)</td>
<td>24.2 (5.76)</td>
<td>25.5 (4.71)</td>
</tr>
<tr>
<td>fasting blood glucose [mmol/L] (IQR)</td>
<td>5.10 (0.76)</td>
<td>5.12 (0.69)</td>
</tr>
<tr>
<td>physical activity (5 hours/week at least)</td>
<td>61.4</td>
<td>73.7</td>
</tr>
<tr>
<td>hypertension ever</td>
<td>29.7</td>
<td>32.9</td>
</tr>
<tr>
<td>never smoked</td>
<td>43.6</td>
<td>40.1</td>
</tr>
<tr>
<td>ex-smoker</td>
<td>27.1</td>
<td>40.7</td>
</tr>
<tr>
<td>smoker</td>
<td>29.2</td>
<td>19.2</td>
</tr>
<tr>
<td>diabetes in family relationship</td>
<td>44.1</td>
<td>34.7</td>
</tr>
<tr>
<td>sum FINDRISK points (IQR)</td>
<td>8.0 (7.0)</td>
<td>8.0 (7.0)</td>
</tr>
<tr>
<td>sum DRT points (IQR)</td>
<td>32.0 (18.5)</td>
<td>37.0 (17.0)</td>
</tr>
</tbody>
</table>

---

**Fig. 1:** Boxplots of the distribution of fasting blood glucose in the risk groups of GDRS and FINDRISK
prediabetes/diabetes with a FBG of 5.6 mmol/L or higher and 7 participants a FBG value above the diabetic cut-off (≥ 7 mmol/L) (Table 3).

This classification was the basis for the following ROC analysis for prediction of prediabetes/diabetes (Figure 2). ROC-AUC for the GDRS was 0.78 (95% confidence interval [95%-CI]: 0.73–0.83) and for FINDRISK 0.73 (95%-CI: 0.68–0.79). This difference of ROC-AUC values was with 0.05 statistically significant (p = 0.0291) with a power of 98.8%. ROC-AUC values for the prediction of undiagnosed diabetes were 0.83 (95%-CI: 0.73–0.94) for the GDRS and 0.80 (95%-CI: 0.61–0.99) for FINDRISK. The difference of ROC-AUCs was with 0.03 not statistically significant (p = 0.67) and the power for this statistical test was 74.6% due to the low number of cases.

For identification of prediabetes/diabetes based on risk categories, we observed for at least elevated risk (GDRS ≥ 40 points, FINDRISK ≥ 12 points) a sensitivity of 70% and a specificity of 78%, respectively. The optimal cut-off for prediabetes/diabetes identified by the YI was 38 points for GDRS and 8 points for FINDRISK with sensitivity of 75% (GDRS) and 80% (FINDRISK) and specificity of 64% (GDRS) and 55% (FINDRISK). For undiagnosed diabetes the optimal cut-off was 40 points for the GDRS and 13 points for FINDRISK with a sensitivity of 100% (GDRS) and 71% (FINDRISK) and specificity of 64% (GDRS) and 83% (FINDRISK).

<table>
<thead>
<tr>
<th>Risk categories FINDRISK</th>
<th>&lt; 7 points</th>
<th>7–11 points</th>
<th>12–14 points</th>
<th>15–20 points</th>
<th>&gt; 20 points</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>low (0–29 points)</td>
<td>105 (26.1)</td>
<td>34 (8.4)</td>
<td>2 (0.5)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>141 (35.0)</td>
</tr>
<tr>
<td>still low (30–39 points)</td>
<td>45 (11.2)</td>
<td>49 (12.2)</td>
<td>14 (3.5)</td>
<td>6 (1.5)</td>
<td>0 (0.0)</td>
<td>114 (28.3)</td>
</tr>
<tr>
<td>increased (40–49 points)</td>
<td>4 (1.0)</td>
<td>48 (11.9)</td>
<td>24 (6.0)</td>
<td>12 (3.0)</td>
<td>0 (0.0)</td>
<td>88 (21.8)</td>
</tr>
<tr>
<td>high (50–59 points)</td>
<td>1 (0.2)</td>
<td>16 (4.0)</td>
<td>7 (1.7)</td>
<td>11 (2.7)</td>
<td>1 (0.2)</td>
<td>36 (8.9)</td>
</tr>
<tr>
<td>very high (&gt; 59 points)</td>
<td>0 (0.0)</td>
<td>7 (1.7)</td>
<td>9 (2.2)</td>
<td>7 (1.7)</td>
<td>1 (0.2)</td>
<td>24 (6.0)</td>
</tr>
<tr>
<td>total</td>
<td>155 (38.5)</td>
<td>154 (38.2)</td>
<td>56 (13.9)</td>
<td>36 (9.0)</td>
<td>2 (0.5)</td>
<td>403 (100)</td>
</tr>
</tbody>
</table>

Tab. 2: Comparison of the risk classification of the “Praxistest Diabetes” (PraDi) population according to German Diabetes Risk Score (GDRS) and FINDRISK

<table>
<thead>
<tr>
<th>Risk categories German Diabetes Risk Score (GDRS)</th>
<th>Healthy (FBG &lt; 5.6 mmol/L)</th>
<th>Prediabetes/Diabetes (FBG ≥ 5.6 mmol/L)</th>
<th>Healthy (FBG &lt; 7 mmol/L)</th>
<th>Diabetes (FBG ≥ 7 mmol/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>low (0–29 points)</td>
<td>133</td>
<td>8</td>
<td>141</td>
<td>0</td>
</tr>
<tr>
<td>still low (30–39 points)</td>
<td>94</td>
<td>20</td>
<td>114</td>
<td>0</td>
</tr>
<tr>
<td>increased (40–49 points)</td>
<td>58</td>
<td>30</td>
<td>85</td>
<td>3</td>
</tr>
<tr>
<td>high (50–59 points)</td>
<td>14</td>
<td>22</td>
<td>33</td>
<td>3</td>
</tr>
<tr>
<td>very high (&gt; 59 points)</td>
<td>11</td>
<td>13</td>
<td>23</td>
<td>1</td>
</tr>
</tbody>
</table>

Tab. 3: Classification of the “Praxistest Diabetes” (PraDi) population into the risk categories of the German Diabetes Risk Score (GDRS) and FINDRISK depending on the case status (healthy vs. prediabetes/diabetes, healthy vs. diabetes)

FBG = fasting blood glucose
The investigation of the PraDi study in the context of application of the diabetes risk scores GDRS and FINDRISK within the framework of health checkups showed that the 5-year risk for developing type 2 diabetes was generally associated with the measured fasting blood glucose. This also yielded in a good predictive performance for the identification of undiagnosed prediabetes or diabetes. In this regard, the predictive performance of the GDRS was more precise when compared to FINDRISK, albeit the difference for undiagnosed diabetes was not significant. Results of the PraDi study showed that participants of health check-ups have a more favorable diabetes risk profile and therefore, more than 50% were allocated to the lower diabetes risk groups. A comparison of risk categorization by the two tests showed an agreement of almost 50%, however, nearly 40% of the participants were classified into a higher risk category by the GDRS than classified by FINDRISK. In line with this, also the proportion of participants classified into at least elevated risk was higher for the GDRS (36.6%) than for FINDRISK (23.3%). Consequently, the sensitivity of the GDRS was considerably higher and all diabetes cases were classified into at least elevated risk with the GDRS. For FINDRISK, sensitivity was only 78%. For the identification of prediabetes/diabetes a similar picture could be observed, meaning that with FINDRISK a much higher proportion of participants was allocated into lower risk categories. Finally, the ROC analysis resulted generally in a more precise predictive performance for the GDRS.

A direct comparison of the two tests is so far lacking, however, in a previous study a similar association of the GDRS with FBG was observed in a subcohort of the EPIC-Potsdam study in which also the predictive performance for an impaired FBG was similar with a ROC-AUC of 0.79 [19]. In the underlying investigation of the PraDi study, diabetes cases were included in the prediction, however this proportion is negligible because only 7 diabetes cases were observed and we assume a limited impact on the overall predictive accuracy in the group of prediabetes cases. For the prediction of undiagnosed diabetes Schulze et al. observed in a cross-sectional study a predictive performance of 0.83 (Tübingen Family Study) and 0.75 (Metabolic Syndrome Berlin Potsdam) [5]; results from our study are between these measures. Differences might be explained by the different populations and the different settings. In German cross-sectional studies, FINDRISK was also already evaluated for the prediction of undiagnosed diabetes [11, 20]. Among people with an adverse risk profile, the predictive accuracy was 0.81 [20] and in a population-based epidemiologic study the predictive accuracy was 0.65 [11]; predictive performance observed in the PraDi study was thereby similar or higher. The different predictive accuracy might be surprising because many

---

**Tab. 4:** Sensitivity, specificity, negative and positive predictive values of German Diabetes Risk Score (GDRS) and FINDRISK at prediabetes/diabetes

<table>
<thead>
<tr>
<th></th>
<th>Prediabetes/Diabetes</th>
<th>Diabetes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Se %</td>
<td>Spe %</td>
</tr>
<tr>
<td>DRT (≥ 40 points)</td>
<td>70</td>
<td>73</td>
</tr>
<tr>
<td>FINDRISK (≥ 12 points)</td>
<td>41</td>
<td>82</td>
</tr>
</tbody>
</table>

NPV = negative predictive value; PPV = positive predictive value; Se = sensitivity; Spe = specificity

---

**Fig. 2:** Receiver-operating-characteristic (ROC) curves and corresponding areas under the curves for German Diabetes Risk Score (GDRS) and FINDRISK to predict prediabetes/diabetes (fasting blood glucose ≥ 5.6 mmol/L)

ROC-AUC GDRS: 0.78 (95%-CI: 0.73–0.83)
ROC-AUC FINDRISK: 0.73 (95%-CI: 0.68–0.79)

CI = confidence interval
risk factors of the two tests are similar. However, FINDRISK additionally includes information on family history of diabetes which is a strong risk factor and was included in an updated version of the GDRS [10, 21] but was not available at the time the PraDi study was conducted. Due to the improvement in prediction of future diabetes risk we would assume that the updated version, consequently, would identify undiagnosed prediabetes or diabetes more precisely than the investigated original version. For FINDRISK, questions from the Finnish language were translated, partly extended and adapted to German habits [7] but in contrast to the GDRS, it was so far not validated for prediction of future diabetes in a German prospective study. Still, it remains unclear whether this might explain the difference in predictive accuracy when compared to the GDRS.

Strengths and limitations

To our knowledge, this is the first study investigating the predictive accuracy for identification of undiagnosed prediabetes or diabetes which directly compared the application of the GDRS and FINDRISK for use in the German context. The results of this study might have an important contribution to diabetes prevention in Germany. Another strength of this study is the application in the setting of health checkups in the general practice. From this, a direct starting point for diabetes prevention evolves. Indeed, the selected general practices do not represent all general practices in Germany but this should not have influenced the comparison between the GDRS and FINDRISK.

A limitation of this study was the low number of undiagnosed diabetes cases. A statistical analysis based on 7 cases is limitedly valid and results need to be interpreted with caution. Due to this, we performed a power calculation which described the validity of the comparison of ROC-AUCs. Additionally, our results are only appropriate for undiagnosed prediabetes/diabetes, diagnosis based on FBG. Conclusions are therefore not transferrable to parameters such as HbA1c or values from oral glucose tolerance tests. Such investigations would be restricted to patients with at least elevated diabetes risk.

Impact of the results and transportability

The identification of undiagnosed diabetes cases and high-risk patients is of high importance for an early therapeutic intervention. Health examinations are an important instrument in this context currently including in Germany the diagnosis of a potential type 2 diabetes [22].

Based on the findings from this study, stepwise screening scenarios are possible in which the GDRS is the preceding step of further diagnosis using FBG, HbA1c, or oral glucose tolerance tests as recommended in the current DDG guidelines [23]. The latter two parameters, however, were not investigated in this study. Studies indicate that the acceptability and participation for screening with connected blood tests increase if the risk was identified with a preceding risk score [24]. Another advantage of the combination of risk scores with measurement of glucose parameters is the strong improvement of prediction of future diabetes than what is possible with the identification of prediabetes using solely glucose measures [1, 25].

In a recent study [26], no impact on behavioral change was observed when a risk score was applied, however, this cannot be achieved only by application of risk scores. In this study, the risk score did not include modifiable lifestyle related risk factors which are directly related to health recommendations. Risk scores such as the GDRS, in contrast, include suggestions on modifiable risk factors which cannot directly be derived from glucose measurements. Based on this, recommendations on behavioral changes which follow the risk score can directly be linked to the obtained risk factors in the test.

Future research

On the basis of these study findings, it would be interesting to evaluate the application of the GDRS and a stepwise screening approach with different glucose parameters Germanwide in a representative population. Furthermore, the effectiveness of such a screening might be investigated with subsequently following interventions for high-risk individuals and undiagnosed diabetes patients.

Conclusions

Overall, application of diabetes risk scores in the context of health checkups in the general practice as partly implemented in practice guidelines seem to be advantageous. For an early identification of undiagnosed diabetes cases and for patients with a high risk of future diabetes, the GDRS might be especially appropriate due to the generally better predictive performance when compared to FINDRISK.

This study was supported by the Federal Ministry of Education and Research (BMBF) und the state Brandenburg (DZD promotion 92DZD00302).

Acknowledgements

We thank the participants of the PraDi study and the participating general practices including physician’s assistants and physicians for the conduct of the PraDi study.

Conflict of Interest

The authors declare no conflict of interest.
References


DOI: 10.4455/eu.2018.041