Introduction

Together with population growth and ageing, both the decrease in physical activity and increase in obesity have led to a significant increase in the number of adults with diabetes worldwide: from 108 million in 1980 to 422 million in 2014 [1]. It has also been estimated that approximately 40% of people with diabetes are unaware of their disease [2], and that the mean time between diabetes development and its diagnosis is approximately 5 years. In addition, and mainly in those with type 2 diabetes (T2D), the silent development of microvascular and macrovascular complications frequently arises, yet it may be prevented or delayed by early screening of high-risk individuals [3].

A variety of approaches have been proposed to identify individuals with an increased risk of diabetes. The Finnish Diabetes Risk Score (FINDRISC) is currently the most validated and widely used [4]. This simple, safe and inexpensive screening test (see
Methods below) was developed to predict drug-treated T2D, and to increase awareness of the modifiable risk factors and benefits of a healthy lifestyle [5].

Skin autofluorescence (SAF) has also been proposed as a useful cost-effective, simple and reproducible screening test for diabetes [6]. SAF is a clinical tool that non-invasively assesses advanced glycation endproducts (AGEs) in the skin of the forearm [7,8]. AGEs are formed by non-enzymatic modification of proteins, lipids and nucleic acids by reducing sugars and reactive carbonyl compounds [9,10]. The accumulation of AGEs is seen during general ageing in healthy individuals, but happens at an advanced rate in people with impaired renal function, inflammatory disease and/or diabetes as a result of oxidative and glycaemic stress [11,12]. In addition to the growing evidence from clinical studies that AGEs and SAF serve as potential biomarkers for diabetic complications [11,13], SAF has been shown to be comparable or superior to HbA1c and fasting plasma glucose in the detection of diabetes in intermediate-risk groups. However, it has also been pointed out that establishing the predictive value of SAF in lower-risk groups would still be useful.

The aim of the present study was to evaluate the combined performance of the FINDRISC and SAF in detecting undiagnosed diabetes. Furthermore, it was investigated whether a simplified model could be devised that would have a similar performance to the full model. Finally, the study also focused on the identification of the optimal cut-off values of the FINDRISC (+SAF) for the general Dutch population.

**Methods**

**Study population**

Subjects included in the present study were participants in the LifeLines Cohort Study [16], a large prospective population-based cohort in the northern part of the Netherlands. LifeLines was established as a resource for research into the complex interactions between genomic, phenotypic and environmental factors in the development of chronic diseases and healthy ageing. At baseline (2006–2013), approximately 167,000 participants completed extensive questionnaires, physical examinations and the collection of biomaterials [17]. All participants provided written informed consent before participating in the study, which was approved by the Medical Ethics Review Committee of the University Medical Center Groningen.

**Case definition**

For the present cross-sectional analysis, adults for whom SAF measurements and plasma glucose or HbA1c values were available (n = 81,286) were evaluated. For the vast majority, two glycaemic measurements had been taken (plasma glucose and HbA1c); for 0.5% of participants, only one glycaemic measurement was available. Individuals who were pregnant (n = 58), who were using systemic glucocorticoid therapy (n = 345) or who had previously been diagnosed with type 1 (T1D; n = 182) or type 2 (T2D; n = 14,553) diabetes were excluded. As SAF measurements may be influenced by the use of skin products [18], patients who had used sunscreen prior to SAF measurement were also excluded. Participants were classified as having screen-detected diabetes if they had fasting blood glucose ≥ 7.0 mmol/L, non-fasting blood glucose ≥ 11.1 mmol/L or HbA1c ≥ 6.5% (48 mmol/mol).

**Skin autofluorescence**

SAF was assessed in a subset of participants in the LifeLines cohort using the AGE Reader (DiagnOptics Technologies B.V., Groningen, Netherlands). This non-invasive desktop device uses the characteristic fluorescent properties of certain AGEs to estimate the level of AGE accumulation in the skin. Technical details concerning the optical technique have been extensively described elsewhere [19].

**Questionnaires and physical examination**

Extensive baseline questionnaires included questions on demographic details, medical history and medication use. Weight was measured to the nearest 0.1 kg and height to the nearest 0.5 cm by trained technicians using calibrated measuring equipment, with participants wearing light clothing and no shoes. Body mass index (BMI) was calculated as weight divided by height squared (kg/m²). For the present study, participants were classified as having a history of high blood glucose if they reported previous T2D, previous gestational diabetes or previous diabetes caused by another medical condition.

**Finnish Diabetes Risk Score**

The original publication of the FINDRISC included two models: a concise model and a full model. Because it was reported that the full model only minimally improved prediction, it was decided to use the concise model for the present analysis. Variables in the concise FINDRISC model are age, BMI, waist circumference, use of antihypertensive agents and history of high blood glucose [5].

In addition, multiple imputations by chained equations were applied to eliminate missing values for BMI (0.02%), waist circumference (0.02%), use of antihypertensive agents (1.3%) and history of high blood glucose (0.2%), resulting in five imputed datasets.

**Biochemical measures**

Blood was collected in the fasting state between 08h00 and 10h00 in the morning, and transported to the LifeLines laboratory facility at room temperature or at 4 °C, depending on the sample requirements. On the day of collection, HbA1c (EDTA-anticoagulated) was analyzed using a National Glycohemoglobin Standardization Program (NGSP)-certified turbidimetric inhibition immunoassay on a COBAS Integra 800 CTS analyzer (Roche Diagnostics Nederlands B.V., Almere, Netherlands). Fasting and non-fasting plasma glucose was measured using a hexokinase method. For the present analysis, blood was collected in a non-fasting rather than fasting state in 1.6% of participants.

**Statistical analysis**

All analyses were performed using SPSS version 22 (IBM Corp., Armonk, NY, USA) and R 3.1.3 (R Foundation for Statistical Computing, Vienna, Austria) software. Data are presented as means ± standard deviation (SD) or as number of participants and percentages. Student’s t test or Chi² test was performed to compare groups. Receiver operating characteristic (ROC) curve analysis was used to define groups for SAF. Cut-off values were chosen based on their 50%, 90% and 97.5% specificity for detecting diabetes.

Logistic regression analysis was performed to determine the association between FINDRISC variables, SAF and diabetes, and also to determine the point-score systems for the SAF and FINDRISC variables, as well as the simplified models. Regression coefficients were multiplied by 4 and rounded off for simplicity, thereby allowing comparisons of these score values with those of FINDRISC. As the five imputed datasets showed comparable regression coefficients and standard errors in the prediction of diabetes, the imputation variation was low. Therefore, one randomly chosen imputed dataset was used for further analyses.
The discriminatory ability of the different models was estimated using the area under the ROC (AUROC) curve. Calibration was on clinically relevant risk categories and defined as: with the actual risk observed in each group. Risk groups were based on the present regression analysis) were directly compared in risk reclassification analyses; in which predicted risk estimates individuals into higher and lower risk categories was assessed using. Accurate stratification of observed and predicted outcomes (based on the present regression analysis) over deciles of risk [20]. Accurate stratification of observed and predicted outcomes (based on the present regression analysis) was assessed with the Hosmer-Lemeshow Chi 2 test by comparing using the area under the ROC (AUROC) curve. Calibration was on clinically relevant risk categories and defined as:

### Results

The demographic and clinical characteristics of the study population are presented in Table 1 for participants with and without diabetes. Overall, the age range was 18–92 years, and the prevalence of undiagnosed diabetes was 1.3%. Those with diabetes were significantly older, had higher BMI scores, larger waist circumferences and higher SAF values. Also, a greater proportion of participants with diabetes was male, was using antihypertensives and had a history of high blood glucose.

**FINDRISC and SAF in detecting diabetes**

Logistic regression analysis showed that all variables of the FINDRISC were significant detectors of diabetes in our study cohort independently of covariates (Table 2). Moreover, the addition of SAF [in categories of arbitrary units (AU) < 1.9, 1.9–2.4, 2.5–2.9 and > 2.9] contributed significantly to the model. People in the highest SAF category had a 3.5-fold increased risk of having diabetes compared with those in the lowest SAF category. To build the FINDRISC + SAF model, the calculated risk scores for SAF were added to the original risk scores of the FINDRISC. Estimates (including calibration) for the different regression models are presented in the footnotes of Table 2. The ROC curve for the FINDRISC model yielded an AUC of 0.802 (95% CI: 0.789–0.815) for detection of diabetes, whereas the addition of SAF significantly improved the discriminatory value of the model to 0.811 (95% CI: 0.798–0.824; P < 0.001).

Table 3 shows the number and percentage of participants initially classified by FINDRISC into four risk categories who were, in fact, reclassified into higher- or lower-risk categories with the addition of SAF. The proportion of participants reclassified was low (< 10%) among those originally classified as having < 1% or > 10% risk of diabetes, and high (> 30%) among those in the intermediate-risk categories. Overall, 15% of all participants were reclassified with the addition of SAF. As shown by the observed diabetes prevalence, most of these participants were reclassified into more accurate risk categories. Reclassification improved by 8.3% in those who had diabetes and worsened by 0.2% in those without diabetes, resulting in a net reclassification of 8.0% (95% CI: 5.2–11.0).

**Simplified model and SAF for detecting diabetes**

ROC curve analysis revealed that a simplified model consisting of age and BMI (using the adjusted point-scores shown in Table II)

### Table 1

Characteristics of the study population.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Diabetes</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes (n=1042)</td>
<td>No (n=78,206)</td>
</tr>
<tr>
<td>Age, years</td>
<td>55 ± 12</td>
<td>44 ± 12</td>
</tr>
<tr>
<td>Male gender, n (%)</td>
<td>561 (53.8)</td>
<td>324 (41.5)</td>
</tr>
<tr>
<td>Body mass index, kg/m²</td>
<td>30.0 ± 5.2</td>
<td>26.0 ± 4.2</td>
</tr>
<tr>
<td>Waist circumference, cm</td>
<td>Female 101 ± 14</td>
<td>87 ± 12</td>
</tr>
<tr>
<td>History of high blood</td>
<td>Male 106 ± 13</td>
<td>95 ± 10</td>
</tr>
<tr>
<td>glucose, n (%)</td>
<td>Use of antihypertensives drugs, n (%)</td>
<td>393 (37.7)</td>
</tr>
<tr>
<td>Skin autofluorescence, arbitrary units</td>
<td>12 (1.2)</td>
<td>172 (0.2)</td>
</tr>
</tbody>
</table>

Data are presented as means ± standard deviation unless otherwise specified.

### Table 2

Multivariable logistic regression models for detection of undiagnosed type 2 diabetes in the LifeLines cohort (n=79,248).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>FINDRISC model</th>
<th>FINDRISC model + SAF</th>
<th>Simplified model + SAF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Regression coefficient</td>
<td>OR (95% CI)</td>
<td>Point-score</td>
</tr>
<tr>
<td>Age, years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 45</td>
<td>0.725</td>
<td>2.06 (1.72–2.48)</td>
<td>2.041</td>
</tr>
<tr>
<td>45–55</td>
<td>1.349</td>
<td>3.85 (3.18–4.67)</td>
<td>1.030</td>
</tr>
<tr>
<td>56–65</td>
<td>1.757</td>
<td>5.80 (4.72–7.12)</td>
<td>1.328</td>
</tr>
<tr>
<td>Body mass index, kg/m²</td>
<td>0.581</td>
<td>1.79 (1.44–2.22)</td>
<td>1.571</td>
</tr>
<tr>
<td>SAF, arbitrary units (AU)</td>
<td>0.408</td>
<td>1.50 (1.17–1.94)</td>
<td>0.409</td>
</tr>
<tr>
<td>Use of antihypertensives</td>
<td>0.716</td>
<td>2.05 (1.58–2.65)</td>
<td>0.712</td>
</tr>
<tr>
<td>Use of antihypertensives drugs, n (%)</td>
<td>0.653</td>
<td>1.92 (1.67–2.22)</td>
<td>0.602</td>
</tr>
</tbody>
</table>

Nagelkerke R²: 0.135 for Finnish Diabetes Risk Score (FINDRISC) model; 0.144 for FINDRISC model + skin autofluorescence (SAF); 0.132 for simplified model + SAF; Model χ²: 1409 (P < 0.001) for FINDRISC model; 1505 (P < 0.001) for FINDRISC model + SAF; 1383 (P < 0.001) for simplified model + SAF; Hosmer–Lemeshow χ²: 16.2 (P = 0.024) for FINDRISC model; 11.7 (P = 0.113) for FINDRISC model + SAF; 16.8 (P = 0.019) for simplified model + SAF.

Scores as published by Lindström et al. (Diabetes Care, 2003; 26(3):725–31) [5].
had an AUC close to that of the FINDRISC model (0.795, 95% CI: 0.781–0.808; P = 0.035). The addition of SAF to age + BMI further improved the discriminatory value of this model to 0.806 (95% CI: 0.793–0.818; P < 0.001). The AUC of the simplified model + SAF was not significantly different from that of either the FINDRISC (P = 0.365) or FINDRISC + SAF (P = 0.062) model. Reclassification of participants due to the addition of SAF to age + BMI is shown in Table 4. Overall, 10% of all participants were reclassified into more accurate risk categories. However, reclassification worsened by 0.4% among those who had diabetes, but improved by 8.2% among subjects without diabetes, resulting in a net reclassification of 7.8% (95% CI: 5.7–9.9).

Prevalence and cut-off points

With the FINDRISC (point range: 0–21), the prevalence of diabetes likewise rises as the score increases (Fig. 1B). Test characteristics of this model are presented in Table S2 (see supplementary materials associated with this article online) for a range of cut-off values. At a score of 9, the sum of sensitivity and specificity is at maximum (sensitivity 75%, specificity 75%).

With the FINDRISC model + SAF (point range: 0–21), the prevalence of diabetes also rises as the score increases (Fig. 1C). Test characteristics of this model are presented in Table S3 (see supplementary materials associated with this article online) for a range of cut-off values. At a score of 8, the sum of sensitivity and specificity is at maximum (sensitivity 83%, specificity 65%).

The optimal cut-off points for the FINDRISC, FINDRISC + SAF and simplified + SAF models did not differ by gender (data not shown).

Discussion

Our present study has shown that SAF improved performance of the FINDRISC model, an accepted tool for diabetes detection, in a large and recently recruited population cohort. Furthermore, a simplified model (comprising age and BMI) in combination with SAF with re-estimated point-scores was shown to have a similar
performance for diabetes detection compared with the FINDRISC model with SAF added.

Discrimination, calibration and reclassification criteria were used to assess the significance of the addition of SAF as a new biomarker to the list of conventional diabetes risk factors. A significant improvement in discrimination was found with the addition of SAF to the FINDRISC model. Calibration also showed a better fit for the FINDRISC model + SAF compared with the FINDRISC alone. More importantly, significant reclassification was demonstrated with the addition of SAF, with an overall NRI index of 8%. This new model was especially useful for reclassifying participants in the intermediate-risk categories, in which > 30% of subjects were reclassified.

Although calibration of the simplified model + SAF was poor (Hosmer–Lemeshow $\chi^2 = 16.2, P = 0.024$), the discriminatory value did not differ significantly from that of the FINDRISC model + SAF. Therefore, the simplified model + SAF may well represent an accurate alternative for settings where answers concerning medication use and history of high blood glucose may be unreliable or inappropriate: this includes screening not only in medical settings such as general practitioner practices and pharmacies, but also in supermarkets and at major public events, thereby reaching those who avoid, or have never been exposed to, healthcare.

Moreover, because it is easy and quick to perform, the simplified model + SAF may even improve the general public’s participation in screening programmes, which is generally low [23].

Previous studies have addressed the value of skin fluorescence in the detection of diabetes. In those at risk of diabetes, skin fluorescence was comparable or superior to HbA1c and fasting plasma glucose for the detection of impaired glucose tolerance and diabetes detected by Oral Glucose Tolerance Tests (OGTTs) [14,24]. To further improve the sensitivity and specificity of diabetes detection, an SAF-based decision tree was developed. Indeed, this decision tree has proved to be as good as or superior for the detection of diabetes and impaired glucose tolerance in comparison to conventional risk predictors in an intermediate-risk group [15]. However, no previous reports have validated skin fluorescence for the detection of diabetes in lower-risk groups.

The FINDRISC is often used as the first assessment tool to identify those who may need further glucose testing. It was found that the optimal cut-off value was a score $\geq 7$ with the FINDRISC model (sensitivity 79%, specificity 68%), $\geq 9$ with the FINDRISC model + SAF (sensitivity 75%, specificity 75%) and $\geq 8$ for the simplified model + SAF (sensitivity 83%, specificity 65%). However, as diabetes prevalence was rather low in the LifeLines population, higher cut-off values may be more cost-effective in Dutch and other populations [25].

Compared with the original publication, the present study’s AUROC for FINDRISC was somewhat lower (0.857 vs. 0.802, respectively). It is known that the performance of risk models is generally higher in the population for which they were designed. In addition, the present study differs from the original Finnish study...

Fig. 1. Incidence of undiagnosed diabetes for different scores of (A) the FINDRISC score, (B) the FINDRISC + SAF score and (C) the simplified model + SAF score; black bars, observed risk; grey bars, expected risk.
with respect to population characteristics as well as the way that diabetes was identified. The discriminatory performance of the FINDRISC in our present cohort was somewhat higher than in other validation studies addressing the performance of FINDRISC in the detection of undiagnosed diabetes. In those studies, the AUROC ranged from 0.70 to 0.78 [26–31]. The difference may be explained by the wider distribution of risk factors in our large cohort study. In fact, as an illustration of this, ROC curve analysis for participants in the age range for which the FINDRISC was originally designed (35–64 years) revealed a significantly lower AUC (results not shown).

Nowadays, screening programmes for early detection of (pre)diabetes are lacking in most countries, despite the fact that several (systematic) reviews performed over the last 5 years have concluded that screening programmes for diabetes are efficient and cost-effective [3,32,33]. On the other hand, glycaemic tests (fasting plasma glucose, HbA1c, OGTT) have been poorly taken up for diabetes screening by the general population, and the most promoted questionnaire-based risk score (the FINDRISC) is still little used. Yet, the model described here, comprising age, BMI and SAF (no fasting required), can be used as a diabetes-screening tool in non-medical settings, with medical referral of only a selection of high-risk patients for confirmatory glycaemic tests, thereby adding to the cost-effectiveness of such screening.

Our present analysis demonstrates that SAF improves the risk classification of individuals at high-risk of diabetes, which opens up the prospects for studies (some already ongoing) of related applications. An important opportunity to address is the possible value of SAF alone or in combination with conventional models for the prediction of diabetes. Moreover, it would be relevant to investigate whether the addition of SAF to the FINDRISC model might more accurately predict the cardiovascular complications of diabetes and cardiovascular disease. The FINDRISC has already proved invaluable for the prediction of cardiovascular events in the population for which it was developed [34], as well as in a randomly selected Finnish population of men aged 45–74 years [35]. Likewise, SAF has proved to be highly useful for the prediction of cardiovascular events in diabetic populations [13,36] and in patients with peripheral artery disease [37]. Finally, it is important to confirm the validity of our present results in other populations, especially non-Caucasians.

One limitation of our study is the identification of diabetes cases using just a single measurement of (fasting) plasma glucose and HbA1c rather than the repeated measurements required for a clinical diagnosis or with OGTTs. However, the main disadvantage of the gold standard OGTT would be the lower rate of participation. Second, a history of high blood glucose was not directly addressed in the LifeLines questionnaire, as participants with such a history were identified only if they filled in a free space following the question about having diabetes. This has probably led to underestimation of the number of patients with a history of high blood glucose. Nevertheless, as the prevalence of such a history was rather low (0.7–1.6%) in other Dutch cohorts in which the FINDRISC was validated [38], this is unlikely to have changed our main results. Finally, a family history of diabetes was incomplete in the LifeLines population, as the corresponding question was only later integrated into the baseline questionnaire. Thus, it was not possible to compare the performance of FINDRISC (+SAF) with the updated FINDRISC [39] and other available risk scores [40–42].

In conclusion, the combination of a simple-to-perform FINDRISC score plus the easy, non-invasive SAF measurement provides an invaluable clinical tool for diabetes detection. This new model is especially useful for reclassifying people in intermediate-risk categories for whom further blood glucose testing should confirm the presence of diabetes. Furthermore, a simplified model including age, BMI and SAF may represent an accurate alternative screening tool for settings where answers about medication use and history of high blood glucose may be unreliable or inappropriate. Indeed, the resultant early identification of patients with diabetes may prevent or delay the development of microvascular and macrovascular complications by enabling intervention at an early stage of the disease.

Funding

LifeLines has been funded by a number of public sources, notably the Dutch government, Netherlands Organization for Scientific Research (NWO; grant 175.010.2007.006), Northern Netherlands Collaboration of Provinces (SNN), European Regional Development Fund (ERDF), Dutch Ministry of Economic Affairs (Piekien in de Delta), the provinces of Groningen and Drente, the Target project, BBMRI-NL, and the University of Groningen and University Medical Center Groningen, Netherlands. The present work was supported by the Netherlands Consortium for Healthy Ageing (NCHA) and by funds from the European Commission’s Seventh Framework Programme (FP7/2007–2013) through BioSHaRE-EU (Biobank Standardisation and Harmonisation for Research Excellence in the European Union) project (grant agreement 261433). The funding organizations had no role in the data analysis, data interpretation or final report of this study.

Trial registration

LifeLines (BRIF4568) is engaged in a Bioresource Research Impact Factor (BRIF) policy pilot study, the details of which can be found at: http://www.bioshare.eu/content/lifelines-cohort-study-biobank.

Author’s contributions

B.T. Fokkens: study design, data analysis, data interpretation, writing. R.P. van Waateringe: study design, data analysis, data interpretation, revising manuscript. D.J. Mulder: data interpretation, revising manuscript. B.H.R. Wolffenbuttel: study design, data interpretation, revising manuscript A.J. Smit: study design, data interpretation, revising manuscript.

Availability of data and material

The manuscript is based on data from the LifeLines Cohort Study. LifeLines adheres to the standards for data availability. The LifeLines data catalogue is publicly accessible on http://www.lifelines.net. All international researchers can apply for data at the LifeLines research office (Llscience@umcg.nl). The LifeLines system allows access for reproducibility of study results.

Disclosure of interest

Prof. Dr. A.J. Smit is founder and shareholder of DiagnOptics Technologies B.V., the company that developed the autofluorescence reader used in our study to assess skin accumulation of AGEs. The other authors declare that they have no competing interest.

Acknowledgements

The authors wish to thank all participants in the LifeLines Cohort Study, and all the people involved in the setting up and implementation of the study.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version, at http://dx.doi.org/10.1016/j.diabet.2017.09.002.