Evaluating risk for cardiovascular diseases—vain or value? How do different cardiovascular risk scores act in real life

Eeva Ketola¹, Tiina Laatikainen², Erkki Vartiainen²

Background: Screening tools to identify persons with high cardiovascular risk exist, but less is known about their validity in different population groups. The aim of this article is to compare the sensitivity and specificity of three different cardiovascular disease risk scores and their ability to detect high-risk individuals in daily practice. Methods: The sensitivity and specificity of risk charts based on Framingham Risk Function, SCORE and cardiovascular disease (CVD) Risk Score were analysed using a large population risk factor survey database in Finland. For different cardiovascular disease end-points in 10-year follow-up true positive, false positive, true negative and false negative cases were identified using different risk charts. Subjects over 40 years (n = 25 059) of the FINRISK Study were used in analyses. Results: Risk scores differed depending on gender, age and cardiovascular outcome. Among men the sensitivity of CVD Risk Score and Framingham Risk Function at risk of ≥10% for each end point was higher than of SCORE or Framingham Risk Function at risk of 20%. The specificity of Framingham Risk Function at risk of 20% was higher than the specificity of other risk charts. Among women in all endpoints the sensitivity was highest in CVD Risk Score and lowest in Framingham Risk Function at risk of ≥20%. Specificity for all different endpoints was highest in SCORE and Framingham Risk Function at risk of 20%. Conclusions: Sensitivity and specificity varied markedly in between three cardiovascular risk evaluation tools. Practitioners should be aware of their limitations especially when estimating risk among women and younger patients.

Keywords: cardiovascular disease, risk score, mortality, sensitivity, specificity.

Introduction

Risk evaluation for cardiovascular disease (CVD) is present in everyday practice of family physicians. This might start an important step in reducing the risk of CVD by effective treatment or be in vain, causing unnecessary worries and useless measurements and laboratory costs. Anyway, the ageing population increases the CVD disease burden and the workload related to cardiovascular diseases and their complications increases. Targeting the resources is essential. Early detection of high-risk individuals and effective working practices in primary and secondary prevention are a corner stone of high quality of care. A good-risk evaluation tool should be sensitive enough to detect high-risk individuals for self-care and nurse led lifestyle interventions and on doctors appointment specific enough to assist in selection of the most effective treatment to high-risk individuals.

There are several different widely used CVD risk evaluation tools based on different local and international population data. Some surveys have been carried out on the external validity of these risk functions. However, very little information exists how these tools fit for clinical practice. Knowing the sensitivity and specificity of different CVD risk evaluation tools might help in using these screening tools. The aim of this study is to compare the sensitivity and specificity of three different cardiovascular disease risk scores as a tool to detect high-risk individuals and evaluate how they fit in daily practice.

Methods

The study population was participants of the National FINRISK Study from years 1972, 1977, 1982, 1987 and 1992. The survey covers the age range from 25 to 64 years. Only participants aged ≥40 years (lowest age range common to all risk scores), were used in the analyses. Participants were followed up for 10 years using the Causes of Death Register and the Finnish Hospital Discharge Register. Regarding cardiovascular diseases Finnish population, especially men, is a high risk population. The CHD mortality among 35- to 64-year-old men in 2001 (which is the last year of the follow-up in these analyses) was 123/100 000 and among women 20/100 000. Despite of major reduction in cardiovascular risk factors during the last decades the cholesterol and blood pressure levels are still considerably high among the Finnish population. Smoking rates are quite low especially among women.

The outcome measures were deaths and events attributable to cardiovascular diseases and more specifically deaths and events attributable to ischaemic heart diseases and stroke. Causes of deaths were classified using the ICD-9 coding until the end of year 1995 and ICD-10 from the beginning of year 1996 as the codes have been in use in Finland. Both underlying and immediate cause of death were taken into account. For cardiovascular diseases, ICD-9 codes 390–459 and 798 and ICD-10 codes I00-I99, R96 and R98 were used for mortality endpoints. Both hospitalizations due to acute cardiovascular events [ICD-9 codes 410, 430, 431, 432, 4330A, 4331A, 4339A, 4340A, 4341A, 4349A, 4376A, 436 and ICD-10 codes I21, I22, I60, I61, I62, I63 (except I636), I64] and deaths attributable to cardiovascular diseases were included in classification of cardiovascular events. ICD-9 codes used for ischaemic heart disease mortality were 410–414 and 798 and ICD-10 codes
for non-fatal ischaemic heart disease events the hospitalizations registered with ICD-9 codes 410 and ICD-10 codes I21 and 122 were used. Ischaemic heart disease events category includes both fatal and non-fatal events. For strokes only a category including both fatal and non-fatal strokes was classified. For hospitalizations attributable to stroke ICD-9 codes 430–432, 4330A, 4331A, 4339A, 4340A, 4341A, 4349A, 4376A, 436 and ICD-10 codes I21, I22, I60, I61, I62, I63 (except I636), I64 were used and for deaths ICD-9 codes 430–438 and ICD-10 codes I60–I69.

Risk scores compared in these analyses were the Framingham Risk Function,7,8 SCORE9 and CVD Risk Score.10–12 Framingham Risk Function is based on the Framingham cohort and predicts the risk for non-fatal and fatal coronary events. SCORE is based on combined cohorts of different European countries and predicts the risk for cardiovascular disease mortality. CVD Risk Score is a modified version of the North Karelia project risk score developed in Finland to detect individuals with high cardiovascular risk (Supplementary Data).12

In the National FINRISK Study, several cardiovascular disease risk factors were measured from each participant. Factors used in all risk estimation functions are serum cholesterol, systolic blood pressure and smoking. In addition to these, BMI, amount of smoked cigarettes and leisure time physical activity are included in CVD Risk Score. Age is taken into account in Framingham Risk Function and SCORE. In the CVD Risk Score same tables are used for both gender, but in SCORE and Framingham the risk charts are separate for men and women.

Risk factor information was fully available for altogether 11 900 men and 13 159 women used in these analyses. The measured risk factor values of each participant in the National FINRISK Study were inserted in each of the selected risk estimation functions using the risk charts with SCORE and Framingham Risk Function and risk factor related point scoring in CVD Risk Score (Table 1). With this method we simulated the practical risk assessment situation at the health professionals’ appointment. We also show with few examples the weaknesses of different risk scores.

In these analyses for SCORE, the risk tables for population at high cardiovascular risk were used. The risk of 5% or higher was used to indicate a high risk person for cardiovascular mortality. Concerning Framingham Risk Function the risk of ≥10% and ≥20% for coronary heart disease events were used. In CVD Risk Score, depending on their risk factor status persons can obtain risk points from zero to sixteen. The cut-off point 4.5 was used to categorize persons at high risk.11

Using different cardiovascular disease end-points in 10-year follow-up true positive, false positive, true negative and false negative cases were identified. The sensitivity and specificity of each risk estimation function to predict cardiovascular disease end-points were assessed with 95% CIs. Logistic regression model was used to analyse how much the different risk factors included in risk estimation functions increased the risk of different CVD end points. Categorization of risk factors used in the model followed the categories of CVD Risk Score. That was because in the CVD Risk Score more risk factors and their categories are used than in other risk estimation functions. In addition to CVD risk factors, age was included in the model. The analyses were done using SAS statistical package, version 8.2 (Cary, NC, USA).

**Ethical issues**

The National FINRISK Study protocols have been approved by the Ethics Committee of the National Public Health Institute and later by the Ethics Committee for the Research in Epidemiology and Public Health. For these analyses no additional ethics approval was required.

**Results**

We detected major differences in sensitivity and specificity in between three cardiovascular risk evaluation tools. They acted also differently in regarding gender and events (IHD and CVD mortality and morbidity and stroke morbidity). According to CVD Risk Score and Framingham Risk Function (at 10% risk) >75% of the Finnish male population aged ≥40 years belong to high-risk group compared with 36% according to SCORE and 25% according to Framingham at 20% risk. Difference in classifying women to high-risk individuals using the different risk estimation tools was even bigger. According to CVD Risk Score 70% of Finnish women over 40 years are at high risk compared with 2% according to Framingham at 20% risk and 7% according to SCORE.

**Men**

Among men the CVD Risk Score and Framingham Risk Function (risk at 10%) was able to find correctly over 90% of true cases for each end-point (sensitivity). In comparison the sensitivity of SCORE varied between 58 and 70% being highest when predicting stroke morbidity in men. If Framingham risk function was used at 20% risk, the sensitivity for each endpoint was ~50%. About 22–30% of true negative cases in men were classified as negative, using CVD Risk Score and Framingham risk function at 10% risk (specificity), respectively. The specificity of SCORE and Framingham at 20% risk were higher than the specificity of other functions (Table 1).

**Women**

Among women the sensitivity of CVD Risk Score was highest compared with other functions in all end-points varying between 81% and 94%. Sensitivity was highest for IHD mortality and lowest for stroke morbidity. Comparatively the sensitivity of Framingham Risk Function at 10% risk was highest for IHD mortality being 80% and lowest for stroke morbidity being 57%. Sensitivity of SCORE and Framingham at 20% risk were among women remarkably less than of other functions being 17 and 4% for stroke, and 27 and 13% for IHD mortality, respectively. Specificity for all different endpoints was about 30% for CVD Risk Score, ~65% for Framingham Risk Function at 10%, ~94% for SCORE and 98% for Framingham at 20% risk. (Table 1)

**Age and risk evaluation**

However, when assessing the sensitivity and specificity in CVD mortality in two different age groups (aged 40–54 years and 55–64 years), the sensitivity of SCORE and Framingham Risk Function (at 10%) was lower in the younger age group compared with the older age group. CVD Risk Score was the most sensitive in the younger age groups both among men and women. Especially in women under 55 years the sensitivity of Framingham and SCORE was very low. According to SCORE none of the women in the FINRISK study population was classified as in high risk (Table 2).

**Risk factors and increase in cardiovascular events**

The factors included in all risk estimation functions, i.e. age, smoking, systolic blood pressure and serum cholesterol increased significantly the risk of different cardiovascular end
points (Table 3). The only exception was serum cholesterol that did not increase the risk of stroke events either among men or women, and among women, in addition, did not increase the risk of CVD mortality. In addition to these BMI increased significantly the risk of all different end points among men and physical inactivity increased the risk of IHD events, CVD mortality and CVD events among women. Though systolic blood pressure was in the model, diastolic blood pressure further increased the risk of IHD and CVD events and stroke.

**Discussion**

In an appointment the physician has many possibilities for preventive activities. In the population strategy the main aim is to reduce the risk factor distribution in the whole population by life style changes. This is mostly done by legislation and health educational activities in mass media, but also physician can promote these ideas in their appointments and encourage patients for self-care. In the high-risk strategy the main target is to screen high risk individuals and offer them special
treatment by supporting life style changes and in many cases optimize treatment by drugs.

Treatments should be started after evaluating total cardiovascular risk based on individual’s risk factor levels. Prevention should be tailored after risk factor clusters, not using single risk factors unless the family history of CVD is strong or if the single risk factor level is grave. However, treatment based on one risk factor may direct treatment also to low risk individual, which might be vain. Usually health professionals direct most effective prevention actions to those at highest risk with same amount of professionals. The consumed time for screening, diagnose and treat both the diseases and their risk factors increase constantly. The resources in health care do not increase in same amount. We have more work to do, with same amount of professionals. New delegation of tasks and self-care are needed. In screening it is important to target both the preventive activities and the treatment to high-risk patients (multiple risk factors on CVD or secondary prevention). These people need heavy intervening. Self-care and lifestyle interventions at nurses’ appointments should be available for those at slightly elevated risk. Our study revealed that the most specific screening instruments should be targeted to heavy (drugs or specialist consultations) and the most sensitive instruments to lighter interventions (self-care, lifestyle interventions, or referral from nurses’ to a doctors’ appointment) and as a preventive measure to raise the...
awareness. The screening instruments should be used carefully and we should be aware what they really detect and predict.

None of the analysed risk factor scales was optimal. Criticism against the available risk charts and their use has also been presented in HUNT2-study. Getz et al. reported, that with available European risk chart based on SCORE risk function 85.9% of Norwegian 40-year-old men were at high risk. It was suggested, that local epidemiological data and size of high-risk population should be estimated and ethical and practical dilemmas related to resource allocation should be discussed before implementation of the guideline. Our results support this finding, because each of the risk charts studied had its limitations. Clinicians should be critical in using them and should understand their limitations and in which situations they are most usable. The reason to evaluate the possible effects of the different clusters of cardiovascular risk factors in each patient should be clear. Is the purpose to screen high risk individuals? If so, are there enough resources to give life-style advice and adequate treatment? Is the purpose to screen high risk individuals? If so, are there enough resources to give life-style advice and adequate treatment? Is the purpose to screen high risk individuals? If so, are there enough resources to give life-style advice and adequate treatment?

Our analyses show, that there are major differences in sensitivity and specificity of different risk estimates to predict CVD events. This is in accordance with earlier validation studies showing that risk charts derived from one population easily either under or overestimate the disease risk in other populations or population groups. It is also obvious, that most of the risk charts are poorly adapted in women and in younger age groups. Due to this, extrapolation of the current risk estimate to 60 years might be miss-leading and overestimate the risk, as shown in HUNT2-study. In our analyses it seemed that single highly elevated risk factor might be higher than the risk estimate received. Weakness of this study is that we have analysed only patients aged ≥40 years but that was the common age group for all risk charts. Accordingly we do not know, how the risk charts would work in younger patients. People under 40 seldom ask their risk and it might be stressful for the patient to be false positive in the test. Another weakness is that the analyses were done using the risk charts, not by the original logistic regression functions, but this is, however, the way clinicians mostly use them.

Risk evaluation in the appointment is usually challenging. The idea of risk might be hard to understand for the patient. It is more clear to say that four out of 10 patients may suffer from an infarction in 10 years than 40% have higher risk of suffering infarction. It is also important to tell, that these methods reveal only estimated risk and are not absolute truths, but tools to describe the possible situations ahead and so to help the patient in motivating oneself to adequate treatment.

As CVD Risk Score visualizes the risk levels one by one, scales are wider than in the two other risk charts and also factors like BMI and physical inactivity, clearly predicting the risk, are included, it might be helpful tool in both selecting persons for preventive activities and in lifestyle counselling. The structure of the CVD Risk Score also helps the patient himself to understand the contribution of possible reductions in each risk factor and acts as tool for self-care. The effect of reduction of single risk factor can easily be calculated and adapted to his or her own lifestyle modification possibilities. Though when using CVD Risk Score, one should keep in mind that the high sensitivity means, that majority of the screened population might get a positive test result. So there must be enough resources to give the lifestyle counselling and adequate treatment to all patients who might need it. In CVD Risk Score it might be relevant to consider using different cut-off points to estimate different levels of risk.

These screening tools are widely used, but none of them is optimal. There were major differences in sensitivity and specificity in between selected cardiovascular risk evaluation tools. They acted also differently in regarding gender and events (IHD and CVD mortality and morbidity and stroke morbidity).

Table 4  Examples of high risk individuals, their outcome and predicted risk using different risk screening tools

<table>
<thead>
<tr>
<th>Sexa</th>
<th>Age</th>
<th>Serum cholesterol (mmol/l)</th>
<th>Blood pressure SBP (mmHg)</th>
<th>DBP (mmHg)</th>
<th>Smoking</th>
<th>BMI (kg/m²)</th>
<th>Physical activity</th>
<th>Cause of death</th>
<th>Diagnosis of event</th>
<th>Risk/points according to different screening tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>41</td>
<td>7.3</td>
<td>134</td>
<td>88</td>
<td>No</td>
<td>33</td>
<td>1/week</td>
<td>410</td>
<td>ICD-9 or ICD-10</td>
<td>Framingham (%)</td>
</tr>
<tr>
<td>F</td>
<td>58</td>
<td>6.1</td>
<td>152</td>
<td>96</td>
<td>No</td>
<td>34</td>
<td>1–2/week</td>
<td>410</td>
<td>ICD-9 or ICD-10</td>
<td>6.5 p</td>
</tr>
</tbody>
</table>

10-year risk of fatal or non-fatal coronary event

<table>
<thead>
<tr>
<th>Sexa</th>
<th>Age</th>
<th>Serum cholesterol (mmol/l)</th>
<th>Blood pressure SBP (mmHg)</th>
<th>DBP (mmHg)</th>
<th>Smoking</th>
<th>BMI (kg/m²)</th>
<th>Physical activity</th>
<th>Cause of death</th>
<th>Diagnosis of event</th>
<th>Risk/points according to different screening tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>45</td>
<td>7.2</td>
<td>158</td>
<td>108</td>
<td>No</td>
<td>29</td>
<td>1–2/week</td>
<td>427</td>
<td>ICD-9 or ICD-10</td>
<td>7.5 p</td>
</tr>
<tr>
<td>F</td>
<td>56</td>
<td>7.7</td>
<td>124</td>
<td>88</td>
<td>No</td>
<td>30</td>
<td>1–2/week</td>
<td>427</td>
<td>ICD-9 or ICD-10</td>
<td>5.5 p</td>
</tr>
</tbody>
</table>

10-year risk of cardiovascular death

<table>
<thead>
<tr>
<th>Sexa</th>
<th>Age</th>
<th>Serum cholesterol (mmol/l)</th>
<th>Blood pressure SBP (mmHg)</th>
<th>DBP (mmHg)</th>
<th>Smoking</th>
<th>BMI (kg/m²)</th>
<th>Physical activity</th>
<th>Cause of death</th>
<th>Diagnosis of event</th>
<th>Risk/points according to different screening tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>56</td>
<td>7.3</td>
<td>128</td>
<td>94</td>
<td>No</td>
<td>40</td>
<td>Never</td>
<td>410</td>
<td>ICD-9 or ICD-10</td>
<td>8 p</td>
</tr>
<tr>
<td>F</td>
<td>58</td>
<td>5.8</td>
<td>142</td>
<td>94</td>
<td>Yes</td>
<td>30</td>
<td>1/week</td>
<td>125</td>
<td>ICD-9 or ICD-10</td>
<td>5.5 p</td>
</tr>
</tbody>
</table>

Table 4. Examples of high risk individuals, their outcome and predicted risk using different risk screening tools

a: M = male, F = female.
Highly specific risk charts, like Framingham Risk Function and SCORE, might be useful in evaluating the intensity of the treatment activities needed. However, it seems that they easily underestimate the risk among women and younger age groups.

A simple risk factor sheet is a valuable tool in delegation of tasks and in self-care. We recommend this tool in teamwork for delegation of tasks and also self-care in patients having light or moderate cardiovascular risk.

In future further research should be targeted in risk analysis of younger patients. More advanced tools for risk analysis are also needed using electronic patient records and electronic decision support devices.

**Supplementary data**

Supplementary data are available at EURPUB online.

**Conflicts of interest:** None declared.

**Key points**

- Practical tools to help the detection of high risk individuals to guide the treatment decisions are clearly needed. Several screening tools to identify persons with high cardiovascular risk are in use. Less is known about the validity of screening tools in different populations and population groups.
- There are major differences in different cardiovascular risk screening functions in detecting high risk individuals.
- Practitioners should be aware of the limitations of different risk screening functions especially when estimating risk among women and younger patients.

**References**


