Performance of Framingham cardiovascular disease (CVD) predictions in the Rotterdam Study taking into account competing risks and disentangling CVD into coronary heart disease (CHD) and stroke

Bob J.H. van Kempen a,b, Bart S. Ferket a,b, Maryam Kavousi a, Maarten J.G. Leening a,e, Ewout W. Steyerberg c, M. Arfan Ikram a,b, Jacqueline C.M. Witteman a, Albert Hofman a, Oscar H. Franco a, M.G. Myriam Hunink a,b,d,*

a Department of Epidemiology, Erasmus MC, Rotterdam, The Netherlands
b Department of Radiology, Erasmus MC, Rotterdam, The Netherlands
c Department of Health Policy and Management, Harvard School of Public Health, Boston, USA
d Department of Epidemiology, Erasmus MC, Rotterdam, The Netherlands
e Department of Public Health, Erasmus MC, Rotterdam, The Netherlands

ARTICLE INFO
Article history:
Received 8 March 2013
Received in revised form 15 November 2013
Accepted 17 December 2013
Available online 27 December 2013

Keywords:
Cardiovascular disease
Risk prediction
Framingham risk score
Rotterdam Study

ABSTRACT

Background: To evaluate the performance of Framingham predictions of cardiovascular disease (CVD) risk corrected for the competing risk of non-CVD death, in an independent European cohort of older individuals and subsequently extend the predictions by disentangling CVD into coronary heart disease (CHD) and stroke separately.

Methods: We used the Rotterdam Study data, a prospective cohort study of individuals aged 55 years and older (N = 6004), to validate the Framingham predictions of CVD, defined as first occurrence of myocardial infarction, coronary death or stroke during 15 years of follow-up, corrected for the competing risk of non-CVD death. We subsequently estimated the risks of CHD and stroke separately, and used the sum as a predictor for the total CVD risk. Calibration plots and c-statistics were used to evaluate the performance of the models.

Results: Performance of the Framingham predictions was good in the low- to intermediate risk (≤ 30%, 15-year CVD risk) (17.5% observed vs. 16.6% expected) but poorer in the higher risk (> 30%) categories (36.3% observed vs. 44.1% expected). The c-statistic increased from 0.66 to 0.69 after refitting. Separately estimating CHD and stroke revealed considerable heterogeneity with regard to the contribution of CHD and stroke to total CVD risk.

Conclusions: Framingham CVD risk predictions perform well in the low- to intermediate risk categories in the Rotterdam Study. Disentangling CVD into CHD and stroke separately provides additional information about the individual contribution of CHD and stroke to total individual CVD risk.

© 2014 Elsevier Ireland. Ltd All rights reserved.

1. Introduction

The use of risk scores as tools to predict cardiovascular disease (CVD) has been widely advocated in primary prevention [1–5]. Guidelines on the prevention of CVD incorporate risk scores in order to make treatment recommendations [6,7]. However, older individuals are at high risk of death due to other causes than CVD. Currently recommended Framingham risk scores tend to overestimate CVD risk in an older population, as non-CVD mortality competes with CVD events [8], and the competing risk is not taken into account in these models.

Although traditional Framingham risk scores have been successfully externally validated in some other populations, recalibration was often necessary to obtain valid estimates [9]. The 30-year CVD risk function developed by Pencina et al. [3], based on the Framingham Offspring cohort was developed to address the need for both long-term CVD prediction and taking into account the competing risk of non-CVD death. The function estimates total CVD as the combination of coronary heart disease (CHD) and stroke. In contrast with more traditional risk scores, this Framingham risk function has not been externally validated.

Both CHD and stroke contribute to the risk of total CVD, but can be regarded as different clinical events, for which different risk factors have been identified [5,10]. As the prevention of both events sometimes is associated with different recommendations [11], disentangling the risk of total CVD into both components could provide clinicians with useful additional information for treatment management.

Therefore, using 15-year follow-up data from the participants of the Rotterdam Study Cohort, a population based cohort study of elderly individuals [12], we aimed to 1) evaluate the performance of Framingham predictions of cardiovascular disease (CVD) risk corrected for the competing risk of non-CVD death, in an independent European cohort and 2) update the predictions by disentangling CVD into coronary heart disease (CHD) and stroke separately.
2. Methods

2.1. Study population

Of the 7983 respondents originally included in the Rotterdam Study, 6871 individuals both visited the research center and signed an informed consent of. Of those, 6004 individuals had no history of CHD and stroke. Individuals have been followed in an ongoing effort from 1990 onwards and consisted of regular examinations with interviews and direct digital linkage to medical files from the general practitioners working in the research area, death registries and other available medical sources, ensuring accurate follow-up of fatal and non-fatal CVD events and cause-specific mortality [12]. The medical records of nursing home were also evaluated. At baseline, participants were interviewed at home by trained research assistants using a computerized questionnaire. Basic data included information on the current health status, history of cardiovascular disease, current medication use, and cardiovascular risk factors. Subsequently, the participants were invited to the research center in order to obtain measurements on cardiovascular risk factors, including body mass index, blood pressure, total cholesterol, high-density lipoprotein (HDL) cholesterol, and non-fasting glucose level. All subjects gave written informed consent, and the study was approved by the medical ethics committee of Erasmus MC.

2.2. Assessment of risk factors

Details of the assessment of CVD risk factors and medical history in the Rotterdam Study are described into more detail elsewhere [13]. In short, participants were categorized with regard to current smoking status (nonsmoker defined as never smoked or abstinence for at least 2 years). Systolic blood pressure was calculated as the mean of two measurements [14]. Serum total and HDL cholesterol levels were determined by an automated enzymatic procedure. Diabetes mellitus was defined as current use of anti-diabetic medication and/or a random or post-load serum glucose ≥200 mg/dL (11.1 mmol/L).

2.3. Clinical end points

Events were classified using ICD-10 codes. We focused on ‘hard’ CVD as the outcome of interest, defined as the composite of hard CHD (consisting of myocardial infarction and coronary death and stroke, both fatal and non-fatal). The definition of coronary heart disease mortality, more specifically the out-of-hospital mortality attributable to CVD, changed slightly during follow-up from 2003 onwards [15], to enhance comparability with other large CVD cohort studies. In order to adjust for the competing risk of non-CVD death, as was done in the Framingham model, we defined non-CVD mortality as any death due to causes other than from CVD events. All events were independently adjudicated by two research physicians. Consensus was met in a separate session and if necessary medical specialists were consulted. We used follow-up information available until January 1, 2007 leading to a maximum follow-up duration of 17 years for an individual.

2.4. Statistical analysis

Complete risk profiles were available in 5436 of the 6004 individuals used in the analysis. We imputed missing values of systolic blood pressure, total and HDL cholesterol, diabetes status, antihypertensive medication use and current smoking status of the Rotterdam Study participants with imputation models that included all risk factors—age, sex, systolic blood pressure, use of antihypertensives, smoking, diabetes, total and HDL cholesterol, and the log cumulative hazard for hard CVD [15]. All continuous variables were log-transformed by taking the natural logarithm in correspondence with the Framingham model, and truncated at their 1st and 99th percentile. Fifteen-year risks of hard CVD and competing non-CVD death for the 6004 Rotterdam Study participants were calculated using the baseline survival at 15 years of both events as reported by Pencina et al. [3], and the linear predictors of CVD and non-CVD death calculated using the published hazard rate ratios (model 1).

A standard Cox model may provide biased estimates of absolute long-term risk because it treats those who die of a non-CVD cause as eligible for the development of a CVD event. We therefore used the model proposed by Rosthoj et al. and Putter et al. [16,17], which is the same statistical model as Pencina et al. used in the 30-year predictions of CVD. This model calculates the cumulative incidence of CVD per individual, by summation of the cause-specific hazard multiplied by the survival of the CVD event and the competing non-CVD death event at each failure time.

We compared the average predicted 15-year risk of CVD, with the average observed outcome in the Rotterdam Study participants [18]. We then recalibrated the Framingham CVD model by updating the 15-year baseline survival of CVD and non-CVD death as well, with the survival as observed in the Rotterdam Study (model 2). To check whether the overall effect of the risk factors based on the Framingham data is valid for the Rotterdam population, we recalibrated model 2 by allowing for a different effect for the slope of the linear predictors of CVD and non-CVD death (model 3). Subsequently, we refitted the Framingham CVD model for CVD and non-CVD death, and compared the coefficients of the risk factors found by fitting the model in the Rotterdam population data, with the original ones published by Pencina et al. (model 4). Finally, we refined the original model by estimating the hazards of hard CHD and stroke separately. This was done as the weights assigned to different risk factors and the shape of the lifetime hazard function may be different for CHD and stroke [2]. Accounting for this difference could potentially further improve CVD risk classification. We therefore fitted three cause-specific Cox models, one for hard CHD, one for stroke and one for the competing event defined as death from any cause other than MI, coronary disease or stroke (model 5). We subsequently calculated the cumulative incidences for hard CHD and stroke, and added the cumulative incidences of hard CHD and stroke to obtain the estimate for (total) CVD.

Table 1

<table>
<thead>
<tr>
<th>Risk factor</th>
<th>Age, years, median (IQR)</th>
<th>Men, n (%)</th>
<th>Systolic BP mm Hg, median (IQR)</th>
<th>Missing data, n (%)</th>
<th>Anti-hypertensive drugs, n (%)</th>
<th>Missing data, n (%)</th>
<th>Anti-thrombotic drugs, n (%)</th>
<th>Missing data, n (%)</th>
<th>Cholesterol lowering drugs, n (%)</th>
<th>Missing data, n (%)</th>
<th>Current smoking, n (%)</th>
<th>Missing data, n (%)</th>
<th>HDL cholesterol, mg/dL, median (IQR)</th>
<th>Missing data, n (%)</th>
<th>Diabetes mellitus, n (%)</th>
<th>Missing data, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years, median (IQR)</td>
<td>68 (52–75)</td>
<td>2251 (37.5)</td>
<td>138 (123–153)</td>
<td>52 (1.0%)</td>
<td>654 (10.9%)</td>
<td>4 (0.0%)</td>
<td>195 (3.2%)</td>
<td>4 (0.0%)</td>
<td>108 (1.8%)</td>
<td>4 (0.0%)</td>
<td>1345 (22.4%)</td>
<td>162 (2.7%)</td>
<td>255.8 (224.8–286.8)</td>
<td>77 (1.3%)</td>
<td>618 (10.3%)</td>
<td>406 (6.7%)</td>
</tr>
</tbody>
</table>

IQR: interquartile range.

Fig. 1. Calibration plot, showing predicted and observed 15-year risk of CVD for each decile of predicted 15-year CVD risk, based on the original Framingham CVD function (model 1, left) and the recalibrated score by adjusting baseline hazards of CVD and non-CVD death (model 2, right)
Discrimination for each model was assessed by the concordance index (c-statistic) adjusted for the competing risks by setting the failure time of an individual who experienced the competing event to infinity. In practice, this was done by adding 1 to the maximum follow-up time i.e. 15 years [8]. Subsequently, calibration of CVD was assessed by calibration plots, comparing predicted risks of CVD with observed incidences, per decile of predicted CVD risk, for each of the five models. We used deciles of predicted CVD risk to make the categories consistent across the plots. The observed incidences were adjusted for competing risks, using the R ‘CumInc’ function, which is included in the R ‘mstate’ library [17]. An Excel risk score calculator was constructed to provide clinicians with a tool to estimate the cumulative incidences of CHD, stroke and CVD conditional on an individual’s risk profile and is available in the online data supplement. All analyses were performed using SPSS version 19 (SPSS for Windows) and R version 2.14 (R Foundation for Statistical Computing, Vienna, Austria).

3. Results

Baseline characteristics of the 6004 Rotterdam Study participants used in this analysis are presented in Table 1. During 15 years of follow-up, 539 (first) hard CHD, 630 (first) stroke and 1719 competing non-CVD deaths occurred in these individuals.

3.1. Calibration

Calibration of the Framingham CVD model was found to be good in the low- to intermediate risk (≤ 30%, 15-year risk) categories (17.5% observed vs. 16.6% expected) but relatively poor in the higher risk (> 30%, 15-year risk) categories (36.3% observed vs. 44.1% expected) (Fig. 1).

Updating the baseline hazards and slope of the linear predictors of CVD and non-CVD death improved calibration in the higher risk categories slightly (36.2% observed, vs. 42.3% expected) but overestimation remained. After refitting the CVD risk function in the Rotterdam data, calibration improved substantially (low to intermediate categories: 16.6% observed vs. 16.6% expected; higher risk categories: 39.3% observed vs. 38.9% expected). Separately estimating CHD and stroke improved calibration even somewhat further (low to intermediate categories: 16.7% observed vs. 16.6% expected; higher risk categories: 38.8% observed vs. 38.8% expected) (Fig. 2). Calibration of the competing non-CVD death event, evaluated by plotting the observed risk of non-CVD death vs. predicted per decile of CVD risk, revealed that the risk of non-CVD is underestimated for the original Framingham CVD function for all categories of individuals, and increased with CVD risk. After refitting, calibration of non-CVD mortality improved as well (Fig. 3).

3.2. Discrimination

C-statistics for the Framingham CVD risk function applied to the Rotterdam Study population for the prediction of 15-year CVD risk was 0.66 and 0.68 after refitting the Framingham CVD risk function in the Rotterdam Study population. Estimating the hazard of CVD separately for CHD and stroke and using the sum as an estimate for total CVD, did not
further increase the c-statistic for the 15-year CVD risk rounded at two decimal points.

3.3. Beta coefficients

Refitting the Framingham CVD risk function in the Rotterdam data led to differences in beta coefficients compared to the original ones published by Pencina et al. [Table 2a]. For CVD, the log of age was found to have a stronger effect on CVD whereas sex, the log of systolic blood pressure, log of total and HDL cholesterol, current smoking status and diabetes were significantly less strong. For the competing risk of non-CVD death, the log of age was also found to have a significantly stronger effect, whereas the log of systolic blood pressure, current smoking and diabetes mellitus had a less strong effect (Table 2b). Separately estimating the hazards CHD and stroke, resulted in different beta coefficients for both events compared to estimating the hazard of CVD as a combined endpoint (Table 2c).

3.4. Fifteen-year risk of CHD, stroke and CVD

To illustrate the effect of different individual risk profiles on CVD risk and on the mixture of CHD and stroke, the cumulative incidences of CHD, stroke and CVD were plotted for a 15-year period for 4 individuals and on the mixture of CHD and stroke, the cumulative incidences of combined endpoint (Table 2c).

4. Discussion

Our analyses show that the Framingham CVD risk predictions perform reasonably well in predicting in the relatively older Rotterdam population for individuals at low to intermediate risk. For the higher risk categories, recalibration by refitting the function in the Rotterdam Study population was required to obtain valid estimates. Disentangling CVD into CHD and stroke separately revealed considerable heterogeneity with regard to the contribution of CHD and stroke to the total risk of CVD.

To our knowledge, this is the first attempt to validate this Framingham CVD risk function corrected for competing death in another population. Previous studies on the validity of Framingham risk functions in the Rotterdam Study focused on 10-year CHD and stroke separately [14,19] and found predictive performance to be reasonable in the lower risk categories for both events, but recalibration was necessary for the apparent overestimation in the higher risk categories. In the current analysis we extended the previous work by incorporating a longer period of follow-up and made adjustments for competing risks. In accordance with the earlier findings for 10-year CHD and stroke, we found that recalibration was especially important in the higher CVD risk categories.

Our study bears some limitations. First, the weights of the risk factors in the original Framingham CVD risk function were estimated over a 30-year period, whereas we validated the risk function for a 15-year period. If the hazard ratios of the risk factors included in the Framingham function would change over time, this could contribute

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Coefficients for the Framingham CVD risk function, for 15-year CVD and non-CVD competing death in the Rotterdam Study data, evaluating a refitted function for CVD as combined endpoint (A, model 4), competing non-CVD death (B, model 4), and CHD and stroke separately (C, model 5).</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>CVD</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Male sex</td>
<td>0.44</td>
</tr>
<tr>
<td>Natural logarithm of age</td>
<td>5.28</td>
</tr>
<tr>
<td>Natural logarithm of systolic blood pressure</td>
<td>1.68</td>
</tr>
<tr>
<td>Natural logarithm of serum total cholesterol</td>
<td>0.24</td>
</tr>
<tr>
<td>Natural logarithm of serum HDL cholesterol</td>
<td>−0.49</td>
</tr>
<tr>
<td>Current smoking</td>
<td>0.33</td>
</tr>
<tr>
<td>Use of antihypertensives</td>
<td>0.23</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>0.46</td>
</tr>
<tr>
<td>B</td>
<td>Non-CVD death</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Male sex</td>
<td>0.37</td>
</tr>
<tr>
<td>Natural logarithm of age</td>
<td>8.49</td>
</tr>
<tr>
<td>Natural logarithm of systolic blood pressure</td>
<td>0.28</td>
</tr>
<tr>
<td>Natural logarithm of serum total cholesterol</td>
<td>−0.92</td>
</tr>
<tr>
<td>Natural logarithm of serum HDL cholesterol</td>
<td>−0.12</td>
</tr>
<tr>
<td>Current smoking</td>
<td>0.58</td>
</tr>
<tr>
<td>Use of antihypertensives</td>
<td>0.05</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>0.19</td>
</tr>
<tr>
<td>C</td>
<td>CHD</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Male sex</td>
<td>0.64</td>
</tr>
<tr>
<td>Natural logarithm of age</td>
<td>5.89</td>
</tr>
<tr>
<td>Natural logarithm of systolic blood pressure</td>
<td>1.006</td>
</tr>
<tr>
<td>Natural logarithm of serum total cholesterol</td>
<td>0.97</td>
</tr>
<tr>
<td>Natural logarithm of serum HDL cholesterol</td>
<td>−0.91</td>
</tr>
<tr>
<td>Current smoking</td>
<td>0.27</td>
</tr>
<tr>
<td>Use of antihypertensives</td>
<td>0.39</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>0.51</td>
</tr>
</tbody>
</table>

Original: coefficients reported by Pencina et al. [3].
to part of the miscalibration we observed of the original function. For the original Framingham function, Pencina et al. did not find evidence for the hazard rate ratios to be time-dependent, which makes different hazard rate ratios for different time-horizons less likely [3]. From a clinical point of view, a 15-year risk is probably of greater interest in older individuals due to the shorter life expectancy and the potential effect of co-morbidities and competing causes of death. Second, when separately analyzing CHD and stroke, we used the same set of risk factors. A further improvement in predictive performance could be expected if we would allow for a different set of risk factors for both events and competing events respectively. Third, as with all validity studies that use external data from other cohorts, our study was limited by marginal differences in outcome classification. In order to enhance comparability with other large CVD cohort studies, the classification of coronary heart disease mortality was slightly modified from 2003 onward, according to standards that have served as a basis for the endorsed international case definition for out-of-hospital CHD mortality in epidemiologic studies [13], and adjusted the baseline hazard function in the current analysis. Fourth, some individuals were using CVD protective drugs at baseline. Ideally these individuals should be excluded from the analysis, but its impact on results would be limited given that only few individuals were using them. Furthermore, including these individuals in the analysis is in line with almost all other established risk scores [20]. Fifth, correcting for initiation of CVD protective drugs during follow-up could address the effect of treatment drop-ins. However, the same review of established CVD risk scores [20] revealed that correcting for treatment drop-ins is generally not accounted for, and follow-up examinations of the Rotterdam Study population showed that statin use was quite limited during later years [21]. Sixth, we did not evaluate the inclusion of novel risk factors, which might further contribute to improvement in risk classification.

As the Framingham population was younger on average than the Rotterdam Study participants, we expected the baseline hazard of CVD to be higher in the Rotterdam data. However, we observed that the Framingham function overestimated CVD risk, especially in the higher risk strata. Part of this overestimation could be explained by the fact that the Framingham function at the same time underestimated the risk of the competing non-CVD death which is of particular importance in older individuals. Underestimation of the competing event results in a higher predicted risk of the CVD event [8]. After adjusting the baseline hazards for both the CVD event and the competing risk of non-CVD death, the overestimation of CVD risk diminished substantially. The hazard rate ratios of the risk factors were sometimes different in magnitudes and significance of the effects from the ones reported by Pencina et al. [3]. Our observation that total cholesterol (in the presence of other factors) did not appear a significant predictor for CVD in the Rotterdam data was supported by earlier analyses from Bos et al. in the Rotterdam Study [22,23]. They found that serum cholesterol has a protective effect on stroke, whereas HDL cholesterol has no significant effect. This is similar to what we found when we analyzed the hazard of stroke separately from CHD. This could explain the non-significant effect for serum total cholesterol on total CVD in our analysis, as the
coefficient for total CVD is a weighted average of the coefficients for stroke and CHD separately. The difference in coefficients for age can be partly explained by the log-transformation [log], together with the older age of the Rotterdam Study cohort compared to Framingham. The increase from log 70 years to log 71 years, a one unit increase on the age scale, is smaller than the log increase from 40 to 41. This implies that the coefficient for age in the Rotterdam data should compensate for these smaller increments in the log-transformed risk factor.

We demonstrated that estimating the hazards for CHD and stroke separately allows for the simultaneous prediction of the risks of these events and found that the weights assigned to the risk factors included in the Framingham risk function are different for both. By separately estimating the hazards of these events, discrimination increased only very little, whereas calibration improved substantially compared to predicting CVD as a combined endpoint. The major contributor to CVD, being either CHD or stroke, differed between individual risk profiles, as illustrated by the four examples. This can have important clinical implications for the allocation of preventive interventions. For example, aspirin is currently recommended in men with a high risk of CHD, while in women the recommendation is only made for those with a high risk of stroke [11].

As we treated CHD, stroke and non-CVD as competing events, our risk function provides information on the separate events and also allows for adding the separate risks of CHD and stroke to obtain an estimate of total CVD risk. This provides clinicians with additional information beyond a risk function which estimates CVD as a single endpoint. The major contributor to CVD, being either CHD or stroke, differed between individual risk profiles, as illustrated by the four examples. This can have important clinical implications for the allocation of preventive interventions. For example, aspirin is currently recommended in men with a high risk of CHD, while in women the recommendation is only made for those with a high risk of stroke [11].

In conclusion, Framingham CVD risk predictions perform well in the low- to intermediate risk categories in the Rotterdam Study. Recalibration is necessary as the Framingham function overestimates CVD risk in the higher risk strata of the Rotterdam Study population. Disentangling CVD into CHD and stroke separately provides additional information about the individual contribution of CHD and stroke to total individual CVD risk and provides clinicians with additional information about the relative contribution of CHD and stroke.

Acknowledgments

This study was not funded by any funding organization or sponsor.

MH had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ijcard.2013.12.036.

References


