

# Predicting the 30-Year Risk of Cardiovascular Disease

## The Framingham Heart Study

Michael J. Pencina, PhD; Ralph B. D'Agostino, Sr, PhD; Martin G. Larson, ScD;  
Joseph M. Massaro, PhD; Ramachandran S. Vasan, MD

**Background**—Present cardiovascular disease (CVD) risk prediction algorithms were developed for a  $\leq 10$ -year follow up period. Clustering of risk factors at younger ages and increasing life expectancy suggest the need for longer-term risk prediction tools.

**Methods and Results**—We prospectively followed 4506 participants (2333 women) of the Framingham Offspring cohort aged 20 to 59 years and free of CVD and cancer at baseline examination in 1971–1974 for the development of “hard” CVD events (coronary death, myocardial infarction, stroke). We used a modified Cox model that allows adjustment for competing risk of noncardiovascular death to construct a prediction algorithm for 30-year risk of hard CVD. Cross-validated survival C statistic and calibration  $\chi^2$  were used to assess model performance. The 30-year hard CVD event rates adjusted for the competing risk of death were 7.6% for women and 18.3% for men. Standard risk factors (male sex, systolic blood pressure, antihypertensive treatment, total and high-density lipoprotein cholesterol, smoking, and diabetes mellitus), measured at baseline, were significantly related to the incidence of hard CVD and remained significant when updated regularly on follow-up. Body mass index was associated positively with 30-year risk of hard CVD only in models that did not update risk factors. Model performance was excellent as indicated by cross-validated discrimination  $C=0.803$  and calibration  $\chi^2=4.25$  ( $P=0.894$ ). In contrast, 30-year risk predictions based on different applications of 10-year functions proved inadequate.

**Conclusions**—Standard risk factors remain strong predictors of hard CVD over extended follow-up. Thirty-year risk prediction functions offer additional risk burden information that complements that of 10-year functions. (*Circulation*. 2009;119:3078-3084.)

**Key Words:** atherosclerosis ■ competing risk ■ lifetime risk ■ obesity ■ risk factors

Identification of risk factors contributing to the incidence of cardiovascular disease (CVD) is 1 of the major accomplishments of 20th century epidemiology. Going a step further, researchers were able to construct multivariable risk prediction algorithms intended to aid clinicians in risk assessment. The importance of these algorithms was underscored by their incorporation into the treatment recommendations of the Third Adult Treatment Panel.<sup>1</sup> Multiple risk scores have been proposed in the literature in the last 20 years.<sup>2–13</sup> These have all been developed for risk assessment over a  $\leq 10$ -year horizon. Some experts articulated the need to know the longer-term risk to better understand the public health burden and the true need for intervention.<sup>14</sup> As an answer to this need, several reports presented the lifetime or long-term risks of CVD, coronary heart disease (CHD), and stroke and their risk factors.<sup>15–23</sup> Some investigators attempted to calculate lifetime and long-term risks within the categories of specific risk factors or their clusters.<sup>15,24–26</sup> Their findings emphasized

the importance of risk factor levels in early adulthood on the long-term risks of CVD as well as the substantial impact of CVD risk factors on all-cause mortality. They also suggested that 10-year functions may underestimate the true risk burden, particularly in younger individuals and women. These results underscore the need for long-term CVD risk prediction models applicable to younger adults that account for the competing cause of non-CVD mortality. The need for long-term CVD risk prediction models was recently articulated by Blumenthal et al<sup>27</sup> and Sniderman and Furberg.<sup>28</sup> However, to the best of our knowledge, no algorithm has been proposed to quantify 30-year risk of CVD as a direct function of risk factors (allowing risk assessment for any combination of risk factors). This can be explained in part by the difficulty of finding a cohort with a sufficiently long and rigorous follow-up and also by the methodological complexities associated with incorporating the competing risk of death due to other causes into the multivariable risk estimation.

Received September 3, 2008; accepted April 6, 2009.

From the Department of Mathematics and Statistics (M.J.P., R.B.D., M.G.L., J.M.M.), School of Medicine (R.S.V.), and Department of Biostatistics (M.J.P., R.B.D., J.M.M.), Boston University, Boston Mass; and Framingham Heart Study, Framingham, Mass (M.J.P., R.B.D., M.G.L., J.M.M., R.S.V.). Guest Editor for this article was Mary Cushman, MD.

The online-only Data Supplement is available with this article at <http://circ.ahajournals.org/cgi/content/full/CIRCULATIONAHA.108.816694/DC1>.

Correspondence to Michael J. Pencina, PhD, Department of Biostatistics, Boston University, 111 Cummington St, Boston, MA 02215. E-mail mpencina@bu.edu

© 2009 American Heart Association, Inc.

*Circulation* is available at <http://circ.ahajournals.org>

DOI: 10.1161/CIRCULATIONAHA.108.816694

## Clinical Perspective on p 3084

In this report, we develop a tool for estimating 30-year risk of hard CVD events among individuals free of the condition at baseline. Our risk estimates allow for an adjustment for the competing risk of non-CVD death and utilize the standard risk factors that can be collected during a physician's office visit. The tool is based on the Framingham Offspring cohort that has contributed to the creation of several successful risk score algorithms<sup>2-5,29</sup> and offers over 35 years of rigorous surveillance for CVD occurrence. Its performance is contrasted with methods based on long-term applications of 10-year risks.

## Methods

The Framingham Heart Study started in 1948 with the enrollment of the "original" cohort of 5209 individuals. In 1971, some 5124 offspring of the original cohort and their spouses were enrolled into the Framingham Offspring Study.<sup>30</sup> Constant monitoring of CVD events and mortality has been performed and was available through the end of 2007 for this investigation. Attendees of the first offspring examination were eligible for analysis if they were  $\geq 20$  and  $< 60$  years of age ( $n=4828$ ), were free of CVD ( $n=4758$ ) and cancer at baseline ( $n=4723$ ), were not lost to follow-up ( $n=4680$ ), and had a complete risk factor profile, yielding a final sample of 4506 individuals (2333 women; mean age, 37 years). All participants gave written informed consent, and the study protocol was approved by the institutional review board of Boston Medical Center.

A detailed physical examination, anthropometry, blood pressure (BP) determination, and phlebotomy for vascular risk factors were conducted at each Heart Study examination, as described by D'Agostino et al.<sup>12</sup> Body mass index (BMI) was calculated as the weight in kilograms divided by the square of height in meters. Antihypertensive medication use was ascertained by the physician examiner at the Heart Study and based on self-report. Serum total and high-density lipoprotein (HDL) cholesterol and triglyceride levels were determined by standardized enzymatic methods. Low-density lipoprotein (LDL) cholesterol levels were calculated with the use of the Friedewald formula.<sup>31</sup> Cigarette smoking in the year preceding the examination was ascertained by self-report. Diabetes mellitus was defined as fasting glucose  $\geq 126$  mg/dL or use of insulin or oral hypoglycemic medications.

Participants were followed for a maximum of 35 years (median, 32 years). We focused on "hard" CVD as the primary outcome of interest and defined it as a composite of hard CHD (coronary death, myocardial infarction) and stroke (fatal and nonfatal). Full CVD, as defined as in D'Agostino et al.<sup>12</sup> (hard CHD plus coronary insufficiency and angina pectoris, stroke plus transient ischemic attack, intermittent claudication, and congestive heart failure), was used as a secondary outcome. Medical histories, physical examinations at the study clinic, hospitalization records, and communication with personal physicians were all used to obtain information about CVD events on follow-up.

## Statistical Analysis

In our primary model, we assessed the effect of risk factors measured at baseline on the long-term (30-year) risk of hard CVD using Cox regression.<sup>32</sup> In a secondary model, we used full CVD as outcome. Considering the extensive length of follow-up and the potential bias due to the competing risk of noncardiovascular mortality in the prediction of long-term risk, we employed the model given by Andersen et al.<sup>33-36</sup> to adjust the risk estimates for the competing risk of non-CVD mortality. The standard Cox model, similar to the standard Kaplan-Meier estimator, may provide biased estimates of absolute long-term risk because it fails to treat those who die of noncardiovascular causes as ineligible for development of CVD events.<sup>37</sup> The competing risk model corrects this shortcoming by calculating the cumulative incidence of CVD in the following manner:

$$\hat{I}_{CVD}(30) = \sum_{t_i < 30} \hat{\lambda}_{CVD}(t_i) \hat{S}(t_{i-1})$$

The quantities under summation denote the instantaneous hazard of CVD at event time  $t_i$  and survival rate from both CVD and noncardiovascular death past event time  $t_{i-1}$ . Further statistical details including estimation techniques are presented in the Technical Appendix in the online-only Data Supplement. The assumption of linearity for all predictors was verified with cumulative sums of martingale residuals as described by Lin et al.<sup>38</sup> Given no significant interactions with sex, the final model was sex-pooled but adjusted for sex. Likewise, no effect modification by baseline age was detected. We also verified that the choice of time scale (time on study versus age of onset<sup>39,40</sup>) did not affect the results. In addition to standard factors (systolic BP [SBP] and antihypertensive treatment, total and HDL cholesterol, smoking, and diabetes mellitus) used in CVD risk prediction (see D'Agostino et al.<sup>12</sup>), we considered diastolic BP (DBP), triglycerides, and LDL cholesterol as a replacement for total cholesterol (see Wilson et al.<sup>3</sup>). Moreover, given recent reports underscoring the usefulness of BMI in cardiovascular risk prediction models, we included it as a candidate risk factor (see References 12, 41, and 42). Continuous variables were log-transformed to decrease the impact of extreme observations.

To assess model performance, we used the discrimination C statistic, which takes into account the timing of events, as proposed by Harrell et al.<sup>43</sup> and Pencina and D'Agostino,<sup>44</sup> and D'Agostino's and Nam's<sup>45</sup> modification of the Hosmer-Lemeshow calibration  $\chi^2$  with survival estimates adjusted for the competing risk of noncardiovascular death.<sup>37</sup> Five-fold cross-validation<sup>46</sup> was used to account for the fact that we evaluated the model on the same data on which it was developed; in this way, we were able to utilize all data available while correcting for potential overoptimism in the assessment of model performance. Additionally, we performed internal validation by randomly splitting the sample 2:1 and developing the function on the first two thirds and evaluating its performance on the remaining third. Net reclassification improvement, as proposed by Pencina and D'Agostino et al.,<sup>47</sup> was used to assess the clinical utility of additional variables and different ways of estimating 30-year risk (see below). Third Adult Treatment Panel-based<sup>1</sup> cutoff points determining categories of low, intermediate, and high risk were adjusted proportionally to the increased duration of follow-up and incidence (from 6% and 20% to 12% and 40%).

An Excel risk score calculator was constructed to facilitate application by clinicians and is available in the online-only Data Supplement. All probability values reported were 2-sided, and a conservative 0.01 level of significance was adopted to avoid inclusion of weak effects. SAS version 9.1 was used to perform all analyses.<sup>48</sup>

## Time-Dependent Analysis

Given the long-term follow-up and the fact that all risk factors were reassessed regularly approximately every 4 years between the 1970s and early 2000s, we performed an additional analysis updating all variables as soon as the new values became available. This resulted in a Cox regression with time-dependent covariates that corresponds to a short-term risk assessment. The results were contrasted with those obtained for the 30-year model developed without updating.

## Comparison With Alternative Approaches for 30-Year Risk Prediction

The following approaches were considered:

1. Naive. This method calculated the 30-year risk as 3 times the 10-year risk from the model that did not account for the competing risks. It ignores aging as a key determinant of CVD risk, and therefore we know a priori that it cannot be correct. However, given its simplicity it might seem attractive, and we wanted to assess the amount of bias that it would introduce.
2. Combined. This approach utilized an application of 10-year CVD risk calculators. For fairness of comparison, we estimated 10-year probabilities of survival based on our data. Three probabilities

**Table 1. Baseline Characteristics and Incident Events**

Clinical Features	Women (n=2333)	Men (n=2173)
Age, y	36.3±9.3	37.3±9.2
Age group, %		
20–29 y	28.0	23.9
30–39 y	33.9	34.9
40–49 y	28.9	29.7
50–59 y	9.2	11.5
SBP, mm Hg	118±16	126±15
DBP, mm Hg	76±10	82±11
Antihypertensive treatment, %	2.7	3.1
Total cholesterol, mg/dL	192±38	202±39
HDL cholesterol, mg/dL	57±15	44±12
LDL cholesterol, mg/dL	120±35	135±35
Triglycerides, mg/dL	77±73	115±97
BMI, kg/m <sup>2</sup>	23.9±4.5	26.5±3.6
Smoking, %	45.0	46.2
Diabetes mellitus, %	0.9	2.6
Incident events, n		
Cardiovascular death	22	67
Myocardial infarction	111	281
Fatal and nonfatal stroke	86	104

Values are mean±SD for continuous variables or percentages for categorical variables.

were calculated for each person: the first using baseline age, second using baseline age plus 10 years, and third using baseline age plus 20 years, with all other risk factors based on the baseline values. The 30-year risk was calculated as the difference of 1 minus the product of these three 10-year probabilities.

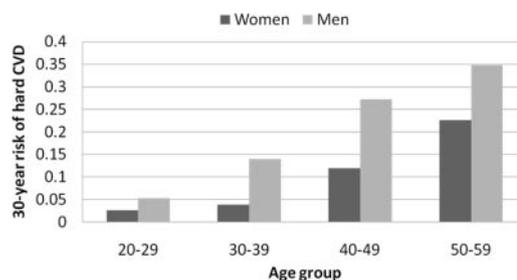
- Unadjusted. Here we applied the standard Cox model to our data with full follow-up, ignoring competing risk of death.
- Adjusted. This is our main approach following the aforementioned model.

The aforementioned 4 methods were applied to individuals with different combinations of risk factors for hard CVD events.

The authors had full access to and take full responsibility for the integrity of the data. All authors have read and agree to the manuscript as written.

## Results

The sex-specific risk factor profile of our sample at baseline as well as number and type of hard CVD events are given in Table 1. In our baseline cohort aged ≥20 but <60 years, men had numerically higher levels of all risk factors except HDL cholesterol. The 3 younger age decades (20 to 29, 30 to 39, 40 to 49 years) had similar representation between sexes, with the fewest people in the oldest age group (50 to 59 years). Over a maximum of 35 years of follow-up, 671 participants (219 women) experienced a first hard CVD event, and 622 (267 women) died of non-CVD causes. Of note, strokes constituted almost 40% of CVD events experienced by women but <25% of events experienced by men. The 30-year Kaplan–Meier rate of hard CVD adjusted for the competing risk of non-CVD death (with the use of the adjustment of Gaynor et al<sup>37</sup>) was 7.6% for women and 18.3% for men. These are lower than the long-term and lifetime risks reported



**Figure 1.** Thirty-year risk of hard CVD in women and men by age group, adjusted for the competing risk of noncardiovascular death.

previously from the Framingham data.<sup>23</sup> This can be explained by the substantially younger ages of our cohort and the different definition of the end point of interest (only hard events were considered in our analysis). Event rates by sex and age decade are presented in Figure 1.

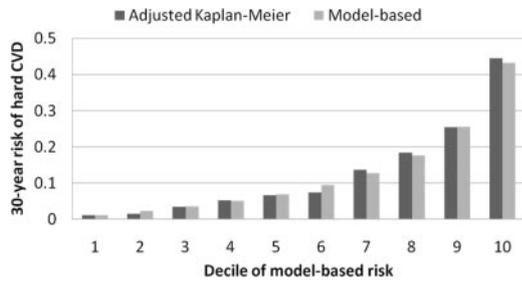
Standard CVD risk factors (male sex, age, SBP, antihypertensive treatment, total and HDL cholesterol, smoking, and diabetes mellitus) were highly significant (0.01 level) in the multivariable model. DBP and triglycerides were not statistically significant, and inclusion of LDL in place of total cholesterol did not improve model performance. BMI was weakly significant in the final model ( $P=0.04$ ); it did not increase the C statistic and had a nonsignificant net reclassification improvement of <1%. Hence, we decided not to include it in the main risk prediction model. However, following the example of D'Agostino et al,<sup>12</sup> we constructed a simplified office-based risk model in which BMI replaced the lipids. It was highly significant in the simple model along with all other risk factors ( $P\leq 0.01$ ). Hazard ratios with CIs for both models are presented in Table 2. Corresponding results for the secondary end point of full CVD are given in the Table in the online-only Data Supplement.

The 30-year risk model offered excellent discrimination (cross-validated C statistic=0.803; 95% CI, 0.786 to 0.820; internally validated C statistic=0.802; 95% CI, 0.772 to 0.832) and calibration (cross-validated Nam-D'Agostino  $\chi^2=4.25$ ;  $P=0.894$ ; Figure 2; internally validated  $\chi^2=3.98$ ;  $P=0.913$ ). It is important to note that our model, which adjusted for the competing risk of noncardiovascular death, improved the model calibration compared with the model that ignored the competing

**Table 2. Hazard Ratios With 95% CIs for 30-Year Risk of Hard CVD**

Variables	Main Model	Simple Model
Male sex	1.73 (1.45, 2.07)	2.08 (1.77, 2.46)
Age	2.09 (1.88, 2.31)	2.22 (2.01, 2.45)
SBP	1.29 (1.19, 1.39)	1.26 (1.16, 1.36)
Antihypertensive treatment	1.48 (1.10, 2.00)	1.48 (1.09, 2.00)
Smoking	2.01 (1.72, 2.35)	2.21 (1.90, 2.58)
Diabetes mellitus	2.49 (1.82, 3.41)	2.82 (2.07, 3.84)
Total cholesterol	1.33 (1.23, 1.44)	...
HDL cholesterol	0.78 (0.72, 0.84)	...
BMI	...	1.20 (1.10, 1.30)

Hazard ratios for continuous risk factors are given per 1-SD increase in the natural logarithm. All  $P\leq 0.01$ .



**Figure 2.** Calibration by decile, adjusted for the competing risk of noncardiovascular death.

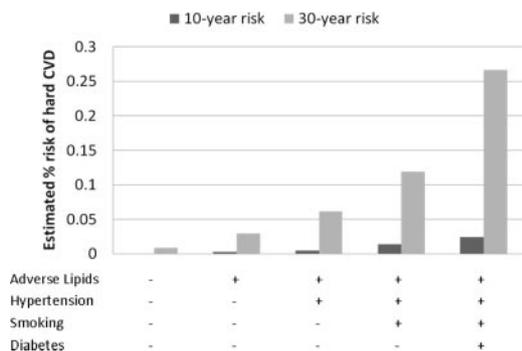
risk (cross-validated Nam-D’Agostino  $\chi^2$  for the model that ignored the competing risk=18.7;  $P=0.027$ ).

Figures 3 and 4 contrast the estimated 30-year risks of hard CVD adjusted for the competing risk of noncardiovascular death with 10-year risks, calculated for 25-year-old women and men, respectively. The corresponding estimates for 45-year-old subjects are given in Figures 5 and 6. These ages were selected to illustrate CVD risk burden in young adults with varied combinations of risk factors as well as in middle-aged adults. The results for 25-year-old subjects are striking, especially for women. The 10-year models suggest negligible risk levels (<2.5% in women and 5% in men), whereas the 30-year models give estimates that are almost 10 times higher. For example, 10-year risk for a 25-year-old smoking woman with adverse lipid profile and hypertension is only 1.4%, but the corresponding 30-year risk reaches 12%.

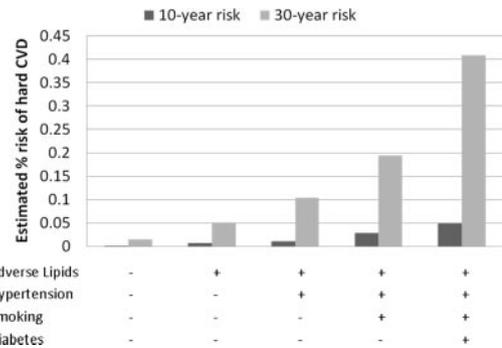
**Results of Time-Dependent Analysis**

In time-dependent analysis updating all variables approximately every 4 years, all standard risk factors remained significantly related to the hard CVD outcome with hazard ratios similar to those obtained in 30-year risk models (Table 3). Smoking had the biggest numerical change: The hazard ratio increased by approximately one third in the time-dependent model. This can be explained by the fact that time-updated models focus on a shorter event follow-up and can account for changes in risk factor levels (smoking cessation in this case).

The most interesting change occurred in the impact of BMI on the risk of hard CVD. BMI was weakly significant in the



**Figure 3.** Ten- vs 30-year risk of hard CVD for 25-year-old women with different risk profiles. No risk factors profile: total cholesterol=150 mg/dL; HDL cholesterol=60 mg/dL; untreated SBP=110 mm Hg; nonsmoker; nondiabetic. Adverse lipids: total cholesterol=260 mg/dL; HDL cholesterol=35 mg/dL. Hypertension: SBP=160 mm Hg, untreated.

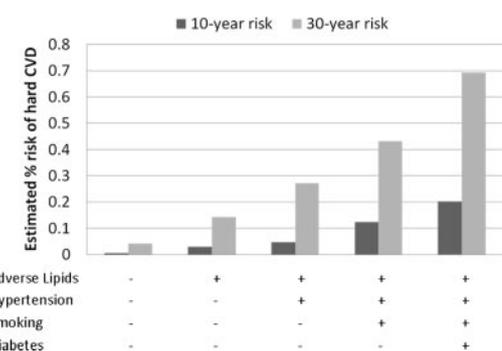


**Figure 4.** Ten- vs 30-year risk of hard CVD for 25-year-old men with different risk profiles. No risk factors profile: total cholesterol=150 mg/dL; HDL cholesterol=60 mg/dL; untreated SBP=110 mm Hg; nonsmoker; nondiabetic. Adverse lipids: total cholesterol=260 mg/dL; HDL cholesterol=35 mg/dL. Hypertension: SBP=160 mm Hg, untreated.

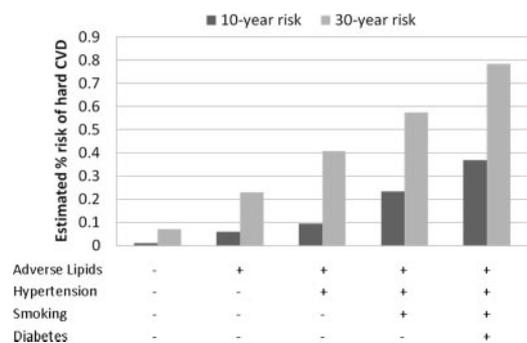
long-term, 30-year model (hazard ratio=1.10 per 1 SD;  $P=0.04$ ) but lost its entire impact in the time-dependent model (hazard ratio=0.99;  $P=0.82$ ). This finding illustrates how the effect of BMI is mediated through other risk factors: It is present in the 30-year risk model when the follow-up is extended for a long period from the baseline, but then it affects the individual risk factors, and after we control for this impact in time-updated models, BMI loses its significance.

**Comparison With Alternative Approaches for 30-Year Risk Prediction**

The mean estimated 30-year risks based on our model (the “adjusted” approach) were 7.9% for women and 18.0% for men (as expected, very close to the incidence rates given above). As expected, the “naive” approach consistently underestimated the true risk; the mean risks were 4.1% for women and 13.3% for men. If we ignored the competing risk of noncardiovascular death (“unadjusted” approach), the mean risks increased to 8.6% and 20.4%, respectively. The risks based on the “combined” approach averaged across our cohort were even higher; however, the relationship varied across individuals with different levels of risk factors. When we applied these approaches to calculate 30-year risks for individuals with different combinations of risk factors, the



**Figure 5.** Ten- vs 30-year risk of hard CVD for 45-year-old women with different risk profiles. No risk factors profile: total cholesterol=150 mg/dL; HDL cholesterol=60 mg/dL; untreated SBP=110 mm Hg; nonsmoker; nondiabetic. Adverse lipids: total cholesterol=260 mg/dL; HDL cholesterol=35 mg/dL. Hypertension: SBP=160 mm Hg, untreated.



**Figure 6.** Ten- vs 30-year risk of hard CVD for 45-year-old men with different risk profiles. No risk factors profile: total cholesterol=150 mg/dL; HDL cholesterol=60 mg/dL; untreated SBP=110 mm Hg; nonsmoker; nondiabetic. Adverse lipids: total cholesterol=260 mg/dL; HDL cholesterol=35 mg/dL. Hypertension: SBP=160 mm Hg, untreated.

unadjusted approach consistently overestimated the correct predictions based on the adjusted model. The combined approach underestimated the true risk for people with lower risk (younger and with fewer risk factors) and overestimated the risk in people with higher risk (older with several risk factors). Differences were more pronounced for higher risk levels (>20%). There was a 10% (95% CI, 6% to 14%) net reclassification improvement resulting from using the adjusted 30-year risk estimates over the tripled 10-year risks (naive approach) but no improvement when compared with the unadjusted or combined approaches.

## Discussion

In this report, we present a simple way to estimate 30-year risk of hard CVD based on risk factors routinely measured during an office visit. The results are based on >30 years of rigorous follow-up and ascertainment of CVD incidence and death. Our algorithm allows for risk assessment for individuals with any combination of continuous and categorical risk factors. It also accounts for the competing risk of noncardiovascular death.

Our approach is based on advanced statistical techniques that allow avoiding bias in the assessment of true absolute risk. Ignoring the competing risk of death inflates the estimates by an average of 1% to 2% on the absolute scale (or 10% on the relative scale), which leads to inferior calibration as demonstrated in the  $\chi^2$  statistics. On the other hand, simpler approaches that try to make 30-year inferences on the basis of a 10-year risk model are inadequate and may lead to underestimation or overestimation of the true risk burden.

The need for long-term risk prediction tools has been articulated for many years<sup>14</sup> as a complement for the shorter-term calculators. There are several reasons why it is necessary. As Sniderman and Furberg<sup>28</sup> point out, studies with shorter follow-up miss cases that would be found if the duration was extended and thus “restrict our appreciation of the true importance of the modifiable factors that cause vascular disease.” As seen here, 30-year risk cannot be adequately replaced by different combinations of 10-year risks. Blumenthal et al<sup>27</sup> raise the issue of high lifetime risk of CVD in women to underscore the need for long-term risk prediction algorithms. This point finds a striking illustration in our data, in which adverse risk factor profile leads to high

**Table 3. Hazard Ratios With 95% CIs for 30-Year Risk of Hard CVD vs Model With Time-Dependent Risk Factors Updated Every 4 Years**

Variables	30-Year Risk Model	Time-Updated Model
Male sex	1.72 (1.44, 2.05)	2.05 (1.72, 2.44)
Age	2.08 (1.88, 2.31)	2.18 (1.97, 2.42)
SBP	1.26 (1.16, 1.37)	1.28 (1.19, 1.39)
Antihypertensive treatment	1.48 (1.10, 2.00)	1.36 (1.14, 1.62)
Smoking	2.04 (1.74, 2.38)	2.74 (2.32, 3.24)
Diabetes mellitus	2.42 (1.77, 3.31)	2.30 (1.89, 2.81)
Total cholesterol	1.32 (1.22, 1.43)	1.23 (1.14, 1.33)
HDL cholesterol	0.80 (0.73, 0.87)	0.75 (0.68, 0.81)
BMI	1.10* (1.00, 1.20)	0.99† (0.91, 1.08)

Hazard ratios for continuous risk factors are given per 1-SD increase in the natural logarithm. All  $P \leq 0.01$  except as noted.

\* $P=0.04$ .

† $P=0.82$ .

30-year risk despite young age, an effect that is entirely missed by the 10-year risk model. Moreover, most contemporary cohorts are confounded by treatment effects that significantly influence short-term prediction. The use of 30-year instruments might partially overcome this problem with its long-term focus. Effective risk communication is another reason why 30-year risk might be helpful. Individuals might be more likely to adopt necessary lifestyle changes on hearing that their 30-year risk of CVD is 1 in 8 (75th percentile of the 30-year risk in men aged <40 years) than when they are told it is 1 in 50 in 10 years (75th percentile of the 10-year risk in men aged <40 years). Moreover, the potent impact of accumulation of risk factors as presented in Figures 3 through 6 may serve as an effective risk communication tool.

We have shown that established CVD risk factors that are significant in models based on shorter follow-up duration<sup>3,4,12</sup> are also significantly related to hard CVD incidence in 30 years. The impact of risk factors measured only at baseline is similar to that of risk factors updated regularly at follow-up. The same is not true for BMI, which loses its independent impact when other covariates are time updated. No significant effect modifications by sex were detected despite differences in hard CVD composition, with strokes comprising almost 40% of all first events in women and <25% in men.

As indicated earlier, this is the first report to our knowledge that presents a risk score for incidence of hard CVD in the 30-year horizon. In their recent publications, researchers from the Chicago Heart Association Detection Project in Industry calculate the remaining lifetime risk until age 85 years adjusted for the competing risk of death for participants aged 40 to 59 with 0 to 5 elevated CVD risk factors<sup>24</sup> and quantify the effect of these standard risk factors on the 30-year risk of CVD, coronary, and all-cause mortality in women aged 18 to 39 years.<sup>25</sup> Although their reports offer valuable insights into the effect of risk factors on the long-term risk of CVD, they were not designed for individual-specific risk prediction in a clinical setting.

The analysis of changes in impact on the risk of CHD and death of baseline risk factors during long-term follow-up undertaken by the Chicago Heart Association authors<sup>24–26</sup> as well as

Menotti and Lanti<sup>49</sup> revealed that single-occasion measurement risk factors remain strong predictors even in the long-term models. This finding has been confirmed in our setting.

A few limitations of our study need to be acknowledged. First, our results and the risk score were derived on the basis of a white cohort, which potentially limits the generalizability to other ethnic groups. Appropriate recalibration (see D'Agostino et al<sup>29</sup>) may correct differences in baseline survival between ethnicities, but further investigation is warranted to determine the impact of relative risks for CVD and the competing risk of non-CVD mortality differing between ethnicities. Second, in the assessment of model performance, we accounted for overoptimism introduced by evaluating the model on the same data on which it was developed using 5-fold cross-validation and internal validation. Although these techniques are well suited for this purpose, they cannot be equated with the preferred method of validation in a different cohort. Third, we considered only standard risk factors and did not investigate the effect of novel biomarkers on the risk of hard CVD because they were not available at the baseline examination in the early 1970s. Wilson et al<sup>50</sup> have shown recently that C-reactive protein measured in the Framingham cohort between the late 1970s and early 1980s might contribute to improvement in risk reclassification, a finding postulated before by Ridker et al<sup>10,13</sup> on the basis of other cohorts. It is plausible that the use of novel biomarkers could help to reduce the amount of risk underestimation with 10-year models compared with 30-year models. Further research is needed to investigate this issue. Fourth, the effect and interplay of risk factors might be different now than it was 30 years ago. However, no other way of constructing a 30-year risk score is possible given the length of the time horizon. Fifth, the nature of our design did not account for changes in risk factor levels that can take place during the course of follow-up. For example, smoking cessation between the early 1970s and the present time period might have reduced the true absolute risk, leading to an underestimation of the effect of continued smoking. However, we have shown that regardless of the absence or presence of risk factor adjustment on follow-up, they remain strong independent predictors of hard CVD. Finally, we did not attempt to predict the 30-year risk of death from noncardiovascular causes; however, this risk was implicitly factored into the calculations as the competing cause. Further research is needed to provide estimators of long-term risks of all-cause and noncardiovascular mortality that would complement the predictions available in this report and allow for patient-specific treatment strategies.

We hope that the simple way of quantifying 30-year risk of hard CVD based on a combination of standard risk factors and additional insights into the nature of their effect presented in this report will complement the currently available 10-year risk algorithms and serve as a useful tool in the clinical and public health settings and provide a useful framework for future research.

### Sources of Funding

This study was supported by the National Heart, Lung, and Blood Institute's Framingham Heart Study, National Institutes of Health.

Framingham Heart Study research is supported by National Institutes of Health/National Heart, Lung, and Blood Institute contract No. N01-HC-25195. Dr Vasan is supported in part by 2K24 HL04334 (National Heart, Lung, and Blood Institute).

### Disclosures

Dr D'Agostino has served as a consultant or on the advisory board for Sanofi and Pfizer. The other authors have nothing to report beyond the funding sources listed above.

### References

- Executive Summary of the Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). *JAMA*. 2001;285:2486–2497.
- Anderson KM, Odell PM, Wilson PWF, Kannel WB. Cardiovascular disease risk profiles. *Am Heart J*. 1991;121:293–298.
- Wilson PWF, D'Agostino RB, Levy D, Belanger A, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation*. 1998;97:1837–1847.
- D'Agostino RB, Russell MW, Huse DM, Ellison RC, Silbershatz H, Wilson PWF, Hartz SC. Primary and subsequent coronary risk appraisal: new results from the Framingham Study. *Am Heart J*. 2000;139:272–281.
- D'Agostino RB, Wolf PA, Belanger A, Kannel WB. Stroke risk profile: adjustment for antihypertensive medication. *Stroke*. 1994;25:40–43.
- Assmann G, Cullen P, Schulte H. Simple scoring scheme for calculating the risk of acute coronary events based on the 10-year follow-up of the Prospective Cardiovascular Munster (PROCAM) Study. *Circulation*. 2002;105:310–315.
- Conroy RM, Pyorala K, Fitzgerald AP, Sans S, Menotti A, DeBacker G, DeBacquer D, Ducimetiere P, Jousilahti P, Keil U, Njolstad I, Oganov RG, Thomsen T, Tunstall-Pedoe H, Tverdal A, Wedel H, Whincup P, Wilhelmens L, Graham IM. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J*. 2003;24:987–1003.
- Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M, Brindle P. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. *BMJ*. 2007;335:136.
- Menotti A, Lanti M, Gabiti-Rosei E, Carratelli L, Cavera G, Dormi A, Gaddi A, Mancini M, Motolese M, Muiiesan ML, Muntoni S, Notarbartolo A, Prati P, Remiddi S, Zanchetti A. Riskard 2005: new tools for prediction of cardiovascular disease risk derived from Italian population studies. *Nutr Metab Cardiovasc Dis*. 2005;15:426–40.
- Ridker PM, Buring JE, Rifai N, Cook RN. Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds risk score. *JAMA*. 2007;297:611–619.
- Woodward M, Brindle P, Tunstall-Pedoe, for the SIGN Group on Risk Estimation. Adding social deprivation and family history to cardiovascular risk assessment: the ASSIGN score from the Scottish Heart Health Extended Cohort (SHHEC). *Heart*. 2007;93:172–176.
- D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain MR, Massaro JM, Kannel WB. General cardiovascular risk profile for use in primary care. *Circulation*. 2008;117:743–753.
- Ridker PM, Paynter NP, Rifai N, Gaziano JM, Cook NR. The Reynolds risk score for men. *Circulation*. 2008;118:2243–2251.
- Grundy SM, Pasternak R, Greenland P, Smith S, Fuster V. Assessment of cardiovascular risk by use of multiple-risk-factor assessment equations. *Circulation*. 1999;100:1481–1492.
- Lloyd-Jones DM, Wilson PFW, Larson MG, Leip E, Beiser A, D'Agostino RB, Cleeman JI, Levy D. Lifetime risk of coronary heart disease by cholesterol levels at selected ages. *Arch Intern Med*. 2003;163:1966–1972.
- Lloyd-Jones DM, Larson MG, Leip EP, Beiser A, D'Agostino RB, Kannel WB, Murabito JM, Vasan RS, Benjamin EJ, Levy D. Lifetime risk for developing congestive heart failure. *Circulation*. 2002;106:3068–3072.
- Seshadri S, Beiser A, Kelly-Hayes M, Kase CS, Au R, Kennel WB, Wolf PA. The lifetime risk of stroke: estimates from the Framingham Study. *Stroke*. 2006;37:345–350.
- Vasan RS, Beiser A, Seshadri S, Larson MG, Kannel WB, D'Agostino RB, Levy D. Residual lifetime risk of developing hypertension in middle-aged women and men. *JAMA*. 2002;287:1003–1010.

19. Vasan RS, Pencina MJ, Cobain MR, Freiberg MS, D'Agostino RB. Estimated risks for developing obesity in the Framingham Heart Study. *Ann Intern Med.* 2005;143:473–480.
20. Cobain MR, Pencina MJ, D'Agostino RB, Vasan RS. Lifetime risk for developing dyslipidemia: the Framingham Offspring Study. *Am J Med.* 2007;120:623–630.
21. Narayan KMV, Boyle JP, Thompson TJ, Sorensen SW, Williamson DF. Lifetime risk for diabetes mellitus in the United States. *JAMA.* 2003;290:1884–1890.
22. Pekkanen J, Nissinen A, Puska P, Punsar S, Karvonen M. Risk factors and 25 year risk of coronary heart disease in a male population with a high incidence of the disease: the Finnish cohorts of the Seven Countries Study. *BMJ.* 1989;299:81–85.
23. Lloyd-Jones DM, Leip EP, Larson MG, D'Agostino RB, Beiser A, Wilson PFW, Wolf PA, Levy D. Prediction of lifetime risk of cardiovascular disease by risk factor burden at 50 years of age. *Circulation.* 2006;113:791–798.
24. Lloyd-Jones DM, Dyer AR, Wang K, Daviglius ML, Greenland P. Risk factor burden in middle age and lifetime risks for cardiovascular and non-cardiovascular death (Chicago Heart Association Detection Project in Industry). *Am J Cardiol.* 2007;99:535–540.
25. Daviglius ML, Stamper J, Pirzada A, Yan LL, Garside DB, Liu K, Wang K, Dyer AR, Lloyd-Jones DM, Greenland P. Favorable cardiovascular risk profile in young women and long-term risk of cardiovascular and all-cause mortality. *JAMA.* 2004;292:1588–1592.
26. Greenland P, Knoll MD, Stamler J, Neaton JD, Dyer AR, Garside DB, Wilson PFW. Major risk factors as antecedents of fatal and nonfatal coronary heart disease events. *JAMA.* 2003;290:891–897.
27. Blumenthal RS, Michos ED, Nasir K. Further improvements in CHD risk prediction for women. *JAMA.* 2007;297:641–643.
28. Sniderman AD, Furberg CD. Age as a modifiable risk factor for cardiovascular disease. *Lancet.* 2008;371:1547–1549.
29. D'Agostino RB, Grundy S, Sullivan LM, Wilson PFW. Validation of the Framingham coronary heart disease prediction scores. *JAMA.* 2001;286:180–187.
30. Kannel WB, Feinleib M, McNamara PM, Garrison RJ, Castelli WP. An investigation of coronary heart disease in families: the Framingham Offspring Study. *Am J Epidemiol.* 1979;110:281–290.
31. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without the use of the preparative ultracentrifuge. *Clin Chem.* 1972;18:499–502.
32. Cox DR. Regression models and life tables. *J Roy Stat Soc B.* 1972;34:187–220.
33. Andersen PK, Borgan O, Gill RD, Keiding N. *Statistical Models Based on Counting Processes.* New York, NY: Springer; 1993.
34. Rosthøj S, Andersen PK, Abildstrom SZ. SAS macros for estimation of the cumulative incidence functions based on a Cox regression model for competing risks survival data. *Comput Meth Programs Biomed.* 2004;74:69–75.
35. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med.* 2007;26:2389–2430.
36. Kalbfleisch JD, Prentice RL. *The Statistical Analysis of Failure Time Data.* 2nd ed. Hoboken, NJ: Wiley; 2002.
37. Gaynor JJ, Feuer EJ, Tan CC, Wu DH, Little CR, Straus DJ, Clarkson BD, Brennan MF. On the use of cause-specific failure and conditional failure probabilities: examples from clinical oncology data. *J Am Stat Assoc.* 1993;88:400–409.
38. Lin D, Wei L, Ying Z. Checking the Cox model with cumulative sums of martingale-based residuals. *Biometrika.* 1993;80:557–572.
39. Korn EL, Graubard BI, Midthune D. Time-to-event analysis of longitudinal follow-up of a survey: choice of time scale. *Am J Epidemiol.* 1997;145:72–80.
40. Pencina MJ, Larson MG, D'Agostino RB. Choice of timescale and its effect on significance of predictors in longitudinal studies. *Stat Med.* 2007;26:1346–1359.
41. Gaziano TA, Young CR, Fitzmaurice G, Atwood S, Gaziano JM. Laboratory-based versus non-laboratory-based method for assessment of cardiovascular disease risk: the NHANES I Follow-up Study cohort. *Lancet.* 2008;371:923–931.
42. Wilson PFW, Bozeman SR, Burton TM, Hoaglin DC, Ben-Joseph R, Pashos CL. Prediction of first events of coronary heart disease and stroke with consideration of adiposity. *Circulation.* 2008;118:124–130.
43. Harrell FE, Lee KL, Mark DB. Tutorial in biostatistics: multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med.* 1996;15:361–387.
44. Pencina MJ, D'Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Stat Med.* 2004;23:2109–2123.
45. D'Agostino RB, Nam BH. Evaluation of the performance of survival analysis models: discrimination and calibration measures. *Handbook of Statistics.* New York, NY: Elsevier; 2004:1–25.
46. Hastie T, Tibshirani RFJ. *The Elements of Statistical Learning.* New York, NY: Springer; 2001.
47. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med.* 2008;27:157–172.
48. *SAS/STAT Software, Version 9.1. of the SAS System for Windows.* Cary, NC: SAS Institute Inc; 2002–2003.
49. Menotti A, Lanti M. Coronary risk factors predicting early and late coronary deaths. *Heart.* 2003;89:19–24.
50. Wilson PFW, Pencina MJ, Jacques P, Selhub J, D'Agostino RB, O'Donnell CJ. C-reactive protein and reclassification of cardiovascular risk in the Framingham Heart Study. *Circ Cardiovasc Qual Outcomes.* 2008;1:92–97.

### CLINICAL PERSPECTIVE

The impact of standard risk factors (male sex, age, systolic blood pressure, antihypertensive treatment, cholesterol levels, smoking, and diabetes mellitus) on 10-year risk of coronary or cardiovascular disease (CVD) has been studied extensively, and reliable algorithms exist for risk prediction. In the present investigation, we evaluated the impact of standard risk factors on CVD incidence on long-term follow-up (ie, over 30 years). Our observations suggest that standard risk factors remain highly predictive of CVD risk over a 30-year follow-up period and that their impact is substantial even if levels are not updated. We also quantified the impact of body mass index on 30-year risk of CVD and observed that its association with increased CVD risk is mediated partly through promoting higher levels of standard risk factors over a long-term follow-up. Additionally, we have formulated a 30-year CVD risk prediction algorithm that adjusts for the competing risk of death on long-term follow-up. Our observations suggest that different applications of 10-year risk prediction functions for estimating 30-year CVD risk may be suboptimal, especially when applied to younger women and men who have an adverse risk factor profile. We also have constructed a simple calculator for estimating 30-year risk of CVD that is based on standard risk factors (with and without knowledge of laboratory values) and that could be implemented easily in primary care settings.

**Predicting the 30-Year Risk of Cardiovascular Disease. The Framingham Heart Study**  
Michael J. Pencina, Ralph B. D'Agostino, Sr, Martin G. Larson, Joseph M. Massaro and  
Ramachandran S. Vasan

*Circulation*. published online June 8, 2009;  
*Circulation* is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231  
Copyright © 2009 American Heart Association, Inc. All rights reserved.  
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the  
World Wide Web at:

<http://circ.ahajournals.org/content/early/2009/06/08/CIRCULATIONAHA.108.816694.citation>

Data Supplement (unedited) at:

<http://circ.ahajournals.org/content/suppl/2009/06/08/CIRCULATIONAHA.108.816694.DC1>

**Permissions:** Requests for permissions to reproduce figures, tables, or portions of articles originally published in *Circulation* can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the [Permissions and Rights Question and Answer](#) document.

**Reprints:** Information about reprints can be found online at:  
<http://www.lww.com/reprints>

**Subscriptions:** Information about subscribing to *Circulation* is online at:  
<http://circ.ahajournals.org/subscriptions/>

## SUPPLEMENTAL MATERIAL

### Technical Appendix

Probabilities of development of a CVD event in the 30-year risk horizon adjusted for the competing risk of non-CVD death based on significant risk factors were calculated following the developments presented in<sup>33</sup> and adapted to our situation. We outline them below.

For CVD event define  $\lambda_{CVD}(t)$  to be the instantaneous hazard function and  $\Lambda_{CVD}(t)$  the cumulative hazard function. Similarly, define the equivalent quantities,  $\lambda_{DTH}(t)$  and  $\Lambda_{DTH}(t)$ , for the death outcome. Now define the probability of not failing from either cause as  $S(t) = \exp[-(\Lambda_{CVD}(t) + \Lambda_{DTH}(t))]$  and the cumulative incidence function for CVD event (the probability of

experiencing CVD event before time  $t$ ) as  $I_{CVD}(t) = \int_0^t \lambda_{CVD}(u)S(u)du$ . This quantity is estimated

at 30 years of follow-up as:

$$\hat{I}_{CVD}(30) = \sum_{t_i < 30} \hat{\lambda}_{CVD}(t_i) \hat{S}(t_{i-1}) \quad (A1)$$

where  $t_i, i = 1, \dots, n_{EV}$  are the ordered event times and  $t_0 = 0$ . The quantities under the summation are estimated as:

$$\hat{S}(t_{i-1}) = \exp[-(\hat{\Lambda}_{CVD}(t_{i-1}) + \hat{\Lambda}_{DTH}(t_{i-1}))] = \hat{S}_{CVD}(t_{i-1}) \cdot \hat{S}_{DTH}(t_{i-1}) \quad (A2)$$

where the last two survival functions are taken directly from the standard Cox models for CVD events and non-CVD deaths, treating the other outcome as censoring event;

$$\hat{\lambda}_{CVD}(t_i) = \hat{\Lambda}_{CVD}(t_i) - \hat{\Lambda}_{CVD}(t_{i-1}) = -\log(\hat{S}_{CVD}(t_i)) + \log(\hat{S}_{CVD}(t_{i-1})). \quad (A3)$$

In our application we used the risk factors significant in the CVD models in the models for non-CVD death. The recursive nature of formula (A1) makes it impossible to give a closed form solution for the estimated risk.

## Appendix Table

### Hazard ratios\* with 95% confidence intervals for 30-year risk of full CVD

Variables	Main Model	Simple Model
<b>Male Sex</b>	1.41 (1.23, 1.62)	1.72 (1.51, 1.96)
<b>Age</b>	1.98 (1.83, 2.14)	2.05 (1.90, 2.21)
<b>Systolic Blood Pressure</b>	1.27 (1.19, 1.35)	1.24 (1.16, 1.32)
<b>Antihypertensive Treatment</b>	1.69 (1.31, 2.17)	1.70 (1.32, 2.19)
<b>Smoking</b>	1.81 (1.60, 1.05)	1.99 (1.76, 2.25)
<b>Diabetes</b>	1.98 (1.49, 2.65)	2.22 (1.66, 2.95)
<b>Total cholesterol</b>	1.24 (1.16, 1.33)	-
<b>HDL cholesterol</b>	0.77 (0.72, 0.82)	-
<b>Body Mass Index</b>	-	1.18 (1.10, 1.26)

\*Hazard ratios for continuous risk factors are given per 1 standard deviation increase in the natural logarithm.

All p-values were less or equal to 0.01.

Title for the interactive risk calculator:

### **30-year risk of cardiovascular disease**

Legend:

Hard CVD includes coronary death, myocardial infarction and fatal and non-fatal stroke.

Full CVD includes hard CVD plus coronary insufficiency, angina pectoris, transient ischemic attack, intermittent claudication and congestive heart failure.