An Updated Coronary Risk Profile
A Statement for Health Professionals

Keaven M. Anderson, PhD; Peter W.F. Wilson, MD; Patricia M. Odell, PhD; and William B. Kannel, MD, MPH

Coronary heart disease (CHD) continues to be the cause of the greatest number of deaths among adult Americans. Although there is a downward trend in cardiovascular mortality rates, morbidity and mortality rates remain high and are of great concern to clinicians and health officials. Using a simple worksheet, a patient’s 5- and 10-year CHD risks can be estimated. The components of the profile were selected because they are objective and strongly and independently related to CHD and because they can be measured through simple office procedures and laboratory results.

In the past, investigators with the Framingham Heart Study developed CHD risk equations for use by clinicians in predicting the development of coronary disease in individuals free of disease. Early efforts reflected the experience of study investigators from 1950 to the mid-1960s. In addition to age and gender, risk factors included systolic blood pressure (SBP), serum cholesterol, cigarette smoking, glucose intolerance, and left ventricular hypertrophy (LVH). A handbook containing CHD risk tables based on Framingham equations was published in 1973. An even simpler approximation of the equations, by which CHD risk could be estimated by following instructions on a pocket-sized card, soon followed.

The equations presented here have several advantages over previous versions. The data base from which they are derived is larger and more recent. In particular, more data for individuals older than 60 years are available. In addition, the influence of high density lipoprotein (HDL) cholesterol, which has been measured in the Framingham Heart Study since 1968, is reflected in these equations. Measurement of the ratio of total cholesterol to HDL cholesterol has been found to be superior to measurement of serum cholesterol as a predictor of CHD.

At the baseline examination for this study (1968-1975), all members of the original Framingham cohort were more than 50 years old. To provide current estimates of CHD risk over a large age range, data from the original cohort have been combined with data from the second-generation study population, the Framingham Offspring Cohort, for which 12 years of follow-up have recently been completed. Together, these groups span the ages of 12–82 years; however, only persons 30–74 years old have been included in this study. Those 75 years of age or older were excluded because of possible differences in risk factors in this older group and its potentially large influence in the algorithm determination. Together, these groups span the ages of 12–82 years; however, only persons 30–74 years old have been included in this study. Those 75 years of age or older were excluded because of possible differences in risk factors in this older group and its potentially large influence in the algorithm determination. Tables 1–4 detail the crude incidence rates of CHD and the distributions of risk factors among the study population by age. From the experience of this group during a 12-year period, estimates of CHD risk have been produced that reflect the approximate combined impacts of total and HDL cholesterol, SBP or diastolic blood pressure (DBP), cigarette smoking, diabetes mellitus, and LVH as measured by electrocardiography (ECG-LVH).

The derivation and uses of the worksheet are detailed in the remainder of this article.

Methods
Population and Risk Factors

Every member of the original and offspring cohorts who met three basic criteria were included in the study. Requirements for inclusion were 1) age 30–74 years at the time of the baseline examination; 2) measurements available for SBP and DBP, cigarette smoking status, total and HDL cholesterol, and diagnoses (yes or no) of diabetes and ECG-LVH (when information on diabetes or LVH was not available, diagnoses were presumed to be negative); and 3) freedom from cardiovascular disease (stroke, transient ischemia, CHD [includes angina pectoris, coronary insufficiency (unstable angina), myocardial infarction, and sudden death], congestive heart failure, and intermittent claudication) until time of risk factor measurement.

Definitions of risk factors and end points are those considered standard in the Framingham study. In the original cohort, diabetes was diagnosed if a casual whole blood glucose measurement was 150 mg/dl or above or the individual was being treated with insulin or oral hypoglycemics. In the offspring study, a more
TABLE 1. Prevalence of Dichotomous Risk Factors and Crude Coronary Heart Disease Incidence by Age for Men

<table>
<thead>
<tr>
<th>Age (yr)</th>
<th>30–39</th>
<th>40–49</th>
<th>50–59</th>
<th>60–69</th>
<th>70–74</th>
<th>70+</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHD incidence (n) (%)</td>
<td>41 (5)</td>
<td>74 (11)</td>
<td>143 (20)</td>
<td>98 (29)</td>
<td>29 (26)</td>
<td>385 (15)</td>
</tr>
<tr>
<td>Cigarette smoking (n) (%)</td>
<td>340 (45)</td>
<td>284 (43)</td>
<td>301 (42)</td>
<td>102 (30)</td>
<td>28 (25)</td>
<td>1,055 (41)</td>
</tr>
<tr>
<td>Diabetes (n) (%)</td>
<td>11 (1)</td>
<td>23 (4)</td>
<td>69 (10)</td>
<td>63 (19)</td>
<td>17 (15)</td>
<td>183 (7)</td>
</tr>
<tr>
<td>ECG-LVH (n) (%)</td>
<td>1 (0.1)</td>
<td>4 (0.6)</td>
<td>2 (0.3)</td>
<td>17 (5)</td>
<td>4 (4)</td>
<td>28 (1)</td>
</tr>
<tr>
<td>Total (n)</td>
<td>759</td>
<td>655</td>
<td>725</td>
<td>340</td>
<td>111</td>
<td>2,590</td>
</tr>
</tbody>
</table>

CHD, coronary heart disease; ECG-LVH, left ventricular hypertrophy by electrocardiography.

Let $T$ denote the time until CHD from the beginning of follow-up, $\mu$ a location parameter, and $\sigma$ a scale parameter for $\log(T)$. Throughout, $\log()$ denotes the natural logarithm function. Assume that $\mu$ and $\sigma$ depend on risk factors, as will be described, and that

$$\log(T) = \frac{\mu}{\sigma}$$

follows an extreme value distribution. This implies that $T$ follows the Weibull distribution, which is often used in analyzing studies of times until event.

The location parameter $\mu$ is assumed to be a sum of the products of the risk factors multiplied by their corresponding coefficients; for example,

$$\mu = \beta_0 + \beta_1 \times \text{female} + \beta_2 \times \text{log (age)} + \beta_3 \times \text{log (SBP)} + \ldots$$

The exact form taken by this function is presented in “Results.” Next, assume that

$$\log(\sigma) = \theta_0 + \theta_1 \times \mu$$

This functional form for $\sigma$, as well as others, is discussed in detail elsewhere and provides a highly statistically significant improvement in model fit over the standard model in which $\sigma$ is held constant. Thus, the unknown parameters to be estimated are $\beta_0, \ldots, \beta_5, \theta_0$, and $\theta_1$.

The models are estimated by the maximum likelihood method, which is implemented with a specialized computer program written by one of the authors. This type of model can also be estimated with PROC NLMIX from SAS Institute, Cary, N.C. To reduce the apparently undue influence of the earliest CHD events, those events occurring during the first 4 years of follow-up are coded as such rather than as occur-

TABLE 2. Prevalence of Dichotomous Risk Factors and Crude Coronary Heart Disease Incidence by Age for Women

<table>
<thead>
<tr>
<th>Age (yr)</th>
<th>30–39</th>
<th>40–49</th>
<th>50–59</th>
<th>60–69</th>
<th>70–74</th>
<th>70+</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHD incidence (n) (%)</td>
<td>4 (1)</td>
<td>32 (5)</td>
<td>105 (12)</td>
<td>68 (15)</td>
<td>32 (20)</td>
<td>241 (8)</td>
</tr>
<tr>
<td>Cigarette smoking (n) (%)</td>
<td>357 (45)</td>
<td>284 (41)</td>
<td>391 (45)</td>
<td>101 (22)</td>
<td>26 (16)</td>
<td>1,159 (39)</td>
</tr>
<tr>
<td>Diabetes (n) (%)</td>
<td>4 (0.5)</td>
<td>13 (2)</td>
<td>58 (7)</td>
<td>55 (12)</td>
<td>24 (15)</td>
<td>154 (5)</td>
</tr>
<tr>
<td>ECG-LVH (n) (%)</td>
<td>2 (0.3)</td>
<td>0 (0)</td>
<td>3 (0.4)</td>
<td>3 (0.7)</td>
<td>7 (4)</td>
<td>15 (0.5)</td>
</tr>
<tr>
<td>Total (n)</td>
<td>798</td>
<td>693</td>
<td>867</td>
<td>462</td>
<td>163</td>
<td>2,983</td>
</tr>
</tbody>
</table>

CHD, coronary heart disease; ECG-LVH, left ventricular hypertrophy by electrocardiography.
Results

Results are presented in two forms; the first is an updated version of the 1973 report equations that incorporates new data on HDL cholesterol with the present study data described above. Then, an easy-to-calculate point-scoring algorithm that is based on one of the equations is given. It is hoped that the algorithm will be easy to use. Predicted risk may be calculated with SBP or DBP. Because of the high correlation between these two measurements, both cannot be included; the redundancy leads to difficulty in interpretation. Although in general results are similar, separate equations are given for SBP and DBP to accommodate strong user preferences. The equation incorporating SBP is recommended because SBP is more accurately determined, has a wider range of values, and is a stronger predictor of CHD, particularly in the elderly.

**Framingham Equations**

Computation with the estimated equation is described below. In this section, SBP is used; the equation incorporating DBP is given next. Explanations of some components of the equation are given after the calculations.

**Systolic Blood Pressure Equation**

There are some differences in equation calculations for men and women, but both calculations begin in the same way. Compute an interim number \( a \) that is based on risk factor measurements.

\[
a = 11.122 - 0.9119 \times \log (SBP) - 0.2767 \times \text{smoking} - 0.7181 \times \log \left(\frac{\text{cholesterol}}{\text{HDL}}\right) - 0.5865 \times \text{ECG-LVH}
\]

(1)

The next step, computing a second interim value \( m \), is different for men and women. For men, compute

\[
m = a - 1.4792 \times \log (age) - 0.1759 \times \text{diabetes}
\]

(2a)

For women, compute

\[
m = a - 5.8549 + 1.8515 \times [\log (age/74)]^2 - 0.3758 \times \text{diabetes}
\]

(2b)

Next, for both sexes, compute

\[
\mu = 4.4181 + m
\]

(3)

\[
\sigma = \exp (-0.3155 - 0.2784 \times m)
\]

(4)
Finally, choose the number of years for which you wish to predict (from 4 to 12) and call it $t$. Compute

$$ u = \frac{\log (t) - \mu}{\sigma} $$

(5)

The predicted probability for $t$ is

$$ p = 1 - \exp (-e^u) $$

(6)

As an example, consider a 55-year-old individual with SBP of 130 mm Hg, total cholesterol of 240 mg/dl, and HDL cholesterol of 45 mg/dl who smokes cigarettes. We assume that neither diabetes nor ECG-LVH has been diagnosed. First, we compute

$$ a = 11.122 - 0.9119 \times \log (130) - 0.2767 - 0.7181 \times \log (240/45) = 5.1947 $$

For a man, then compute

$$ m = 5.1947 - 1.4792 \times \log (55) = -0.7329 $$

$$ \mu = -0.7329 + 4.4181 = 3.685 $$

$$ \sigma = \exp [-0.3155 - 0.2784 \times (-0.7329)] = 0.894 $$

For a woman, compute

$$ m = 5.1947 - 5.8549 + 1.8515 \times [\log (55/74)]^2 = -0.4972 $$

$$ \mu = -0.4972 + 4.4181 = 3.921 $$

$$ \sigma = \exp [-0.3155 - 0.2784 \times (-0.4972)] = 0.8377 $$

We let $t$ be 10 years and compute for a man

$$ p = 1 - \exp (-e^{-1.546}) = 0.192 $$

For a woman, compute

$$ p = 1 - \exp (-e^{-1.932}) = 0.135 $$

### Diastolic Blood Pressure Equation

The equations incorporating DBP instead of SBP are precisely analogous to those given above.

$$ a = 11.0938 - 0.8670 \times \log (DBP) - 0.2789 \times \text{smoking} - 0.7142 \times \log (\text{cholesterol/HDL}) - 0.7195 \times \text{ECG-LVH} $$

(7)
Again, the next step is different for men and women. For men, compute
\[ m = a - 1.6346 \times \log (\text{age}) - 0.2082 \times \text{diabetes} \quad (8a) \]
For women, compute
\[ m = a - 6.5306 + 2.1059 \times [\log (\text{age}/74)]^2 - 0.4055 \times \text{diabetes} \quad (8b) \]
Next, compute
\[ \mu = 4.4284 + m \quad (9) \]
\[ \sigma = \exp (-0.3171 - 0.2825 \times m) \quad (10) \]
Proceed as with the SBP equation to compute predicted probabilities. For example, if DBP is 90 mm Hg, \( p = 0.22 \) for a man and \( p = 0.16 \) for a woman.

**Model Selection**

There are several points to consider in the derivations of the models. First, the natural logarithms of continuous covariates were used rather than actual values. Likelihood analysis indicates that such use improves the fit of the model; one reason is that extremely large covariate measurements now receive less emphasis, allowing risk to change more slowly for very large values than for the range of the bulk of the data. This explains the uneven covariate intervals in the risk prediction worksheet.

Second, the addition of a quadratic term for age for women produces a significantly improved model fit. This term accommodates a rapid increase in risk at younger ages and little change at older ages. The maximum age effect in women was arbitrarily forced to occur at the oldest age in the population; otherwise, this would have occurred at a slightly younger age. This change is not statistically significant. Because of the quadratic form of the function, extrapolation to older and younger age groups is particularly questionable.

Third, the ratio of total cholesterol to HDL cholesterol was used because no improved fit was found when the covariates were used separately. Note that the equations do not contain a cholesterol–age interaction term as did some previous prediction models. When serum cholesterol–age and HDL cholesterol–age interaction terms were added to the model, their contributions were negligible. Thus, it appears that including the ratio of total cholesterol to HDL cholesterol in the equations eliminates the need for such interaction terms.

Fourth, except for differences in aging effects, no significantly different effects were found for men and women. A separate coefficient for diabetes was fit for women based on significant differences between men and women in previous studies. Although female smokers were not found to be at increased risk in some previous Framingham reports, the offspring cohort provides evidence of such a relation. Thus, no distinction has been made between smoking effects for men and those for women. Using the number of cigarettes rather than a simple yes-or-no code for cigarette smoking did not provide an improvement in model fit. This may be largely due to the fact that older smokers in this study reported smoking fewer cigarettes than did younger smokers. Presumably, this confounded possible dose effects with age effects.

Finally, the estimated effect of ECG-LVH is very large but has a large standard error because of the small prevalence of the condition at baseline.

**Point Score Algorithm**

Equations 1–6 were used to devise a worksheet that enables users to estimate their CHD risk by assigning a point score to each risk factor (Table 5). Clinicians or patients can insert the appropriate points and add as indicated to obtain a good approximation of CHD risk during a 5- or 10-year period. For comparison, average risk values, by age and sex, for the Framingham population are also given.

If the values used in the sample computations are used in the worksheet, results are as given in Table 6. These results are a close approximation of those obtained with Equations 1–6 (19.2% for a man and 13.5% for a woman).

**Discussion**

The CHD risk equations and point score prediction probability algorithm attempt to elucidate the multifactorial etiology of CHD and produce possible comparisons and interpretations for clinicians and their patients at risk. The point-scoring technique allows estimation of an individual’s 5- and 10-year risks of CHD, whereas the equations may be used for 4- to 12-year estimations. These calculations are based on the Framingham experience for individuals free of cardiovascular disease at baseline, and such calculations are not appropriate for individuals with coronary disease.

Generalization of the Framingham equations to the population at large is always a matter of concern and should be done cautiously. However, it has been demonstrated repeatedly that the Framingham risk model is effective in predicting heart disease in other large population samples in the United States. For example, the 1973 equations were used in the Multiple Risk Factor Intervention Trial to predict the number of cases of CHD to be expected during the course of the trial. The trial was administered during the early 1970s, when there was a sharp decrease in CHD mortality rates in the United States and extensive intervention against risk factors was being practiced. Hence, as could be anticipated, the estimates were high for both the usual-care and the special-intervention groups. But they were systematically high, and the relative weightings of risk factors distinguished low- from high-risk individuals.

Nonetheless, cautions should be observed when using the equations or worksheet. First, the equations can, of course, be used only with all risk factors...
### Table 5. Framingham Heart Study Coronary Heart Disease Risk Prediction Chart

1. **Find points for each risk factor**

<table>
<thead>
<tr>
<th>Age (if female) (yr)</th>
<th>Age (if male) (yr)</th>
<th>HDL cholesterol</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Points</td>
<td>Points</td>
</tr>
<tr>
<td>Age</td>
<td>Points</td>
<td>Age</td>
</tr>
<tr>
<td></td>
<td>Points</td>
<td>HDL</td>
</tr>
<tr>
<td>30</td>
<td>12</td>
<td>41</td>
</tr>
<tr>
<td>31</td>
<td>-11</td>
<td>42-43</td>
</tr>
<tr>
<td>32</td>
<td>-9</td>
<td>44</td>
</tr>
<tr>
<td>33</td>
<td>-8</td>
<td>45-46</td>
</tr>
<tr>
<td>34</td>
<td>-6</td>
<td>47-48</td>
</tr>
<tr>
<td>35</td>
<td>-5</td>
<td>49-50</td>
</tr>
<tr>
<td>36</td>
<td>-4</td>
<td>51-52</td>
</tr>
<tr>
<td>37</td>
<td>-3</td>
<td>53-55</td>
</tr>
<tr>
<td>38</td>
<td>-2</td>
<td>56-60</td>
</tr>
<tr>
<td>39</td>
<td>-1</td>
<td>61-67</td>
</tr>
<tr>
<td>40</td>
<td>0</td>
<td>68-74</td>
</tr>
</tbody>
</table>

**Total cholesterol (mg/dl)**

<table>
<thead>
<tr>
<th>Chol Points</th>
<th>Chol Points</th>
<th>SBP Points</th>
<th>SBP Points</th>
<th>Other factors</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>139-151</td>
<td>-3</td>
<td>220-239</td>
<td>2</td>
<td>98-104</td>
<td>-2</td>
<td>150-160</td>
</tr>
<tr>
<td>152-166</td>
<td>-2</td>
<td>240-262</td>
<td>3</td>
<td>105-112</td>
<td>-1</td>
<td>161-172</td>
</tr>
<tr>
<td>167-182</td>
<td>-1</td>
<td>263-288</td>
<td>4</td>
<td>113-120</td>
<td>0</td>
<td>173-185</td>
</tr>
<tr>
<td>183-199</td>
<td>0</td>
<td>289-315</td>
<td>5</td>
<td>121-129</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>200-219</td>
<td>1</td>
<td>316-330</td>
<td>6</td>
<td>130-139</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

2. **Add points for all risk factors**

\[(\text{Age}) + (\text{Total chol}) + (\text{HDL}) + (\text{SBP}) + (\text{Smoking}) + (\text{Diabetes}) + (\text{ECG-LVH}) = (\text{Total})\]

Note: Minus points subtract from total.

3. **Look up risk corresponding to point total**

<table>
<thead>
<tr>
<th>Points</th>
<th>Probability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability (%)</td>
<td>5 yr</td>
</tr>
<tr>
<td>≤1</td>
<td>&lt; 1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
</tr>
</tbody>
</table>

4. **Compare with average 10-year risk**

<table>
<thead>
<tr>
<th>Age (yr)</th>
<th>Probability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability (%)</td>
<td>Women</td>
</tr>
<tr>
<td>30-34</td>
<td>&lt; 1</td>
</tr>
<tr>
<td>35-39</td>
<td>&lt; 1</td>
</tr>
<tr>
<td>40-44</td>
<td>2</td>
</tr>
</tbody>
</table>

HDL, high density lipoprotein; SBP, systolic blood pressure; ECG-LVH, left ventricular hypertrophy by electrocardiography.

measured. These were chosen because they can be measured objectively and because they make independent contributions to the risk equations. Other important risk influences have not been included in the equations, but they should not be underestimated. For example, heredity is a factor of major importance, but information on this subject is difficult to quantify or even obtain accurately. Individuals with family histories of heart disease should view increased risk factors with particular concern and deal with them more vigorously than should persons without such backgrounds. Obesity is another exam-
ple of a significant risk factor in long-term predictions; however, in shorter-term studies such as this one, its effects tend to be mediated by other risk factors. Hence, it is not included in the equations.

Second, the predictions may not be appropriate for individuals with extremely elevated risk factors such as malignant hypertension and severe diabetes mellitus or extremely high cholesterol and HDL cholesterol levels that place them in the top or bottom few percentiles of the distributions.

Third, the equations may not be directly applicable to populations with very low CHD incidence rates. Although these equations are considered more accurate than earlier versions when incidence rates are low, this has not been tested extensively. Hence, the equations may be inappropriate for use with populations from countries or ethnic groups that have CHD incidence rates that are much lower or higher than the range presented here.

When generalization seems reasonable, estimation of CHD risk can be useful in projecting patient progress in clinics at which preventive cardiology is the goal, such as those concentrating on lipid or blood pressure influence. Risk factor scores and calculations can be discussed with patients and provide a framework for intervention. Clinicians can predict possible rewards in the form of improved risk profiles for patients who make the appropriate although often difficult changes in smoking, eating, and exercise habits.

References
