

Validation of the Framingham Coronary Heart Disease Prediction Scores

Results of a Multiple Ethnic Groups Investigation

Ralph B. D'Agostino, Sr, PhD

Scott Grundy, MD, PhD

Lisa M. Sullivan, PhD

Peter Wilson, MD

for the CHD Risk Prediction Group

THE FRAMINGHAM HEART STUDY has developed mathematical functions for predicting risk of clinical coronary heart disease (CHD) events.¹⁻⁵ These are derived multivariable mathematical functions that assign weights to major CHD risk factors such as sex, age, blood pressure, total cholesterol (TC), low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), smoking behavior, and diabetes status. For a person free of cardiovascular disease, his/her CHD risk factors are entered into the function to produce a probability estimate of developing CHD within a certain time period (eg, the next 5 years). Recently, Framingham investigators developed a simplified model that incorporates blood pressure and cholesterol categories proposed by the Fifth Joint National Committee on Hypertension (JNC-V) and the National Cholesterol Education Program, Adult Treatment Panel II (NCEP-ATP II).⁵⁻⁸

The Framingham functions were developed to assess the relative importance of CHD risk factors and to quantify the absolute level of CHD risk for individual patients. The report of the third adult treatment panel (NCEP-

Context The Framingham Heart Study produced sex-specific coronary heart disease (CHD) prediction functions for assessing risk of developing incident CHD in a white middle-class population. Concern exists regarding whether these functions can be generalized to other populations.

Objective To test the validity and transportability of the Framingham CHD prediction functions per a National Heart, Lung, and Blood Institute workshop organized for this purpose.

Design, Setting, and Subjects Sex-specific CHD functions were derived from Framingham data for prediction of coronary death and myocardial infarction. These functions were applied to 6 prospectively studied, ethnically diverse cohorts (n=23 424), including whites, blacks, Native Americans, Japanese American men, and Hispanic men: the Atherosclerosis Risk in Communities Study (1987-1988), Physicians' Health Study (1982), Honolulu Heart Program (1980-1982), Puerto Rico Heart Health Program (1965-1968), Strong Heart Study (1989-1991), and Cardiovascular Health Study (1989-1990).

Main Outcome Measures The performance, or ability to accurately predict CHD risk, of the Framingham functions compared with the performance of risk functions developed specifically from the individual cohorts' data. Comparisons included evaluation of the equality of relative risks for standard CHD risk factors, discrimination, and calibration.

Results For white men and women and for black men and women the Framingham functions performed reasonably well for prediction of CHD events within 5 years of follow-up. Among Japanese American and Hispanic men and Native American women, the Framingham functions systematically overestimated the risk of 5-year CHD events. After recalibration, taking into account different prevalences of risk factors and underlying rates of developing CHD, the Framingham functions worked well in these populations.

Conclusions The sex-specific Framingham CHD prediction functions perform well among whites and blacks in different settings and can be applied to other ethnic groups after recalibration for differing prevalences of risk factors and underlying rates of CHD events.

JAMA. 2001;286:180-187

www.jama.com

ATP III) endorses knowledge of absolute CHD risk as a means of identifying those patients who are likely to benefit from aggressive primary prevention strategies and as a tool motivating patients to comply with them.⁹

The Framingham Heart Study consists of white middle-class individuals. Concern exists as to the generaliz-

Author Affiliations: Departments of Mathematics and Statistics (Drs D'Agostino and Sullivan) and Medicine (Dr Wilson), Boston University, Boston, Mass; the Framingham Study, Framingham, Mass (Drs D'Agostino, Sullivan, and Wilson); and Center for Human Nutrition, University of Texas Southwestern Medical Center, Dallas (Dr Grundy).

Members of the CHD Risk Prediction Group are listed at the end of this article.

Corresponding Author and Reprints: Ralph B. D'Agostino, Sr, PhD, Department of Mathematics and Statistics, Boston University, 111 Cummington St, Boston, MA 02215 (e-mail: ralph@bu.edu).

ability of its CHD risk function to populations such as other whites, blacks, Asian Americans, Hispanics, and Native Americans. In January 1999 the National Heart, Lung, and Blood Institute convened a CHD Prediction Workshop to evaluate the performance of Framingham functions in non-Framingham populations.¹⁰

METHODS

Framingham Models

Sex-specific Framingham CHD risk functions were derived from 2439 men and 2812 women, 30 to 74 years of age, who were free of cardiovascular disease (CVD) at the time of their Framingham Heart Study examinations in 1971 to 1974. Participants attended either the 11th examination of the original Framingham Cohort^{4,5} or the initial examination of the Framingham Offspring Study.¹¹ Coronary heart disease risk factors were routinely and systematically evaluated during these examinations as described in detail elsewhere.⁵ Twelve-year follow-up was obtained for the development of "hard" CHD events, defined as coronary death or myocardial infarction. Sex-specific Cox proportional hazards regression functions were computed that relate JNC-V blood pressure and NCEP-ATP II cholesterol categories, along with age, current smoking, and presence of diabetes to the occurrence of hard CHD events.

Non-Framingham Cohorts

Six non-Framingham cohorts were identified for evaluation.¹²⁻¹⁷ Criteria for selection were similar age range, systematic measurement of CHD risk factors, and adequate length of follow-up for hard CHD events. The selected cohorts were participants in the Atherosclerosis Risk in Communities Study (ARIC, 1987-1988), the Physicians' Health Study (PHS, 1982), the Honolulu Heart Program (HHP, 1980-1982), the Puerto Rico Heart Health Program (PR, 1965-1968), the Strong Heart Study (SHS, 1989-1991), and the Cardiovascular Health Study (CHS, 1989-1990). The PHS is a prospective,

nested case-control study with 1-to-4 matching of cases to controls for age and smoking.

For each cohort, sex-specific Cox regression functions were derived using the same variables as in the Framingham functions but using data from the individual cohorts. We call these the cohorts' "own" functions. They represent the best possible Cox prediction functions for each cohort based on specific prevalences of risk factors and CHD event rates. For the ARIC study, which includes white and black subjects, the Cox regression functions were sex- and race-specific. The CHS cohort included subjects 65 to 88 years old. We used only CHS subjects aged 65 to 74 years for the CHS' own functions.

Statistical Analysis

All analyses were sex- and race-specific. The performance of the Framingham prediction functions among the non-Framingham cohorts was assessed according to 3 evaluations: equality of regression coefficients (relative risk [RR] comparison), discrimination, and calibration.

Relative Risk Comparison. For each risk factor, Cox proportional hazards modeling yielded regression coefficients for the Framingham and non-Framingham cohorts. To compare these coefficients a test statistic z was calculated, where $z = (b[F] - b[O])/SE$, and where $b(F)$ and $b(O)$ are, respectively, the regression coefficients of the Framingham and the other cohort's model, while SE is the standard error of the difference in the coefficients. This was computed as the square root of the sum of the squares of the SEs for the 2 coefficients. Because the RR of a variable is computed by exponentiating its regression coefficient, the z statistic tests the equality of RRs between Framingham and non-Framingham cohorts.

Discrimination. Discrimination is the ability of a prediction model to separate those who experience hard CHD events from those who do not. We quantified this by calculating the c statistic, analogous to the area under a receiver operating characteristic (ROC)

curve¹⁸⁻²⁰; this value represents an estimate of the probability that a model assigns a higher risk to those who develop CHD within a 5-year follow-up than to those who do not.^{18,19} For each non-Framingham cohort 2 c statistics were computed, one applying the Framingham function to the cohort and the other from the cohort's own prediction function. These were compared using a test developed by Nam.²⁰

Calibration. Calibration measures how closely predicted outcomes agree with actual outcomes. For this we used a version of the Hosmer-Lemeshow χ^2 statistic.^{19,20} For each non-Framingham cohort, the Framingham function's predicted risks were used to divide subjects into deciles of predicted risk for experiencing a hard CHD event within 5 years. Plots were constructed showing predicted and actual event rates for each decile. A χ^2 statistic was calculated to compare the differences between predicted and actual event rates; small values indicated good calibration. Values exceeding 20 indicate significant lack of calibration ($P < .01$). For further evaluation of calibration, we compared this χ^2 statistic with one derived from each cohort's own prediction function. All statistical analyses were performed in SAS version 6.12 (SAS Institute, Cary, NC). Because the PHS is a case-control study it is not suitable for calibration comparisons.

Recalibration

When there is a systematic overestimation or underestimation of risk, transporting a prediction function from one setting to another requires a process known as recalibration. The Framingham Cox regression models have the form $S_0(t)^{\exp(f(x,M))}$ where $f(x,M) = b_1(x_1 - M_1) + \dots + b_p(x_p - M_p)$. Here b_1, \dots, b_p are the regression coefficients (logs of the RRs), x_1, \dots, x_p represent an individual's risk factors, and M_1, \dots, M_p are the means of the risk factors of the Framingham cohort. $S_0(t)$ is the Framingham average incidence rate at time t or, more precisely, the survival rate at the mean values of the risk factors. With recalibration, the

Framingham mean values of the risk factors (M_1, \dots, M_p) are replaced by the mean values of the risk factors from a non-Framingham cohort, while the Framingham average incidence rate $S_0(t)$ is replaced by the cohort's own average incidence rate. We used Kaplan-Meier estimates to determine average incidence rates.^{20,21} It is important to note that recalibration does not affect RR comparisons or discrimination evaluations.

RESULTS

Baseline Characteristics and Framingham Cohort Coefficients

The Framingham Study cohort consisted of 2439 men and 2812 women free of cardiovascular disease. The 5- and 10-year hard CHD event rates were 3.7% and 8.0% for men and 1.4% and 2.8% for women.

TABLE 1 shows the Cox regression coefficients for the sex-specific Framingham regression models.

TABLE 2 contains the racial compositions, sample sizes, age ranges, mean ages, risk factor distributions, and 5-year incidence rates for the Framingham and non-Framingham cohorts.

Relative Risk Comparisons

TABLE 3 and TABLE 4 contain the RRs of the CHD risk factors for each cohort's sex- and race-specific regression model. First, we considered each function separately. Among men most risk factors had statistically significant coefficients, whereas among women a number of coefficients were not statistically significant, presumably because of low event rates (eg, for ARIC black women there were only 38 events among 2333 subjects). Nonetheless, within risk factor categories, trends were significant. For example, except for the SHS Native Americans, for all cohorts the risk for hard CHD events increased as blood pressure went from optimal to stage II-IV hyperten-

sion, as TC increased, and as HDL-C decreased ($P < .01$ for all). In the SHS there were some unexpected but not statistically significant elevated risks in high HDL-C groups.

Among men, there were no significant differences between Framingham RRs and those of ARIC white and black men. There were differences in RRs for smoking and age in the PHS cohort, which may reflect the matching scheme that was used in that study. Among HHP Japanese American men and PR Hispanic men, RRs were lower for optimal blood pressure. Smoking was associated with a lower RR among HHP Japanese American men, while diabetes and TC 280 mg/dL (7.25 mmol/L) or higher were associated with much higher RRs among SHS Native American men. Also, in this cohort HDL-C in the range of 50 to 59 mg/dL (1.30-1.53 mmol/L) had an unexpected elevated risk and stage I hypertension had an unexpected low risk, both resulting in significant differences from the Framingham function. Among the more elderly CHS white men, cholesterol abnormalities were associated with a lower RR.

Among women, there were no differences between Framingham RRs and those of white ARIC or CHS women. Black women in the ARIC cohort had higher RRs for high normal blood pressure and stage II-IV hypertension. In the SHS there were significant differences for diabetes and smoking. Also, HDL-C greater than or equal to 60 mg/dL (1.55 mmol/L) carried an elevated risk in the SHS, resulting in a significantly different RR than that of the Framingham function.

Discrimination

TABLE 5 contains the *c* statistics for both men and women. The "FHS" row refers to the discrimination achieved by applying the Framingham prediction functions to the non-Framingham cohorts, while the "Best Cox" row contains the *c* statistics resulting from the cohort's own Cox regression function. Since they are based on the same cohorts for which the scores were developed, the latter *c* statistics are overestimates. For non-

Table 1. Framingham Functions (Cox Regression Coefficients) for Hard CHD Events (Coronary Death or MI)*

Risk Factor	Men	Women
Age, y	0.05	0.17
Age ²	NA	-0.001
Blood pressure, mm Hg		
Optimal (S<120, D<80)	0.09	-0.74
Normal (S<130, D<85)†	0.00	0.00
High normal (S<140, D<90)	0.42	-0.37
Stage I Htn (S<160, D<100)	0.66	0.22
Stage II-IV Htn (S≥160, D≥100)	0.90	0.61
Total cholesterol, mg/dL‡		
<160	-0.38	-0.21
160-199†	0.00	0.00
200-239	0.57	0.44
240-279	0.74	0.56
≥280	0.83	0.89
HDL-C, mg/dL‡		
<35	0.61	0.73
35-44	0.37	0.60
45-49§	0.00	0.60
50-59	0.00	0.00
≥60	-0.46	-0.54
Diabetes	0.53	0.87
Current smoking	0.73	0.98

*CHD indicates coronary heart disease; MI, myocardial infarction; NA, not applicable; S, systolic; D, diastolic; Htn, hypertension; and HDL-C, high-density lipoprotein cholesterol.

†Referent category for men and women.

‡To convert mg/dL to mmol/L, multiply values for total cholesterol and HDL-C by 0.0259.

§Referent category for men.

||Referent category for women.

Framingham white men (ARIC, PHS, and CHS) and women (ARIC and CHS), the Framingham functions discriminated well, almost always achieving the same discrimination as best Cox functions. Overall, within sampling fluctuations the Framingham functions discriminated nearly as well as the best Cox functions of the non-Framingham cohorts, with the exception of the SHS Native Americans.

Calibration

Table 5 also contains the χ^2 statistics for evaluation of the calibration of the

Framingham prediction functions applied to non-Framingham cohorts. For white men and women, including the more elderly subjects in the CHS cohort, both the Framingham functions ("Unadjusted" row) and the individual cohort's own functions ("Best Cox" row) showed a statistically acceptable calibration. FIGURE 1 contains calibration plots for white and black men and women from the ARIC study. In general, actual CHD event rates were similar to event rates predicted by Framingham functions among white and black men and women.

For the HHP Japanese American men and the PR Hispanic men, the calibration χ^2 statistics of 66.0 and 142.0, respectively, indicate poor calibration (Table 5, "Unadjusted" row and FIGURE 2A and B, "Unadjusted" panels). The Framingham prediction function systematically overestimated risk in both cohorts, in which the overall CHD event rates were substantially lower. Model recalibration using the non-Framingham cohorts' mean values for risk factors and CHD incidence rates substantially improved the performance of the Framingham prediction

Table 2. Description of Studies Used in Evaluation*

	Men								Women					
	FHS		ARIC		PHS†	HHP		PR	SHS		CHS	SHS		CHS
	White (n = 2439)	White (n = 4705)	Black (n = 1428)	White (n = 901)	Japanese American (n = 2755)	Hispanic (n = 8713)	Native American (n = 1527)	White (n = 956)	White (n = 2812)	White (n = 5712)	Black (n = 2333)	Native American (n = 2255)	White (n = 1601)	
Age range, y	30-74	44-66	44-66	40-74	51-81	35-74	45-75	65-74	30-74	44-66	44-66	45-75	65-74	
Mean age, y	48.3	54.6	53.7	57.6	61.9	54.1	55.4	69.7	49.6	53.9	53.3	56.5	69.3	
Blood pressure, mm Hg														
Optimal (S<120, D<80)	20	50	27	13	18	24	27	23	35	58	33	37	28	
Normal (S<130, D<85)	24	23	21	34	18	23	25	20	21	18	22	22	19	
High normal (S<140, D<90)	20	14	18	30	20	18	22	22	15	11	17	17	18	
Stage I Htn (S<160, D<100)	23	10	22	21	29	21	20	23	19	9	19	17	24	
Stage II-IV Htn (S≥160, D≥100)	13	3	12	3	14	13	7	12	10	2	9	7	9	
Total cholesterol, mg/dL‡														
<160	7	8	9	5	4	14	22	11	8	6	8	20	4	
160-199	31	32	35	25	29	37	41	42	30	29	30	42	23	
200-239	39	39	33	39	41	33	29	36	33	39	34	25	41	
240-279	17	16	16	22	21	12	7	9	20	19	20	10	24	
≥280	6	4	7	10	5	4	1	2	9	7	9	3	8	
HDL-C, mg/dL‡														
<35	19	25	13	18	15	0	25	10	4	6	4	12	2	
35-44	36	36	28	31	36	50	40	33	15	18	18	34	14	
45-49	15	14	13	14	15	50	11	19	12	12	13	16	13	
50-59	19	16	24	21	20	0	13	21	28	25	25	21	27	
≥60	11	9	22	16	14	0	10	16	41	40	40	16	45	
Diabetes	5	6	14	5	14	7	42	15	4	6	17	51	10	
Current smoking	40	24	38	16	32	44	40	12	38	25	25	30	15	
5-Year CHD Rate	0.0373	0.0317	0.0322	0.0200	0.0279	0.0123	0.0301	0.0743	0.0139	0.0091	0.0163	0.0102	0.0275	

*Table entries are percentages of sample with the exception of age. FHS indicates Framingham Heart Study; ARIC, Atherosclerosis Risk in Communities Study; PHS, Physicians' Health Study; HHP, Honolulu Heart Program; PR, Puerto Rico Heart Health Program; SHS, Strong Heart Study; CHS, Cardiovascular Health Study; S, systolic; D, diastolic; Htn, hypertension; and HDL-C, high-density lipoprotein cholesterol.

†PHS is a nested case-control study with 1-to-4 matching of cases to controls.

‡To convert mg/dL to mmol/L, multiply values for total cholesterol and HDL-C by 0.0259.

functions (Figure 2, “Adjusted” panels, and Table 5). In the SHS, calibration was good for men ($\chi^2=10.6$), but less good for women ($\chi^2=22.7$). Recalibration resulted in improved performance of the Framingham functions (Figure 2 and Table 5).

COMMENT

The Framingham CHD prediction functions were developed to help clinicians estimate the absolute risk of any individual patient developing clinically manifest disease. We sought to demonstrate the external validity of the

Framingham functions by examining their performance in 6 different well-described population-based cohorts that reflect a wide range of ethnic diversity.

As shown in Tables 3 and 4, RRs for major CHD risk factors were remarkably similar to those derived from the Framingham Heart Study cohort among white men and women and black men in the ARIC cohort. Among black women in the ARIC cohort, RRs for elevated blood pressure were somewhat higher. In the cohorts that were made up of other ethnic groups, however, we

did note some differences in RRs. For example, smoking was associated with a much lower RR in HHP Japanese American men for reasons that are not clear. In the CHS cohort, cholesterol abnormalities and smoking had lower RRs, possibly due to age interactions. In the SHS Native American cohort, there were RR differences for cholesterol abnormalities and diabetes. Some cholesterol differences are unexplained, with high HDL-C levels carrying an increased risk in the SHS cohort. Because the prevalence of diabetes among Native Americans is quite high,

Table 3. Relative Risks for CHD Risk Factors: Men by Study*

	Best Cox Model From Other Studies, RR (95% CI)							
	FHS	ARIC		PHS	HHP	PR	SHS	CHS
	White	White	Black	White	Japanese American	Hispanic	Native American	White
Age	1.05 (1.04-1.07)	1.05 (1.02-1.07)	1.05 (1.01-1.10)	1.01† (1.00-1.02)	1.06 (1.04-1.08)	1.03† (1.00-1.05)	1.05 (1.02-1.09)	1.07 (0.99-1.16)
Blood pressure, mm Hg								
Optimal (S<120, D<80)	1.10 (0.67-1.82)	0.75 (0.53-1.06)	1.25 (0.48-3.23)	1.15 (0.83-1.59)	0.48† (0.23-0.98)	0.57† (0.35-0.91)	0.66 (0.28-1.54)	0.69 (0.29-1.65)
Normal (S<130, D<85)	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference
High normal (S<140, D<90)	1.53 (0.98-2.36)	0.93 (0.60-1.44)	2.67 (1.10-6.48)	1.25 (0.97-1.61)	1.23 (0.73-2.08)	1.22 (0.82-1.83)	1.01 (0.46-2.21)	2.05 (1.02-4.11)
Stage I Htn (S<160, D<100)	1.93 (1.28-2.92)	1.66 (1.11-2.47)	3.07 (1.31-7.18)	1.67 (1.28-2.16)	1.45 (0.90-2.34)	1.79 (1.25-2.57)	0.70† (0.30-1.64)	2.28 (1.16-4.51)
Stage II-IV Htn (S≥160, D≥100)	2.45 (1.59-3.79)	1.72 (0.94-3.17)	1.55 (0.52-4.65)	2.05 (1.21-3.49)	1.42 (0.82-2.45)	3.26 (2.29-4.65)	1.40 (0.52-3.76)	2.90 (1.40-5.99)
Total cholesterol, mg/dL								
<160	0.69 (0.31-1.52)	0.92 (0.51-1.67)	0.85 (0.31-2.31)	0.73 (0.44-1.22)	1.04 (0.49-2.23)	0.94 (0.59-1.49)	1.56 (0.69-3.56)	1.02 (0.51-2.07)
160-199	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference
200-239	1.77 (1.25-2.50)	1.40 (1.00-1.97)	1.32 (0.72-2.43)	1.07† (0.83-1.38)	0.94† (0.63-1.39)	1.21 (0.89-1.65)	2.80 (1.35-5.79)	1.03† (0.65-1.64)
240-279	2.10 (1.43-3.10)	1.93 (1.31-2.83)	1.10 (0.50-2.46)	1.60 (1.21-2.11)	1.72 (1.13-2.61)	1.36 (0.90-2.06)	3.01 (1.05-8.59)	0.96† (0.42-2.17)
≥280	2.29 (1.39-3.76)	2.23 (1.25-3.97)	2.23 (0.97-5.16)	1.57 (1.12-2.21)	2.17 (1.14-4.13)	2.08 (1.26-3.44)	8.97† (2.48-32.37)	2.33 (0.83-6.57)
HDL-C, mg/dL								
<35	1.84 (1.17-2.88)	2.66 (1.55-4.56)	1.82 (0.76-4.37)	1.90 (1.37-2.64)	1.28 (0.75-2.19)	1.00 (1.00-1.00)	1.83 (0.52-6.45)	0.84† (0.40-1.75)
35-44	1.45 (0.94-2.21)	1.97 (1.16-3.36)	1.16 (0.51-2.65)	1.24 (0.91-1.69)	1.33 (0.84-2.09)	0.73† (0.50-1.07)	1.48 (0.43-5.06)	0.83† (0.49-1.39)
45-49	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference
50-59	1.00 (0.62-1.60)	1.40 (0.76-2.59)	0.71 (0.28-1.78)	0.88 (0.62-1.24)	0.62 (0.35-1.11)	1.00 (1.00-1.00)	3.51† (0.97-12.66)	0.41† (0.20-0.83)
≥60	0.63 (0.34-1.18)	0.55 (0.21-1.40)	0.71 (0.28-1.78)	0.76 (0.52-1.10)	0.67 (0.35-1.27)	1.00 (1.00-1.00)	1.48 (0.33-6.69)	0.61 (0.29-1.25)
Diabetes	1.69 (1.11-2.57)	2.42 (1.69-3.46)	1.40 (0.75-2.62)	1.54 (1.05-2.26)	2.55 (1.82-3.57)	2.07 (1.50-2.85)	4.29† (2.27-8.10)	1.47 (0.89-2.44)
Current smoking	2.07 (1.60-2.68)	1.94 (1.47-2.56)	1.65 (1.00-2.71)	1.02† (0.79-1.33)	1.27† (0.90-1.79)	1.74 (1.38-2.20)	1.66 (0.94-2.91)	1.58 (0.90-2.78)

*RR indicates relative risk; CI, confidence interval. For expansion of other terms and for cholesterol conversion factor, see Table 2 footnote.
†RR significantly different from that of the Framingham Study ($P<.10$).

it is possible that the different RRs we observed may be due to interactions with other risk factors and with other factors unique to diabetes, such as albuminuria, that were not considered in our analyses.

The ability of the Framingham prediction functions to discriminate be-

tween subjects who developed clinical CHD and those who did not was reasonably good for most of the non-Framingham cohorts (Table 5). Among ARIC black women, the Framingham *c* statistic was numerically, but not significantly, lower than that derived from the model based on that same cohort's

data. Since this ARIC *c* statistic is based on the same data with which its function was developed, it is an overestimate and the difference may relate to this. It may also be due to the small number of CHD events. The Framingham *c* statistic for CHS men was also numerically, but not significantly, lower

Table 4. Relative Risks for CHD Risk Factors: Women by Study*

	Best Cox Model From Other Studies, RR (95% CI)				
	FHS	ARIC		SHS	CHS
	White	White	Black	Native American	White
Age†	1.19 (0.97-1.45)	1.67 (0.73-3.83)	0.88 (0.33-2.33)	1.47 (0.71-3.05)	2.83 (0.02-530.85)
Blood pressure, mm Hg					
Optimal (S<120, D<80)	0.48 (0.22-1.05)	0.42 (0.24-0.73)	0.71 (0.24-2.13)	1.16 (0.43-3.18)	0.60 (0.25-1.45)
Normal (S<130, D<85)	Reference	Reference	Reference	Reference	Reference
High normal (S<140, D<90)	0.69 (0.34-1.42)	1.12 (0.61-2.08)	2.82‡ (1.10-7.23)	0.44 (0.11-1.73)	1.13 (0.50-2.59)
Stage I Htn (S<160, D<100)	1.24 (0.69-2.24)	1.17 (0.59-2.28)	1.94 (0.75-5.02)	0.81 (0.28-2.34)	1.80 (0.87-3.70)
Stage II-IV Htn (S≥160, D≥100)	1.84 (1.00-3.39)	2.80 (1.24-6.36)	8.86‡ (3.81-20.63)	2.19 (0.81-5.89)	2.07 (0.87-4.93)
Total cholesterol, mg/dL					
<160	1.23 (0.27-5.64)	0	1.46 (0.51-4.17)	0.94 (0.33-2.68)	0.53 (0.07-4.18)
160-199	Reference	Reference	Reference	Reference	Reference
200-239	1.55 (0.81-2.96)	0.90 (0.52-1.56)	1.17 (0.56-2.48)	1.29 (0.54-3.04)	1.88 (0.88-4.01)
240-279	1.74 (0.90-3.40)	1.16 (0.64-2.10)	1.36 (0.61-2.99)	3.28 (1.33-8.08)	2.66 (1.22-5.79)
≥280	2.44 (1.21-4.93)	1.48 (0.75-2.93)	2.98 (1.35-6.58)	3.21 (0.71-14.63)	1.71 (0.57-5.13)
HDL-C, mg/dL					
<35	2.08 (1.00-4.31)	2.34 (1.21-4.52)	3.08 (1.20-7.86)	5.36 (1.10-26.17)	1.21 (0.16-9.30)
35-44	1.82 (1.05-3.16)	0.84 (0.44-1.61)	1.80 (0.88-3.71)	3.86 (0.89-16.82)	1.59 (0.78-3.27)
45-49	1.82 (1.05-3.14)	1.80 (0.99-3.28)	1.08 (0.44-2.63)	4.09 (0.85-19.82)	1.35 (0.62-2.98)
50-59	Reference	Reference	Reference	Reference	Reference
≥60	0.58 (0.33-1.02)	0.47 (0.23-0.93)	0.87 (0.41-1.86)	2.87‡ (0.52-15.83)	0.90 (0.48-1.69)
Diabetes	2.38 (1.40-4.06)	3.62 (2.21-5.94)	2.01 (1.16-3.48)	8.63‡ (2.55-29.16)	2.29 (1.23-4.23)
Current smoking	2.65 (1.77-3.97)	3.92 (2.56-5.99)	2.78 (1.64-4.73)	0.55‡ (0.21-1.44)	1.48 (0.78-2.80)

*RR indicates relative risk; CI, confidence interval. For expansion of other terms and cholesterol conversion factor, see Table 2 footnote.

†Values for age are given in 1-year increments. Relative risks for age² are uninformative and are not given.

‡RR significantly different from that of the Framingham Study (*P*<.10).

Table 5. Summary of Discrimination and Calibration Evaluations*

	Men							Women							
	FHS		ARIC		PHS	HHP	PR	SHS	CHS	FHS		ARIC		SHS	CHS
	White	White	Black	White	White	Japanese American	Hispanic	Native American	White	White	White	Black	Native American	White	
Discrimination: <i>c</i> statistic†															
FHS	0.79	0.75	0.67	0.63	0.72	0.69	0.69	0.63	0.83	0.83	0.79	0.75	0.75	0.66	
Best Cox	0.79	0.76	0.70	0.64	0.74	0.72	0.77	0.69	0.83	0.84	0.85	0.86	0.86	0.68	
Calibration: χ^2 statistic‡															
Unadjusted	...	13.8	6.2	...	66.0	142.0	10.6	13.2	3.7	5.3	5.0	22.7	10.4	...	
Adjusted	12.0	10.0	8.2	9.8	
Best Cox	3.3	5.4	7.2	NA§	2.6	7.2	2.7	6.8	3.7	5.2	3.4	3.5	6.8	...	

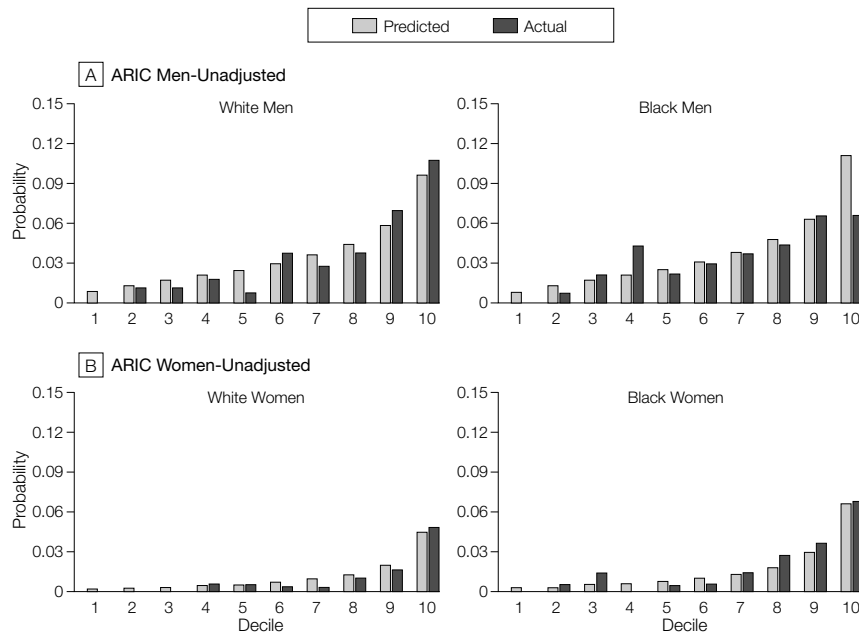
*NA indicates not applicable; ellipses (...), that there is no need for adjustment when unadjusted values are 20 or less. For expansion of other terms, see Table 2 footnote.

†FHS *c* statistic indicates discrimination achieved by applying Framingham function to the other study's data; Best Cox *c* statistic indicates discrimination achieved by applying the Cox model developed on the other study's data.

‡Unadjusted χ^2 statistic indicates calibration achieved by applying the Framingham function to the other study's data (using the Framingham means on the risk factors and the Framingham coronary heart disease [CHD] incidence); adjusted χ^2 statistic indicates calibration achieved by applying the Framingham function to the other study's data with the other study's means on the risk factors and the other study's CHD incidence; Best Cox χ^2 statistic indicates calibration achieved by applying the other study's Cox regression model (using the other study's means on the risk factors and the other study's CHD incidence).

§PHS is a nested case-control study with 1-to-4 matching of cases to controls and is thus not appropriate for calibration evaluation.

Figure 1. Five-Year Predictions for Hard CHD Events: Performance Measures for ARIC Men and Women

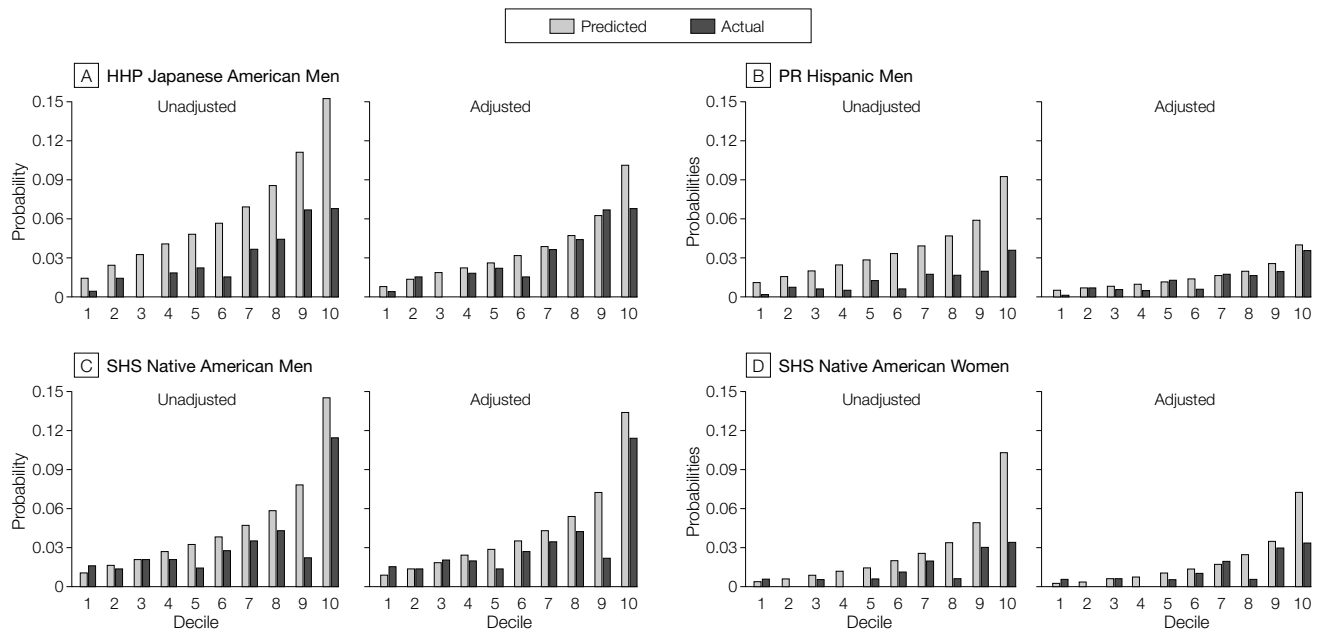


X-axes refer to decile of predicted risk based on the Framingham Heart Study function. ARIC indicates Atherosclerosis Risk in Communities Study. Hard CHD events were coronary death or myocardial infarction.

than the CHS' own function c statistic. The difference may relate to the overestimate of the CHS cohort's own function c statistic. The reasons why both the Framingham and CHS cohort's own function c statistics are low may be a consequence of the relatively small sample size and the narrow age distribution. The c statistics were appreciably decreased for SHS Native Americans. Why discrimination was worse for Native Americans compared with that for white and black men and women of the ARIC cohort, PR Hispanic men, and HHP Japanese American men is not clear. It is possible that the markedly different RR estimates for cholesterol and diabetes among the SHS Native Americans may have adversely affected the ability of the Framingham prediction function to discriminate CHD risk.

In our model calibration analyses, we found reasonably good agreement between predicted and actual CHD event rates for all of the non-Framingham

Figure 2. Five-Year Predictions for Hard CHD Events: Performance Measures for HHP, PR, and SHS Subjects



X-axes refer to decile of predicted risk based on the Framingham Heart Study function. HHP indicates Honolulu Heart Program; PR, Puerto Rico Heart Health Program; and SHS, Strong Heart Study. Hard CHD events were coronary death or myocardial infarction.

cohorts studied (Figures 1 and 2, and Table 5) except for HHP Japanese American men, PR Hispanic men, and SHS Native American women. In these groups, the Framingham prediction functions systematically overestimated CHD risk (Figure 2A, B, and D). This overestimation was corrected by using a process of recalibration. In order to apply this to other such populations, it would be necessary to obtain cross-sectional data on risk factor prevalences as well as population data on CHD event rates over time.

Authors of treatment guidelines have recognized the need to have an accurate and reliable multivariable-based estimate of absolute CHD risk in order to best identify those most in need of aggressive preventive treatment.⁷⁻⁹ Thus, the recently released NCEP-ATP III guidelines specifically recommend that the level of treatment should relate to the level of CHD risk.⁹ They specifically incorporate Framingham prediction functions to aid clinicians and patients in determining optimal strategies.

Simple charts can be used to aid in this activity.^{5,9}

In order for multivariable risk assessment and treatment guidelines to have optimal use and acceptability, clinicians need to be confident that absolute risk prediction functions can be transported to other settings beyond where they were originally developed. We have demonstrated that the FHS prediction functions work reasonably well among white and black men and women. When applied to Japanese American and Hispanic men and Native American women, recalibration was needed. Future work is needed to devise practical schemes by which clinicians can confidently apply the FHS prediction functions in these groups.

Author Contributions: Study concept and design: D'Agostino, Grundy, Wilson.

Acquisition of data: D'Agostino, Grundy, Wilson.

Analysis and interpretation of data: D'Agostino, Grundy, Sullivan, Wilson.

Drafting of the manuscript: D'Agostino, Grundy.

Critical revision of the manuscript for important intellectual content: D'Agostino, Grundy, Sullivan, Wilson.

Statistical expertise: D'Agostino, Sullivan, Wilson.

Obtained funding: D'Agostino, Grundy.

Administrative, technical, or material support: D'Agostino, Grundy, Sullivan.

Study supervision: D'Agostino, Wilson.

Members of the CHD Risk Prediction Group: Boston University and the Framingham Study: Philip A. Wolf, MD, Daniel Levy, MD, Joseph Massaro, PhD, Byung-Ho Nam, PhD; CHD Risk Prediction Planning Committee (NIH): National Heart, Lung, and Blood Institute: James L. Cleeman, MD, Jeffrey A. Cutler, MD, Lawrence Friedman, MD, Edward Rocella, MD; CHD Risk Prediction Planning Committee (non-NIH): Scott M. Grundy, MD (chair), Ralph B. D'Agostino, Sr, PhD, Gregory Burke, MD, Lori Mosca, MD, Daniel Rader, MD, Peter W. F. Wilson, MD; Atherosclerosis Risk in Communities Study: Lloyd E. Chambless, PhD, David J. Couper, PhD; Physicians' Health Study: Meir J. Stampfer, MD, Jing Ma, PhD; Honolulu Heart Program: David Curb, MD, B. R. Rodriguez, MD, Robert Abbott, PhD; Puerto Rico Heart Health Program: Mario R. Gacia-Palmieri, MD, Paul Sorlie, PhD, Sean Coady, PhD; Strong Heart Study: Elisa Lee, PhD, Barbara Howard, MD; Cardiovascular Health Study: Richard Kronmal, PhD, Thomas Lumney, PhD.

Funding/Support: The Framingham Study work and analyses for the workshop carried out at Boston University were supported by a contract (N01-HC-380380) and funds from the National Heart, Lung, and Blood Institute of the National Institutes of Health.

Acknowledgment: We especially thank Claude Lenfant, MD, Director of the National Heart, Lung, and Blood Institute, for his support and organization of the Workshop on CHD Risk Assessment, January 1999, and to the cardiovascular studies (Atherosclerosis Risk in Community Study, Physicians' Health Study, Honolulu Health Project, Puerto Rico Heart Health Program, Strong Heart Study, and the Cardiovascular Health Study) who supported the workshop and this investigation by contributing their data and performing numerous analyses.

REFERENCES

1. *Coronary Risk Handbook: Estimating Risk of Coronary Heart Disease in Daily Practice*. New York, NY: American Heart Association; 1973:1-35.
2. Kannel WB, McGee D, Gordon T. A general cardiovascular risk profile: the Framingham Study. *Am J Cardiol*. 1976;38:46-51.
3. Gordon T, Kannel WB. Multiple risk functions for predicting coronary heart disease: the concept, accuracy, and application. *Am Heart J*. 1982;103:1031-1039.
4. Anderson M, Wilson PW, Odell PM, Kannel WB. An updated coronary risk profile: a statement for health professionals. *Circulation*. 1991;83:356-362.
5. Wilson PWF, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation*. 1998;97:1837-1847.
6. The National Cholesterol Education Program Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel II). National Cholesterol Education Program, second report. *Circulation*. 1994;89:1333-1445.
7. The National Cholesterol Education Program Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel II). Summary of the second report of the NCEP Expert Panel (Adult Treatment Panel II). *JAMA*. 1993;269:3015-3023.
8. Joint National Committee on Detection, Evaluation, and Treatment of High Blood Pressure (JNC-V). The fifth report of the Joint National Committee on Detection, Evaluation, and Treatment of High Blood Pressure (JNC-V). *Arch Intern Med*. 1993;153:154-183.
9. National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). 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.
10. Grundy SM, D'Agostino RB Sr, Mosca L, et al. Cardiovascular risk assessment based on US cohort studies: findings from a National Heart, Lung, and Blood Institute workshop. *Circulation*. 2001;104:1-6.
11. Kannel WB, Feinlieb 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.
12. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. *Am J Epidemiol*. 1989;129:687-702.
13. Stampfer MJ, Sacks FM, Salvini S, Willett WC, Hennekens CH. A prospective study of cholesterol, apolipoproteins, and the risk of myocardial infarction. *N Engl J Med*. 1991;325:373-381.
14. Kagan A, Gordon T, Rhoads GG, Schiffman JC. Some factors related to coronary heart disease incidence in Honolulu Japanese men: the Honolulu Heart Study. *Int J Epidemiol*. 1975;4:271-279.
15. Garcia-Palmieri MR, Costas R. Risk factors of coronary heart disease: a prospective epidemiologic study in Puerto Rico. In: Yu PH, Goodwin JF, eds. *Progress in Cardiology*. Vol 14. Philadelphia, Pa: Lea & Febiger; 1986:101-190.
16. Lee ET, Welty TK, Fabsitz R, et al. The Strong Heart Study: a study of cardiovascular disease in American Indians: design and methods. *Am J Epidemiol*. 1990;136:1141-1155.
17. Fried LP, Borhani NO, Enright P, et al. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol*. 1991;1:1263-1276.
18. Hanley JA, McNeill BJ. The measure and use of the area under the receiver operating characteristic (ROC) curve. *Radiology*. 1982;143:29-36.
19. D'Agostino RB Sr, Griffith JL, Schmid CH, Terrin N. Measures for evaluating model performance. From: Annual Meeting of the American Statistical Association; Chicago, Ill, August 1996. In: *American Statistical Association 1996 Proceedings of the Section on Biometrics*. Alexandria, Va: American Statistical Association; 1997:253-258.
20. Nam B-H. *Discrimination and Calibration in Survival Analysis* [dissertation]. Boston, Mass: Boston University; 2000.
21. Colette D. *Modeling Survival Data in Medical Research*. London, England: Chapman & Hall; 1994.