

Development and Validation of Improved Algorithms for the Assessment of Global Cardiovascular Risk in Women

The Reynolds Risk Score

Paul M Ridker, MD, MPH

Julie E. Buring, ScD

Nader Rifai, PhD

Nancy R. Cook, ScD

IN THE DECADE BETWEEN 1956 AND 1966, investigators in Framingham, Mass, defined age, hypertension, smoking, diabetes, and hyperlipidemia as major determinants of coronary heart disease and coined the term *coronary risk factors*.¹⁻⁵ Over time, these markers were codified into global risk scores for assessment of cardiovascular risk.⁶⁻⁸ However, for women, up to 20% of all coronary events occur in the absence of these major risk factors,⁹ whereas many women with traditional risk factors do not experience coronary events.¹⁰ Furthermore, over the past half-century, understanding of the biological processes underlying atherothrombosis has markedly shifted to encompass the complex biology of hemostasis, thrombosis, inflammation, endothelial dysfunction, and plaque instability.^{11,12}

Despite this changing view of pathophysiology, variables included in current risk algorithms for women are largely unchanged from those recommended 40 years ago. Additional risk markers that have been proposed include alternative lipid measures, such as apolipoproteins A-I and B-100, non-high-density lipoprotein cholesterol (HDL-C), and lipoprotein(a); inflammatory biomarkers such as high-sensitivity C-reactive protein (hsCRP), soluble intercellular adhesion molecule 1 (sICAM-1), and fibrino-

For editorial comment see p 641.

Context Despite improved understanding of atherothrombosis, cardiovascular prediction algorithms for women have largely relied on traditional risk factors.

Objective To develop and validate cardiovascular risk algorithms for women based on a large panel of traditional and novel risk factors.

Design, Setting, and Participants Thirty-five factors were assessed among 24 558 initially healthy US women 45 years or older who were followed up for a median of 10.2 years (through March 2004) for incident cardiovascular events (an adjudicated composite of myocardial infarction, ischemic stroke, coronary revascularization, and cardiovascular death). We used data among a random two thirds (derivation cohort, n=16 400) to develop new risk algorithms that were then tested to compare observed and predicted outcomes in the remaining one third of women (validation cohort, n=8158).

Main Outcome Measure Minimization of the Bayes Information Criterion was used in the derivation cohort to develop the best-fitting parsimonious prediction models. In the validation cohort, we compared predicted vs actual 10-year cardiovascular event rates when the new algorithms were compared with models based on covariates included in the Adult Treatment Panel III risk score.

Results In the derivation cohort, a best-fitting model (model A) and a clinically simplified model (model B, the Reynolds Risk Score) had lower Bayes Information Criterion scores than models based on covariates used in Adult Treatment Panel III. In the validation cohort, all measures of fit, discrimination, and calibration were improved when either model A or B was used. For example, among participants without diabetes with estimated 10-year risks according to the Adult Treatment Panel III of 5% to less than 10% (n=603) or 10% to less than 20% (n=156), model A reclassified 379 (50%) into higher- or lower-risk categories that in each instance more accurately matched actual event rates. Similar effects were achieved for clinically simplified model B limited to age, systolic blood pressure, hemoglobin A_{1c} if diabetic, smoking, total and high-density lipoprotein cholesterol, high-sensitivity C-reactive protein, and parental history of myocardial infarction before age 60 years. Neither new algorithm provided substantive information about women at very low risk based on the published Adult Treatment Panel III score.

Conclusion We developed, validated, and demonstrated highly improved accuracy of 2 clinical algorithms for global cardiovascular risk prediction that reclassified 40% to 50% of women at intermediate risk into higher- or lower-risk categories.

JAMA. 2007;297:611-619

www.jama.com

gen; markers of glycemic control such as glycated hemoglobin A_{1c}; and plasma creatinine and homocysteine levels.¹³ However, data are scant evaluating whether improved risk prediction algorithms can be developed that use these markers.¹⁴⁻¹⁶

We assayed all of these novel biomarkers as well as a large number of tradi-

Author Affiliations: Donald W. Reynolds Center for Cardiovascular Research and the Center for Cardiovascular Disease Prevention (Drs Ridker, Cook, and Buring), Division of Preventive Medicine (Drs Ridker, Buring, and Cook), and the Division of Cardiovascular Diseases (Dr Ridker), Brigham and Women's Hospital, Boston, Mass, and the Department of Laboratory Medicine, Children's Hospital, Boston, Mass (Dr Rifai).

Corresponding Author: Paul M Ridker, MD, MPH, Center for Cardiovascular Disease Prevention, Brigham and Women's Hospital, 900 Commonwealth Ave E, Boston, MA 02215 (pridker@partners.org).

tional risk determinants at baseline in a cohort of 24 558 initially healthy US women who were prospectively followed up for a median 10.2 years for incident myocardial infarction, stroke, coronary revascularization, or cardiovascular death. In a random subset comprising two thirds of these women (model derivation cohort, $n=16\,400$), we developed 2 novel algorithms for global risk prediction. We then tested the effectiveness of these new prediction models in the remaining one third of the women (test validation cohort, $n=8158$).

METHODS

Study Participants, Laboratory Evaluation, and End Point Ascertainment

Study participants were derived from the Women's Health Study (WHS), a nationwide cohort of US women 45 years and older free of cardiovascular disease and cancer at study entry initiated in September 1992.¹⁷ Women eligible for the current analysis were those who provided an adequate baseline plasma sample ($n=27\,939$) and had complete ascertainment of all blood covariates of interest ($n=24\,558$). Exposure data were collected for age, race/ethnicity, diabetes, blood pressure, blood pressure treatment, smoking status, cholesterol treatment, menopausal status, postmenopausal hormone therapy use, height, weight, alcohol use, exercise frequency, parental history of myocardial infarction before age 60 years, and current multivitamin use. All participants self-reported race/ethnicity as white, black, Hispanic American, Asian American, or other. All women were followed up through March 2004 for a median period of 10.2 years (interquartile range, 9.7-10.6 years) for incident myocardial infarction, ischemic stroke, coronary revascularization, and cardiovascular deaths; these were adjudicated by an end-points committee after medical record review. All study participants provided written informed consent. The study protocol was approved by the institutional review board of Brigham and Women's Hospital (Boston, Mass).

All women had baseline plasma samples, 76% of whom had fasting blood

samples. The plasma samples were measured in a core laboratory facility for total cholesterol, HDL-C, low-density lipoprotein cholesterol (LDL-C), lipoprotein (a), apolipoproteins A-I and B-100, hsCRP, sICAM-1, fibrinogen, creatinine, hemoglobin A_{1c}, and plasma homocysteine concentration. The core laboratory is certified by the National Heart, Lung, and Blood Institute/Centers for Disease Control and Prevention Lipid Standardization Program. Assay characteristics and coefficients of variation are available upon request.

Derivation of Novel Risk Prediction Algorithms

Two thirds of the study participants ($n=16\,400$) were randomly assigned to a model derivation data set and one third ($n=8158$) were reserved as an independent validation data set.

Among women allocated to the model derivation set, the best overall prediction algorithm (model A) was fit using Cox proportional hazards models. All available exposure variables and all blood biomarkers were considered for this initial model, as were all potential transformations and interactions between them. Both stepwise selection procedures and multiple additive regression trees¹⁸ were used for variable selection, assessment for interactions, and model development. Partial dependence plots were examined for evidence of interaction, even in the absence of main effects. These interaction terms were then further tested in the Cox models.

The final criterion for inclusion in model A was minimization of the Bayes Information Criterion (BIC).¹⁹ The BIC is a likelihood-based measure in which lower values indicate better fit and in which a penalty is paid for increasing the number of variables. Thus, the variables selected for inclusion should provide not only the best fit but also a parsimonious prediction model. The BIC is not influenced by the number of covariates, so models can be directly compared.

Once variables were selected for model A, we created a second model (model B) that was simplified for the purpose of clinical application and ef-

iciency. For example, in these data non-HDL-C [total cholesterol -HDL-C] is highly correlated with apolipoprotein B-100 ($r=0.87$), and HDL-C is highly correlated with apolipoprotein A-I ($r=0.80$).²⁰ Thus, model B substituted total cholesterol and HDL-C. Simplified model B also eliminated lipoprotein(a) because prior work in this cohort has found the predictive utility of lipoprotein(a) to be limited to those with extremely high values (>90th percentile) and concomitant hyperlipidemia.²¹

To allow for direct comparison, the BIC was calculated using data from the derivation cohort for models A and B, as well as for models based exclusively on covariates used in the current Adult Treatment Panel III (ATP-III) risk prediction algorithm⁷ or in the Framingham Risk Score,⁶ but with coefficients reestimated in the WHS data.

Testing and Validation of Novel Risk Prediction Algorithms

Once determined in final form, models A and B were prospectively tested in the validation data set of 8158 women. In this validation stage, 3 global measures were used to evaluate each prediction model: Entropy (a likelihood-based function for dichotomous outcomes for which smaller values indicate better fit); the Yates slope (the difference in predicted risk between cases and noncases for which larger values indicate better fit); and the Brier score (which computes the sum of squared differences between the observed outcome and fitted probabilities and for which smaller values indicate better concordance between predicted and observed outcomes).^{22,23} Because all women were followed up for at least 8 years, observed status and predicted risk were evaluated and compared as of 8 years of follow-up for all measures.

In addition to these global measures, we assessed the predictive accuracy of each derived model by looking at 2 components of accuracy: discrimination and calibration. Discrimination was evaluated using the *C* statistic that represents the area under the receiver operating characteristic curve (for which larger values indicate bet-

ter discrimination). To assess model calibration (or how closely the predicted probabilities reflect actual risk), the Hosmer-Lemeshow calibration statistic comparing observed and predicted risk was computed based on categories defined by 2% increments in predicted risk.

To compare the performance of models A and B to current risk prediction algorithms, we also computed each of these summary statistics in the test cohort using models limited exclusively to covariates defined in the current ATP-III or Framingham Risk Scores, but with coefficients reestimated in the WHS cohort. We additionally computed each of these summary statistics for predicted outcomes based on formal application of the published ATP-III and Framingham Risk scoring systems as estimated from Framingham data.^{6,7}

Risk Stratification, Reclassification, and Clinical Application

For ease of interpretation and to address the critical clinical issues of reclassification and risk stratification, we divided all participants in the test cohort into the 10-year risk groups of less than 5%, 5% to less than 10%, 10% to less than 20%, and 20% or higher using covariates currently included in the ATP-III risk prediction model. We then calculated the proportion of participants in the test cohort who were reclassified into either higher- or lower-risk categories using models A or B rather than the covariates in the ATP-III model and then compared observed to predicted events during the follow-up period.

Finally, to mimic clinical practice, we repeated these latter analyses using the published ATP-III risk prediction score to determine 10-year risk groups rather than the refitted model using the ATP-III covariates; because diabetes is considered a coronary risk equivalent in current ATP-III guidelines, this final analysis was restricted to nondiabetic study participants.

Analyses were conducted using SAS version 9.1 (SAS Institute Inc, Cary, NC), SPlus version 7.0 (Insightful Corp, Seattle, Wash), and Treenet version 2.0 (Salford Systems, San Diego, Calif).

RESULTS

TABLE 1 shows baseline characteristics and biomarker levels for women in the derivation and validation cohorts.

During follow-up, 504 cardiovascular events occurred in the derivation cohort and 262 in the validation cohort.

Table 1. Baseline Clinical Characteristics and Plasma Biomarker Levels for Women Initially Free of Cardiovascular Disease and Cancer in the Model Derivation Cohort and the Model Testing and Validation Cohort*

	Derivation Cohort (n = 16 400)	Validation Cohort (n = 8158)
Age, median (IQR), y	52 (48-58)	52 (49-59)
Race, No. (%)		
White	15 500 (95.2)	7710 (95.3)
Black	310 (1.9)	151 (1.9)
Hispanic	169 (1.0)	82 (1.0)
Asian	220 (1.4)	123 (1.5)
Other	77 (0.5)	23 (0.3)
Smoking status, No. (%)		
Current	1895 (11.6)	927 (11.4)
Past	5961 (36.4)	3007 (36.9)
Never	8544 (52.1)	4424 (51.8)
Height, median (IQR), in	65 (63-66)	65 (63-66)
Weight, median (IQR), lb	148 (132-170)	148 (132-170)
Body mass index, median (IQR)†	24.9 (22.5-28.3)	24.8 (22.5-28.3)
Alcohol use, >once/wk, No. (%)	6890 (42.0)	3571 (43.8)
Exercise, >once/wk, No. (%)	7110 (43.4)	3492 (42.8)
Blood pressure, median (IQR), mm Hg		
Systolic	125 (115-135)	125 (115-135)
Diastolic	80 (70-80)	80 (70-80)
Risk factors, No. (%)		
Diabetes	442 (2.7)	238 (2.9)
History of hypertension	4061 (24.8)	2061 (25.3)
Parental history of MI	2112 (12.9)	1039 (12.7)
Menopausal	8911 (54.4)	4423 (54.3)
Medication use, No. (%)		
Hormone therapy	7233 (44.2)	3523 (43.3)
Lipid-lowering therapy	531 (3.2)	257 (3.2)
Current multivitamin use	4805 (29.7)	2321 (28.8)
Cholesterol, median (IQR), mg/dL		
Total	208 (183-235)	208 (184-235)
LDL-C	121.0 (100.1-144.1)	121.3 (100.9-143.8)
HDL-C	51.9 (43.1-62.5)	52.2 (43.4-62.5)
Non-HDL-C	153.9 (128.7-181.6)	153.8 (129.4-181.2)
Apolipoprotein A-I, median (IQR), mg/dL	148.9 (132.5-167.7)	149.6 (132.7-168.6)
Apolipoprotein B-100, median (IQR), mg/dL	99.7 (83.5-120.8)	100.1 (84.2-120.8)
Lipoprotein(a), median (IQR), mg/dL	10.5 (4.4-32.0)	10.7 (4.3-32.6)
hsCRP, median (IQR), mg/L	2.0 (0.8-4.3)	2.0 (0.8-4.4)
Fibrinogen, median (IQR), mg/dL	349.8 (306.7-402.6)	351.7 (308.0-402.7)
sICAM-1, median (IQR), ng/mL	343.1 (301.1-394.2)	341.5 (300.8-394.8)
Homocysteine, median (IQR), μ mol/L	10.4 (8.7-12.8)	10.5 (8.8-13.0)
Creatinine, median (IQR), mg/dL	0.71 (0.63-0.80)	0.71 (0.63-0.80)
HbA _{1c}	5.0 (4.8-5.2)	5.0 (4.8-5.2)

Abbreviations: HbA_{1c}, hemoglobin A_{1c}; HDL-C, high-density lipoprotein cholesterol; hsCRP, high-sensitivity C-reactive protein; IQR, interquartile range; LDL-C, low-density lipoprotein cholesterol; MI, myocardial infarction; sICAM-1, soluble intercellular adhesion molecule 1.

SI conversion factors: to convert creatinine to μ mol/L, multiply by 88.4; fibrinogen to μ mol/L, multiply by 0.0294; homocysteine to mg/dL, divide by 7.397; inches to centimeters, multiply by 2.54; pounds to kilograms, multiply by 0.45; lipoprotein(a) to μ mol/L, multiply by 0.0357; and total cholesterol, HDL-C, and LDL-C to mmol/L, multiply by 0.0259.

*Percentages may not sum to 100 due to rounding and some numbers may not add to the total due to missing information. †Body mass index is calculated as weight in kilograms divided by height in meters squared.

Table 2. Best-Fitting Model A and Clinically Simplified Model B for Global Cardiovascular Risk Prediction Based on Data From the Model Derivation Cohort (n = 16 400)

	Best-Fitting Model A, β (SE)	χ^2	P Value
Age	0.078 (0.006)	186.6	<.001
HbA _{1c} , % with diabetes	0.134 (0.017)	62.9	<.001
Natural logarithm			
Systolic blood pressure	3.271 (0.420)	60.6	<.001
Current smoking	0.825 (0.109)	57.0	<.001
[Lp(a) - 10] ₊ if Apo-B-100 \geq 100*	0.0074 (0.0013)	34.8	<.001
Apolipoprotein B-100	0.0082 (0.0016)	25.9	<.001
Natural logarithm			
hsCRP	0.202 (0.042)	22.7	<.001
Apolipoprotein A-I	-0.0077 (0.0018)	17.5	<.001
Parental history of MI <age 60 y	0.427 (0.118)	13.0	<.001
Simplified Model B, β (SE)			
Age	0.080 (0.006)	193.5	<.001
HbA _{1c} , % with diabetes	0.134 (0.017)	62.3	<.001
Current smoking	0.818 (0.109)	55.9	<.001
Natural logarithm			
Systolic blood pressure	3.137 (0.423)	55.1	<.001
HDL-C	-1.172 (0.172)	46.2	<.001
Total cholesterol	1.382 (0.239)	33.3	<.001
hsCRP	0.180 (0.043)	17.5	<.001
Parental history of MI <age 60 y	0.438 (0.118)	13.7	<.001

Abbreviations: Apo, apolipoprotein; HbA_{1c}, hemoglobin A_{1c}; HDL-C, high-density lipoprotein cholesterol; hsCRP, high-sensitivity C-reactive protein; Lp(a), lipoprotein(a); MI, myocardial infarction.

*[Lp(a) - 10]₊ = Lp(a) - 10 if Lp(a) is greater than 10; otherwise = 0.

Model Derivation and Development

In the model derivation cohort, 35 potential variables (and all possible interactions between them) were evaluated for model inclusion. Of these, only 9 were included in model A, the best-fitting predictive model with the smallest BIC value; age, systolic blood pressure, current smoking, apolipoprotein B-100, hsCRP, apolipoprotein A-I, parental history of myocardial infarction before age 60 years, and 2 interaction terms, hemoglobin A_{1c} if diabetes was present and lipoprotein(a) level if apolipoprotein B-100 was 100 mg/dL or higher. The β coefficients, standard errors, and P values for each of these covariates in best-fitting model A are shown in TABLE 2.

Given selection of these 9 variables, some markers, such as homocysteine and sICAM-1, appeared to predict risk, but did not satisfy the BIC criterion for model inclusion. Other notable variables that did not further minimize the BIC once the above variables were taken

into account included body mass index, alcohol use, exercise frequency, menopausal status, hormone therapy, fibrinogen, and creatinine.

Table 2 also presents β coefficients, standard errors, and P values for simplified model B, which was otherwise identical to model A, but substituted total and HDL-C for apolipoproteins B100 and A-I, and eliminated the interaction term requiring measurement of lipoprotein(a) if apolipoprotein B-100 was 100 mg/dL or higher.

In the derivation data set, the BIC value for model B (BIC=9067.5) was not as small as that of the best-fitting model A (BIC=9039.4), suggesting some loss of predictive ability with clinical simplification. However, model B nevertheless was associated with smaller BIC values than were models based on covariates used in the ATP-III prediction model (BIC=9098.5) or those based on covariates used in the Framingham Risk Score (BIC=9161.2). Thus, in the model derivation set, both model A and model B appeared to improve risk prediction

over that achieved with currently measured covariates BOX.

Model Testing and Validation

TABLE 3 presents summary statistics regarding the performance of models A and B in terms of predicting risk among the 8158 women reserved in the prospective validation data set. For each prespecified global summary statistic (Entropy, Yates Slope, Brier Score, and C statistic), models A and B provided improvement over prediction models based on covariates used in the ATP-III or Framingham models or when the published ATP-III or Framingham Scores were directly applied. With regard to comparisons of predicted and observed risk, P values for the Hosmer-Lemeshow statistics for model A and B indicated good calibration. Although calibration was suboptimal for the 3 published score models, part of this effect was due to a difference in end-point definition.

Reclassification and Clinical Application

Although formal statistical testing provides a method of evaluating model superiority, we believe the critical issue for clinical application is the proportion of patients reclassified using a new risk algorithm and whether the magnitude of this reclassification is large enough to alter physician behavior with regard to prevention.²⁴

To address this issue, TABLE 4 presents the proportion of women in the validation cohort initially classified as having a 10-year risk of less than 5%, 5% to less than 10%, 10% to less than 20%, and 20% or higher based on ATP-III covariates (with coefficients reestimated in the WHS data) who would be reclassified to higher- or lower-risk categories by model A and model B. As shown for model A, the proportion of women reclassified was small for those with a 10-year risk of less than 5% (2.5%). However, 43% of all women estimated to be at 5% to less than 10% risk or at 10% to less than 20% risk using ATP-III covariates were reclassified to higher or lower clinical risk categories when model A was used instead. Table 4 also shows that actual event rates for model A matched well

with predicted rates in nearly all groups; of the 681 participants reclassified by model A, all but 93 were placed into more accurate risk categories.

TABLE 5 presents similar analyses for women who did not have diabetes with direct application of the published ATP-III risk score. As shown, about 50% of all women with an estimated 10-year risk for coronary heart disease of 5% to less than 10% or 10% to less than 20% according to ATP-III were reclassified to higher or lower risk categories when model A was used instead. Again, there was excellent matching of actual and predicted rates for model A; of the 722 participants without diabetes who were reclassified by model A, all but 2 were placed into more accurate risk categories.

As also presented in Table 4 and Table 5, similar effects were achieved for clinically simplified model B limited to age, systolic blood pressure, hemoglobin A_{1c} if diabetic, current smoking, total and HDL-C, hsCRP, and parental history of myocardial infarction before age 60 years. Although the proportion of individuals at intermediate-risk reclassified by model B (30%-45%) was smaller than that of model A (43%-50%), there was still excellent matching of actual to predicted event rates in nearly all groups. For example, of the 647 participants without diabetes in Table 5 who were reclassified by model B, all but 6 were placed into more accurate risk categories. Neither new algorithm added substantive information for women at very low initial risk (<5% 10-year risk based on published ATP-III risk scores).

Examples for Outpatient Clinical Practice: The Reynolds Risk Score

As a practical example, TABLE 6 provides estimated 10-year risks based on variables in our most parsimonious model (model B, the Reynolds Risk Score) for a 50-year-old women smoker without diabetes with an ATP-III estimated risk of 11.5%. As shown, 10-year risk estimates based on model B range from a low of 4.9% to a high of 18.4% for this hypothetical patient.

With regard to reclassification, as shown in the FIGURE for a representa-

Box. Computational Formulas for 10-Year Risk Using Best-Fitting Model A and Clinically Simplified Model B

Model A

10-year cardiovascular disease risk (%) = $[1 - 0.98756^{(\exp [A - 19.848])}] \times 100\%$ where
 $A = 0.0785 \times \text{age} + 3.271 \times \text{natural logarithm (systolic blood pressure)} + 0.202 \times \text{natural logarithm (high-sensitivity C-reactive protein)} + 0.00820 \times \text{apolipoprotein B-100} - 0.00769 \times \text{apolipoprotein A-1} + 0.134 \times \text{hemoglobin A}_{1c} (\%) \text{ (if diabetic)} + 0.825 \text{ (if current smoker)} + 0.427 \text{ (if family history of premature myocardial infarction)} + 0.00742 \times (\text{lipoprotein(a)-10}) \text{ (if lipoprotein(a)} > 10 \text{ and apolipoprotein B-100} \geq 100)$

Model B, the Reynolds Risk Score

10-year cardiovascular disease risk (%) = $[1 - 0.98634^{(\exp [B - 22.325])}] \times 100\%$ where
 $B = 0.0799 \times \text{age} + 3.137 \times \text{natural logarithm (systolic blood pressure)} + 0.180 \times \text{natural logarithm (high-sensitivity C-reactive protein)} + 1.382 \times \text{natural logarithm (total cholesterol)} - 1.172 \times \text{natural logarithm (high-density lipoprotein cholesterol)} + 0.134 \times \text{hemoglobin A}_{1c} (\%) \text{ (if diabetic)} + 0.818 \text{ (if current smoker)} + 0.438 \text{ (if family history of premature myocardial infarction)}$

tive population of 100 000 US women without diabetes at intermediate risk (80 000 at 5% to less than 10% and 20 000 at 10% to less than 20% 10-year risk by ATP-III), use of the clinically simplified Reynolds Risk Score would place 13 500 of these women at low risk, 48 500 at low to moderate risk, 32 500 at moderate to high risk, and 5400 at high risk.

COMMENT

In this study of 24 558 initially healthy US women followed up for a median of 10.2 years, we developed and validated risk prediction algorithms that reclassified 40% to 50% of women currently predicted to be at intermediate risk into higher- or lower-risk categories and did so with greatly improved accuracy when compared with models based on current ATP-III prediction scores. This effect was present not only for our best-fitting model (model A) but also for a simplified clinical model limited to age, systolic blood pressure, hemoglobin A_{1c} if diabetic, current smoking, total and HDL-C, hsCRP, and parental history of myocardial infarction before age 60 years (model B, the Reynolds Risk Score).

In addition to providing opportunity for improved risk stratification, we believe these data have clinical impli-

cations for the targeting of preventive therapies. In these analyses, large proportions of women with 10-year risk estimates of 5% to less than 10% or of 10% to less than 20% based on current ATP-III risk scores were reclassified at either higher or lower risk of total cardiovascular disease when either of the new algorithms was used. In current US treatment guidelines that take into account the benefits, risk, and cost of lipid-lowering therapy, statins are considered an option for those with 10-year risk estimates of 10% or greater²⁵; a more conservative approach taken in Europe typically limits statin therapy to those with 10-year risks of 20% or more.⁸ In both settings, application of the models described herein should allow more accurate targeting of statin prescriptions to those patients with the most appropriate level of risk so as to minimize toxicity and maximize benefit and cost efficacy.

We also believe these data provide optimism regarding novel cardiovascular risk factors. In our best-fitting model, hemoglobin A_{1c}, hsCRP, lipoprotein(a), apolipoproteins A-I and B-100, and parental history were included because each contributed to minimization of the BIC. However, homocysteine, fibrinogen, sICAM-1, and creatinine were not included in our par-

simonious models despite univariate risk associations. Similarly, neither body mass index nor exercise frequency added further prognostic information on overall global risk.^{26,27} By contrast, we observed that glucose control as evaluated by hemoglobin A_{1c} was an effective biomarker in these women that

modified the risk associated with diabetes.

Our findings might appear to conflict with a recent report from the Framingham Heart Study in which only marginal utility for novel risk factors was described.¹⁶ However, instead of seeking evidence of reclassification, that analysis re-

lied solely on the C statistic, a technique known to have limited utility for evaluating prediction models for which the task is to assess future risk in a currently healthy population.²⁸ Equally important, that analysis relied on data from 1712 women who experienced only 68 vascular events, many of which were coded as heart

Table 3. Summary Statistics Comparing 2 Novel Risk Prediction Algorithms to Prediction Based on Covariates in the ATP-III and Framingham Scores, and to Direct Application of These Latter Global Risk Algorithms, Based on Data From the Validation Cohort (n = 8158)*

	Best-Fitting Model A	Simplified Model B	ATP-III Covariates†	Framingham Covariates		ATP III Model, 2001	Wilson Framingham	
				Total Cholesterol	LDL-C		Total Cholesterol, 1998	LDL-C, 1998
Global measures								
Entropy	779.8	778.0	784.2	793.0	791.4	823.3	936.9	919.6
Yates slope, %	5.74	5.49	5.13	5.13	5.13	2.58	4.75	4.65
Brier score	0.02246	0.02243	0.02249	0.02254	0.02253	0.02308	0.02418	0.02396
Discrimination								
C statistic	0.809	0.808	0.805	0.791	0.791	0.787	0.752	0.751
Calibration								
Hosmer-Lemeshow P value‡	.38	.62	.45	.18	.16	<.001	<.001	<.001

Abbreviations: ATP, Adult Treatment Panel; LDL-C, low-density lipoprotein cholesterol.

* Lower values of entropy and Brier score and higher values of Yates Slope and C statistic indicate better fit.

†Including history of diabetes.

‡A significant value for the Hosmer-Lemeshow statistic indicates a significant deviation between predicted and observed outcomes.

Table 4. Cardiovascular Risk Reclassification in Validation Cohort Comparing Models A and B to Models Based on Current Adult Treatment Panel III Covariates*

ATP-III 10-Year Risk Categories, %	10-Year Risk Best-Fitting Model A†				Total	No. (%) Reclassified
	<5%	5% to <10%	10% to <20%	≥20%		
<5%						
No. (%) of participants	6778 (97.5)	168 (2.4)	8 (0.1)	0	6954	176 (2.5)
Actual event rate	1.4	9.8	0	0		
5% to <10%						
No. (%) of participants	232 (29.1)	455 (57.0)	103 (12.9)	8 (1.0)	798	343 (43.0)
Actual event rate	4.3	6.2	13.6	30.2		
10% to <20%						
No. (%) of participants	3 (1.0)	85 (27.2)	178 (56.9)	47 (15.0)	313	135 (43.1)
Actual event rate	0	16.7	17.7	22.2		
≥20%						
No. (%) of participants	0	0	27 (32.1)	57 (67.9)	84	27 (32.1)
Actual event rate	0	0	18.2	42.0		
10-Year Risk Simplified Model B†						
<5%						
No. (%) of participants	6837 (98.3)	117 (1.7)	0	0	6954	117 (1.7)
Actual event rate	1.4	10.7	0	0		
5% to <10%						
No. (%) of participants	158 (19.8)	559 (70.0)	81 (10.2)	0	798	239 (30.0)
Actual event rate	4.9	6.4	14.0	0		
10% to <20%						
No. (%) of participants	0	54 (17.3)	221 (70.6)	38 (12.1)	313	92 (29.4)
Actual event rate	0	20.8	15.4	28.5		
≥20%						
No. (%) of participants	0	0	21 (25.0)	63 (75.0)	84	21 (25.0)
Actual event rate	0	0	18.4	39.1		

Abbreviation: ATP, Adult Treatment Panel.

*All estimated and observed risks have been extrapolated to 10-year rates (number of events per 100 people per 10 years of observation). Nine missing values are for treatment of hypertension, a variable in the Adult Treatment Panel III model.

†Percentages may not sum to 100 due to rounding.

Table 5. Cardiovascular Risk Reclassification in the Validation Cohort of Women Without Diabetes Comparing Models A and B to Models Based on the Published Adult Treatment Panel III Algorithm*

ATP-III 10-Year Risk Categories	10-Year Risk Best-Fitting Model A†				Total	No. (%) Reclassified
	<5%	5% to <10%	10% to <20%	≥20%		
<5%						
No. (%) of participants	6803 (95.2)	314 (4.4)	25 (0.4)	2 (0)	7144	341 (4.8)
Actual event rate	1.4	7.0	14.8	0		
5% to <10%						
No. (%) of participants	133 (22.1)	303 (50.2)	151 (25.0)	16 (2.7)	603	300 (49.8)
Actual event rate	2.9	8.4	13.8	32.1		
10% to <20%						
No. (%) of participants	8 (5.1)	36 (23.1)	77 (49.4)	35 (22.4)	156	79 (50.6)
Actual event rate	0	3.7	12.2	32.1		
≥20%						
No. (%) of participants	0	0	2 (25.0)	6 (75.0)	8	2 (25.0)
Actual event rate	0	0	0	39.8		
	10-Year Risk Simplified Model B†					
<5%						
No. (%) of participants	6836 (95.7)	297 (4.2)	11 (0.1)	0	7144	308 (4.3)
Actual event rate	1.4	6.9	22.2	0		
5% to <10%						
No. (%) of participants	96 (15.9)	336 (55.7)	162 (26.9)	9 (1.5)	603	267 (44.3)
Actual event rate	4.0	8.3	13.0	30.2		
10% to <20%						
No. (%) of participants	6 (3.8)	31 (19.9)	86 (55.1)	33 (21.2)	156	70 (44.9)
Actual event rate	24.3	4.1	10.7	31.3		
≥20%						
No. (%) of participants	0	0	2 (25.0)	6 (75)	8	2 (25.0)
Actual event rate	0	0	0	39.8		

Abbreviation: ATP, Adult Treatment Panel.

*All estimated and observed risks have been extrapolated to 10-year rates (number of events per 100 people per 10 years of observation).

†Percentages may not sum to 100 due to rounding.

failure or coronary insufficiency. By contrast, the risk algorithms described herein rely on data from 24 558 women who experienced 766 hard cardiovascular end points. We also note that in a separate Framingham Heart Study analysis addressing the additive value of hsCRP, use of this biomarker alone reclassified 25% of those with ATP-III risks between 5% and 20%, data fully consistent with those presented herein.²⁹

Despite advantages of sample size and power, limitations of our analysis merit discussion. First, because our data are limited to women and our cohort is largely white with a relatively narrow socioeconomic range, care should be taken before generalizing to other populations. We note, however, that all components of models A and B have previously been found to predict cardiovascular risk in men³⁰⁻³⁴ and that both hsCRP and parental history of vascular disease have previously been shown to predict risk within the Framingham cohort itself.^{29,35,36}

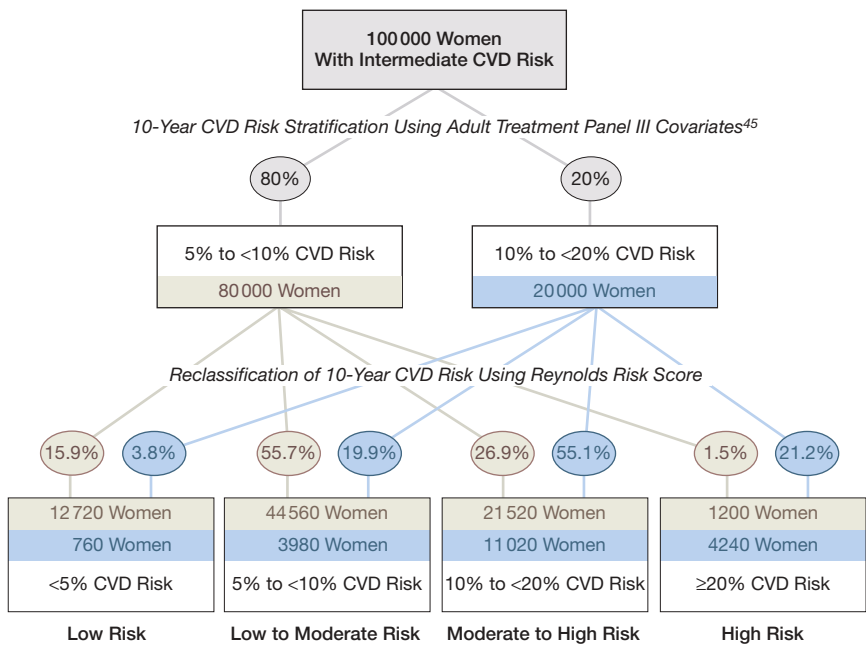
Table 6. Clinical Example: Estimated 10-Year Risk for a 50-Year-Old Nonsmoking Woman Without Diabetes, According to ATP-III or to Clinically Simplified Model B (the Reynolds Risk Score)

Clinical Variables	Estimated 10-Year Risk, %	
	ATP-III Model	Simplified Model B
Blood Pressure, mm Hg		
Cholesterol, mg/dL		
Total		
HDL		
non-HDL		
hsCRP, mg/L		
Parental History*		
155/85	11.5	4.9
155/85	11.5	6.5
155/85	11.5	7.4
155/85	11.5	8.9
155/85	11.5	9.7
155/85	11.5	10.5
155/85	11.5	10.9
155/85	11.5	12.3
155/85	11.5	7.5
155/85	11.5	9.9
155/85	11.5	11.2
155/85	11.5	13.4
155/85	11.5	14.6
155/85	11.5	15.8
155/85	11.5	16.4
155/85	11.5	18.4

Abbreviations: ATP, Adult Treatment Panel; HDL, high-density lipoprotein; hsCRP, high-sensitivity C-reactive protein. SI conversion factor: To convert cholesterol from mg/dL to mmol/L, multiply by 0.02589.

*Parental myocardial infarction event before age 60 years.

Figure. Reclassification of Risk Using the Reynolds Risk Score for a Representative Population of 100 000 Intermediate-Risk US Women Without Diabetes



Percentages shown in ovals indicate the proportion of women distributed to risk categories based on Adult Treatment Panel III (top) and the Reynolds Risk Score (bottom). Reclassification using the Reynolds Risk Score is based on data shown in Table 5, Model B. CVD indicates cardiovascular disease.

Second, our data on blood pressure, obesity, and family history were based on self-report. However, the WHS is composed of female health professionals who are known to provide accurate reports of lifestyle factors and health status, including blood pressure and weight.^{37,38} In addition, self-reported blood pressure, body mass index, and family history have previously been shown in the WHS to be strong predictors of cardiovascular risk, with odds ratios consistent in magnitude with those observed in other major studies.³⁹⁻⁴¹ Regarding parental history, we used a conservative cut point of age younger than 60 years to be consistent with prior findings in this cohort and in recent analyses from Framingham.^{41,42} The inclusion of family history in these algorithms underscores the importance of genetic influences on risk among women; in a recent study of women with low Framingham risk who had premature coronary disease in a first-degree relative, nearly a third had significant subclinical atherosclerosis and 17% had ath-

erosclerotic burden exceeding the 90th percentile.⁴³

Third, following recent recommendations,⁴⁴ we elected in our analysis to use a combined end point of myocardial infarction, ischemic stroke, coronary revascularization, and cardiovascular mortality. We believe this is an appropriate choice because this end point has typically been used in major cardiovascular clinical trials evaluating interventions for primary prevention, including recent trials of aspirin and statin therapy.

Finally, we limited our analysis to blood-based biomarkers and traditional epidemiological risk factors, in part to ensure a cost-effective approach for primary prevention that could be directly compared with the ATP-III algorithm. These data thus do not examine the potential for atherosclerotic imaging tests to serve as an alternative method for evaluating risk. However, we believe the methods developed herein—variable selection in a derivation data set to minimize the BIC followed by prospective testing in a second validation cohort—

should provide a structure for the formal evaluation of emerging risk predictors, including potential imaging tests.

As 8 to 10 million US women have an ATP-III estimated 10-year risk between 5% and 20%, application of these data could have an immediate effect on cardiovascular prevention.⁴⁵ A user-friendly calculator for the Reynolds Risk Score can be freely accessed at <http://www.reynoldsriskscore.org>.

Author Contributions: Drs Ridker and Cook had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Ridker, Cook.

Acquisition of data: Ridker, Buring, Rifai, Cook.

Analysis and interpretation of data: Ridker, Buring, Rifai, Cook.

Drafting of the manuscript: Ridker, Cook.

Critical revision of the manuscript for important intellectual content: Ridker, Buring, Rifai, Cook.

Statistical analysis: Cook.

Obtained funding: Ridker.

Administrative, technical, or material support: Rifai.

Financial Disclosures: Dr Ridker reports that he currently or in the past 5 years has received research funding support from multiple not-for-profit entities including the National Heart, Lung, and Blood Institute, the National Cancer Institute, the American Heart Association, the Doris Duke Charitable Foundation, the Leducq Foundation, the Donald W. Reynolds Foundation, and the James and Polly Annenberg La Vea Charitable Trusts. Dr Ridker also reports that currently or in the past 5 years he has received investigator-initiated research support from multiple for-profit entities including AstraZeneca, Bayer, Bristol-Myers Squibb, Dade-Behring, Novartis, Pharmacia, Roche, Sanofi-Aventis, and Variagenics. Dr Ridker reports being listed as a coinventor on patents held by the Brigham and Women's Hospital that relate to the use of inflammatory biomarkers in cardiovascular disease and has served as a consultant to Schering-Plough, Sanofi-Aventis, AstraZeneca, Isis Pharmaceutical, Dade-Behring, and Vascular-Biogenics. Dr Buring reports that she currently or in the past 5 years has received investigator-initiated research funding and support from the National Heart, Lung, and Blood Institute, the National Cancer Institute, the National Institute on Aging, and Dow Corning Corp; research support for pills and/or packaging from Bayer Health Care and the Natural Source Vitamin E Association; and honoraria from Bayer for speaking engagements. Dr Rifai reports receiving research grant support from Merck Research Laboratories, serving as a consultant to Merck Research Laboratories and Sanofi/Aventis, and receiving honoraria for speaking from Merck Research Laboratories, Dade Behring, Abbott Laboratories, Ortho Diagnostics, Denka Seiken, and Roche Diagnostics. Dr Cook reports having received funding from the National Heart, Lung, and Blood Institute, the National Cancer Institute, and Roche Diagnostics, and has served as a consultant to Bayer Health Care.

Funding/Support: The Reynolds Risk Score Project was supported by investigator-initiated research grants from the Donald W. Reynolds Foundation (Las Vegas, Nev) with additional support from the Doris Duke Charitable Foundation (New York, NY), and the Leducq Foundation (Paris, France). The Women's Health Study cohort is supported by grants from the National Heart, Lung, and Blood Institute and the National Cancer Institute (Bethesda, Md).

Role of the Sponsor: The funding agencies had no involvement in the design and conduct of the study, the collection, management, analysis, and interpretation of the data, or in the drafting of the manuscript.

REFERENCES

- Dawber TR, Moore FE Jr, Mann GV. Coronary heart disease in the Framingham Study. *Am J Public Health.* 1957;47:4-24.
- Dawber TR, Kannel WB, Revotskie N, Stokes J, Kagan A, Gordon T. Some factors associated with the development of coronary heart disease: six years' follow-up experience in the Framingham study. *Am J Public Health.* 1959;49:1349-1356.
- Kannel WB, Dawber TR, Kagan A, Revotskie N, Stokes J. Factors of risk in the development of coronary heart disease—six years follow-up experience: the Framingham Study. *Ann Intern Med.* 1961;55:33-50.
- Doyle JT, Dawber TR, Kannel WB, Kinch SH, Kahn HA. The relationship of cigarette smoking to coronary heart disease. *JAMA.* 1964;190:886-890.
- Dawber TR, Kannel WB. The Framingham study: an epidemiological approach to coronary heart disease. *Circulation.* 1966;34:553-555.
- Wilson PW, 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.
- 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.
- Wood D, DeBacker G, Faergeman O, Graham I, Mancía G, Pyörälä K. Prevention of coronary heart disease in clinical practice: recommendations of the Second Joint Task Force of European and other Societies on Coronary Prevention. *Atherosclerosis.* 1998;140:199-270.
- Khot UN, Khot MB, Bajzer CT, et al. Prevalence of conventional risk factors in patients with coronary heart disease. *JAMA.* 2003;290:898-904.
- Greenland P, Knoll MD, Stamler J, et al. Major risk factors as antecedents of fatal and nonfatal coronary heart disease events. *JAMA.* 2003;290:891-897.
- Hansson GK. Inflammation, atherosclerosis, and coronary artery disease. *N Engl J Med.* 2005;352:1685-1695.
- Libby P, Theroux P. Pathophysiology of coronary artery disease. *Circulation.* 2005;111:3481-3488.
- Ridker PM, Brown NJ, Vaughan DE, Harrison DG, Mehta JL. Established and emerging plasma biomarkers in the prediction of first atherothrombotic events. *Circulation.* 2004;109(25 suppl 1):IV6-IV19.
- Folsom AR, Chambless LE, Ballantyne CM, et al. An assessment of incremental coronary risk prediction using C-reactive protein and other novel risk markers: the Atherosclerosis Risk in Communities Study. *Arch Intern Med.* 2006;166:1368-1373.
- Voss R, Cullen P, Schulte H, Assman G. Prediction of risk of coronary events in middle-aged men in the Prospective Cardiovascular Munster Study (PRO-CAM) using neural networks. *Int J Epidemiol.* 2002;31:1253-1262.
- Wang TJ, Gona P, Larson MG, et al. Multiple biomarkers for the prediction of first major cardiovascular events and death. *N Engl J Med.* 2006;355:2631-2639.
- Ridker PM, Cook NR, Lee IM, et al. A randomized trial of low-dose aspirin in the primary prevention of cardiovascular disease in women. *N Engl J Med.* 2005;352:1293-1304.
- Friedman JH, Meulman JJ. Multiple additive regression trees with application in epidemiology. *Stat Med.* 2003;22:1365-1381.
- Harrell FE Jr. *Regression Modeling Strategies.* New York, NY: Springer-Verlag; 2001.
- Ridker PM, Rifai N, Cook NR, Bradwin G, Buring JE. Non-HDL cholesterol, apolipoproteins A-I and B100, standard lipid measures, lipid ratios, and CRP as risk factors for cardiovascular disease in women. *JAMA.* 2005;294:326-333.
- Suk Danik J, Rifai N, Buring JE, Ridker PM. Lipoprotein(a), measured with an assay independent of apolipoprotein(a) isoform size, and risk of future cardiovascular events among initially healthy women. *JAMA.* 2006;296:1363-1370.
- Gail MH, Pfeiffer RM. On criteria for evaluating models of absolute risk. *Biostatistics.* 2005;6:227-239.
- D'Agostino RB, Griffith JL, Schmidt CH, Terrin N. *Measures for Evaluating Model Performance.* Alexandria, Va: Proceedings of the Biometrics Section, American Statistical Assoc; 1997:253-258.
- Cook NR, Buring JE, Ridker PM. The effect of including C-reactive protein in cardiovascular prediction models for women. *Ann Intern Med.* 2006;145:21-29.
- Grundty SM, Cleeman JJ, Merz CN, et al; Coordinating Committee of the National Cholesterol Education Program. Implications of recent clinical trials for the National Cholesterol Education Program Adult Treatment Panel III guidelines. *Circulation.* 2004;110:227-239.
- Lee IM, Rexrode KM, Cook NR, Manson JE, Buring JE. Physical activity and coronary heart disease in women: is "No pain, no gain" passe? *JAMA.* 2001;285:1447-1454.
- Mora S, Lee IM, Buring JE, Ridker PM. Association of physical activity and body mass index with novel and traditional cardiovascular biomarkers in women. *JAMA.* 2006;295:1412-1419.
- Cook N. Use and misuse of the ROC curve in risk prediction. *Circulation.* In press.
- Wilson PW, D'Agostino RB, Sullivan L, O'Donnell CJ. Increased CRP and long term risk for cardiovascular events in middle age men and women [abstract 4070]. *Circulation.* 2006;114(suppl II):II-877.
- Danesh J, Collins R, Peto R. Lipoprotein(a) and coronary heart disease: meta-analysis of prospective studies. *Circulation.* 2000;102:1082-1085.
- Ridker PM, Cushman M, Stampfer MJ, Tracy RP, Hennekens CH. Inflammation, aspirin, and the risk of cardiovascular disease in apparently healthy men. *N Engl J Med.* 1997;336:973-979.
- Koenig W, Lowel H, Baumert J, Meisinger C. C-reactive protein modulates risk prediction based on the Framingham Score: implications for future risk assessment: results from a large cohort study in southern Germany. *Circulation.* 2004;109:1349-1353.
- Lamarche B, Moorjani S, Lupien PJ, et al. Apolipoprotein A-I and B levels and the risk of ischemic heart disease during a five-year follow-up of men in the Quebec Cardiovascular Study. *Circulation.* 1996;94:273-278.
- Pischon T, Girman CJ, Sacks FM, Rifai N, Stampfer MJ, Rimm EB. Non-high-density lipoprotein cholesterol and apolipoprotein B in the prediction of coronary heart disease in men. *Circulation.* 2005;112:3375-3383.
- Lloyd-Jones DM, Nam BH, D'Agostino RB, et al. Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: a prospective study of parents and offspring. *JAMA.* 2004;291:2204-2211.
- Rost NS, Wolf PA, Kase CS, et al. Plasma concentration of C-reactive protein and risk of ischemic stroke and transient ischemic attack: the Framingham Study. *Stroke.* 2001;32:2575-2579.
- Ascherio A, Rimm EB, Giovannucci EL, et al. A prospective study of nutritional factors and hypertension among US men. *Circulation.* 1992;86:1475-1484.
- Willett W, Stampfer MJ, Bain C, et al. Cigarette smoking, relative weight, and menopause. *Am J Epidemiol.* 1983;117:651-658.
- Glynn RJ, L'Italiani GJ, Sesso HD, Jackson EA, Buring JE. Development of prediction models for long-term cardiovascular risk associated with systolic and diastolic blood pressure. *Hypertension.* 2002;39:105-110.
- Kurth T, Gaziano JM, Rexrode KM, et al. Prospective study of body mass index and risk of stroke in apparently healthy women. *Circulation.* 2005;111:1992-1998.
- Sesso HD, Lee IM, Gaziano JM, Rexrode KM, Glynn RJ, Buring JE. Maternal and paternal history of myocardial infarction and risk of cardiovascular disease in men and women. *Circulation.* 2001;104:393-398.
- Wang TJ, Nam BH, D'Agostino RB, et al. Carotid intima-media thickness is associated with premature parental coronary heart disease: the Framingham Heart Study. *Circulation.* 2003;108:572-577.
- Michos ED, Vasamreddy CR, Becker DM, et al. Women with a low Framingham risk score and a family history of premature coronary heart disease have a high prevalence of subclinical coronary atherosclerosis. *Am Heart J.* 2005;150:1276-1281.
- Grundty SM. The changing face of cardiovascular risk. *J Am Coll Cardiol.* 2005;46:173-175.
- Ford ES, Giles WH, Mokdad AH. The distribution of 10-year risk for coronary heart disease among US adults: findings from the National Health and Nutrition Examination Survey III. *J Am Coll Cardiol.* 2004;43:1791-1796.

Author in the Room Teleconference

Join Dr Ridker, an author of this article, on Wednesday, March 21, 2007, from 2 to 3 PM eastern time for "Author in the Room."

To register for Author in the Room, please visit <http://www.ihl.org/authorintheroom>. You can listen to past conferences or subscribe to the podcast at <http://jama.ama-assn.org/authorintheroom/authorindex.dtl>.