15 January 1993 Volume 118 Number 2

Annals of Internal Medicine

Value of the History and Physical in Identifying Patients at Increased Risk for Coronary Artery Disease

David B. Pryor, MD; Linda Shaw, AB; Charles B. McCants, BS; Kerry L. Lee, PhD; Daniel B. Mark, MD, MPH; Frank E. Harrell, Jr., PhD; Lawrence H. Muhlbaier, PhD; and Robert M. Califf, MD

■ *Objective:* To determine whether information from the physician's initial evaluation of patients with suspected coronary artery disease predicts coronary anatomy at catheterization and 3-year survival.

■ *Design:* Prospective validation of regression model estimates in an outpatient cohort.

■ *Setting:* University medical center.

■ *Patients:* A total of 1030 consecutive outpatients referred for noninvasive testing for suspected coronary artery disease; 168 of these patients subsequently underwent catheterization within 90 days.

■ *Measurements:* Information from the initial history, physical examination, electrocardiogram, and chest radiograph was used to predict coronary anatomy (the likelihood of any significant coronary disease, severe disease [left main or three-vessel], and significant left main disease) among 168 catheterized patients and to estimate 3-year survival among all patients. These estimates were compared with those based on treadmill testing. Cardiovascular testing charges were calculated for all patients.

■ *Results:* Predicted coronary anatomy and survival closely corresponded to actual findings. Compared with the treadmill exercise test, initial evaluation was slightly better able to distinguish patients with or without any coronary disease and was similar in the ability to identify patients at increased risk for dying or with anatomically severe disease. Based on arbitrary definitions, 37% to 66% of patients were at low risk and responsible for 31% to 56% of the charges for cardiovascular testing.

■ *Conclusions:* The physician's initial evaluation, despite the subjective nature of much of the information gathered, can be used to identify patients likely to benefit from further testing. The development of strategies for cost-conscious quality care must begin with the history, physical examination, and simple laboratory testing.

Annals of Internal Medicine. 1993;**118:**81-90.

From Duke University Medical Center, Durham, North Carolina. For current author addresses, see end of text.

Physicians frequently evaluate patients with symptoms that may represent angina. The initial assessment usually begins with a history, physical examination, electrocardiogram, and chest radiograph. On the basis of this initial assessment, the physician must decide whether to begin empiric therapy or to consider further evaluation with noninvasive testing, cardiac catheterization, or both. Additional testing is often justified on the grounds that much of the information collected in the initial assessment is "soft" data and not sufficiently precise to permit the accurate identification of patients at high or low risk.

Further testing, although often justified, exposes the patient to additional risk and cost. Strategies for evaluating patients with suspected ischemic heart disease that maximize the quality of care while minimizing the use of unnecessary tests depend on the accurate identification of patients who need further evaluation. The accurate identification of high- and low-risk patients based on the physician's initial assessment would permit the development of cost-efficient strategies for evaluating patients with suspected ischemic heart disease.

Stored in the Duke Database for Cardiovascular Disease is the accumulated experience at Duke of all patients with suspected coronary artery disease who were referred for cardiac catheterization (1-7). At the time of cardiac catheterization, findings from the history, physical examination, electrocardiogram, chest radiograph, noninvasive tests, and catheterization are recorded. Patients are then prospectively followed at regular intervals. We have previously developed statistical models that use a subset of this information—the history, physical examination, electrocardiogram, and chest radiograph—to estimate the anatomic severity of catheterization findings and to estimate long-term survival.

Outpatients with chest pain who are evaluated in a physician's office might differ substantially from patients subsequently referred for cardiac catheterization (8). Thus, we were not certain that models developed in the catheterization cohort would perform well when applied to outpatients.

We describe the performance of models based on information from the physician's initial assessment when prospectively applied to a cohort of outpatients. We wished to determine whether a physician's office evaluation of a patient with nonacute chest pain could

identify high- and low-risk patients and to evaluate the potential importance of this information in the delivery of cost-effective quality care.

Methods

Patients

Our study sample included 1030 consecutive, symptomatic patients who had not had previous cardiac catheterization and who were referred for outpatient noninvasive testing at the Duke University Medical Center between 28 March 1983 and 31 January 1985. All patients had complete baseline evaluations that were done prospectively before testing. The sample included 602 patients referred by Duke cardiologists or fellows and 428 patients referred by other physicians at Duke or in the surrounding community. Our study sample comprised a consecutive series of patients with suspected coronary artery disease for whom the physician felt noninvasive testing was warranted. Baseline evaluations were done by a cardiology fellow or physician assistant who completed a standardized form containing all descriptors. The evaluation was facilitated by two other forms: a self-administered questionnaire completed by each patient and a referral form completed by the Duke staff cardiologist (for patients referred by cardiology staff) that together provided all descriptors. Chest pain histories were classified at the time of the patient interview by the examiner. Definitions and further descriptors have been previously described (6, 9-11).

The methods of data management and follow-up have been reported previously (6, 9-11). In brief, baseline information was entered prospectively into the Duke Database for Cardiovascular Disease. Because missing information interrupts the clinical report process, descriptors were complete on all patients. Follow-up information was obtained at 1 and 3 years using a mailed, self-administered patient questionnaire. Patients not returning the questionnaire were contacted via telephone by trained interviewers. For patients who died, we obtained death certificates as well as physician and hospital records (including autopsy information when available), and we conducted telephone interviews with the next of kin to discuss the circumstances of the patient's death. All deaths were classified by an independent events committee (blinded to baseline information).

Analysis

We examined three diagnostic outcomes and one prognostic outcome. The diagnostic outcomes (available only in the 168 patients subsequently referred for cardiac catheterization within 90 days) were the presence of significant coronary artery disease (≥75% luminal diameter narrowing of at least one major coronary artery); the presence of severe coronary artery disease (the presence of significant obstruction of all three major coronary arteries or of the left main coronary artery); and the presence of significant left main coronary artery obstruction. Survival at 3 years was the prognostic end point. In the survival model, patients who were referred for angioplasty or coronary artery bypass graft surgery or who were dying of noncardiovascular causes were censored (withdrawn alive) the first time one of these events occurred.

The development of the predictive models evaluated in our study has been described previously (1-7), and model details are included in the Appendix. In brief, the models were developed in consecutive series of patients referred for cardiac catheterization between 1969 and 1983; none of these patients were included in the present study. The strategy used to develop the models required the division of patients into "training" and "test" samples to minimize spurious associations. Model development in each case was done entirely in the "training" sample. Logistic multiple regression (12) was used for diagnostic outcomes, and the Cox proportional hazards regression model (13, 14) was used for survival. All candidate predictor variables were examined graphically to ensure that their relation with the outcome was modeled appropriately. When nonlinearities were present that would violate model assumptions, appropriate recoding or transformation of the variables was carried out so that model assumptions were satisfied in each case. To decrease the risk for spurious relations and "overfitting" the models, a series of clinical indexes were developed to reflect important areas of pathophysiology

Table 1. Characteristics Used To Estimate Outcomes*

Characteristics	Any Disease	Severe Disease	Left Main Disease	Survival
History				
Age	X	X	X	X
Gender	X	X	X	X
Chest pain				
Type	X	X	X	
Frequency		X		X
Course		X		X
Nocturnal		X		X
Length of time present		X	X	
Diabetes mellitus	X	X		
Smoking	X	X		
Hyperlipidemia	X	X		
Hypertension		X		
Previous history of myocardial infarction	X	X		X
Peripheral or cerebral vascular disease		X	X	X
Congestive heart failure severity				X
Physical examination				
Carotid bruit		X	X	X
Ventricular gallop				X
Electrocardiogram				
Significant Q waves	X	X		X
ST-T wave changes	X	X		X
Conduction abnormalities†				X
Premature ventricular contractions				X
Chest radiograph				
Cardiomegaly				X

* X indicates that the variable is a significant predictor in the multivariable regression model.
† Conduction abnormalities included left bundle-branch block, right bundle-branch block, intraventricular conduction delay, and left axis deviation.

Table 2. Baseline Characteristics in Outpatients

Variable	All Study Patients (n = 1030)	Patients with Cardiac Catheterization (n = 168)
Discrete characteristics, %		
Male sex	63	69
Symptoms		
Typical angina	28	49
Atypical angina	52	47
Nonanginal pain	20	4
Progressive angina	18	24
Nocturnal angina	22	24
Risk factors		
Smoking	44	53
History of hypertension	41	42
Diabetes	10	10
Hyperlipidemia	11	13
Other clinical descriptors		
ST-T wave changes on electrocardiogram	35	42
History of myocardial infarction	18	33
Q waves on electrocardiogram	8	11
History of congestive heart failure	14	11
Class IV congestive heart failure	0	0
Ventricular gallop	1	1
Peripheral vascular disease	3	4
Cerebral vascular disease	3	2
Continuous characteristics*		
Age, *y*	55 (45,63)	56 (48,65)
Pain frequency, *episodes/wk*	2 (1,7)	2 (1,7)
Duration of coronary artery disease symptoms, *mo*	12 (3,36)	7 (3,34)

* Median (25th, 75th percentiles).

(4). Forward stepwise variable selection was used to aid in identifying important baseline predictors. Selected interactions among predictor variables were also examined. When a final model had been developed, it was tested and validated in the independent "test" sample. Baseline variables important for estimating each of the diagnostic and prognostic outcomes are listed in Table 1. Baseline descriptors collected for each patient were entered into each model to generate a patient-specific estimate of the probability of each outcome. Model predictions of the likelihood of significant coronary artery disease, severe coronary artery disease, left main coronary artery disease, and survival at 3 years were generated for each outpatient in this study at the time of his or her initial evaluation based solely on information collected before noninvasive testing.

Assessing the quality of predictions requires the use of statistics unfamiliar to most clinicians. Two components of predictive quality were examined. Reliability, the concordance between predicted and observed outcomes, was assessed by grouping all patients into quantiles of predicted risk and graphically comparing the observed prevalence of the outcome as a function of the mean predicted risk for each quantile group. Discrimination, the ability to separate patients with and without the outcome of interest, was assessed in two ways. First, the distribution of predictions for patients with and for patients without each outcome was graphically compared. Second, a concordance probability or c-index was computed (5). The c-index is calculated by pairing each patient who has the outcome with each patient who does not have the outcome and determining the proportion of patient pairs in which the patient with the outcome had a higher estimated probability. A c-index of 0.80, for example, can be interpreted as follows: Eighty percent of the time a patient with the outcome was given a higher predicted probability of the outcome than the patient without the outcome. The c-index ranges from 1 to 0, with 1 corresponding to perfect discrimination, 0.5 to random performance of a predictor, and 0 to perfectly incorrect discrimination. For a binary outcome, the c-index equals the area under the receiver-operating-characteristic (ROC) curve (15). To further show the discrimination of the survival model, the sample was divided into subgroups of equal size based on the risk for dying within 3 years, and Kaplan-Meier (16) empirical survival curves were calculated.

Placing the Results in Perspective

The two approaches to describing the discriminatory ability of the models (the distribution of predictions for patients with and without the outcome and the c-index) do not effectively communicate a perspective on the importance of information. A traditional approach to showing the discriminatory ability of two tests is to compare the ROC curves of each test. Receiver-operating-characteristic curves show the tradeoff between sensitivity (among patients with the outcome, the proportion with a positive test) and specificity (among patients without the outcome, the proportion with a negative test), as the threshold value above which the test is considered positive is varied. When many characteristics are used in a model to provide the probability of an outcome, ROC curves are generated by first calculating the probability of the outcome for each patient in the sample using his or her individual characteristics. Next, the sensitivity and specificity of the model is calculated several times, as the threshold for what is considered a positive test is "allowed" to vary. For example, when a probability threshold of 0% for a "positive" test is selected, virtually all patients will have higher estimates and sensitivity will be 1.00 and specificity will be 0. At the other extreme, when a threshold of 100% is selected, virtually all patients will have lower estimates and sensitivity will be 0 and specificity will be 1.00. At probability thresholds between 0% and 100%, the ROC curve shows the tradeoff between sensitivity and specificity by illustrating sensitivity as a function of $1 - specificity$. A "better" test, defined as a test with a higher sensitivity *and* specificity will produce a curve that is shifted toward the upper left. A test providing random estimates will produce a 45-degree line.

To show the importance of the initial assessment, we compared the discriminatory ability of the model estimates with the discriminatory ability of a test intuitively more familiar to the clinicians, the treadmill exercise test. The ROC curves for models based on information from the initial assessment were compared with ROC curves for models based on only treadmill exercise test data.

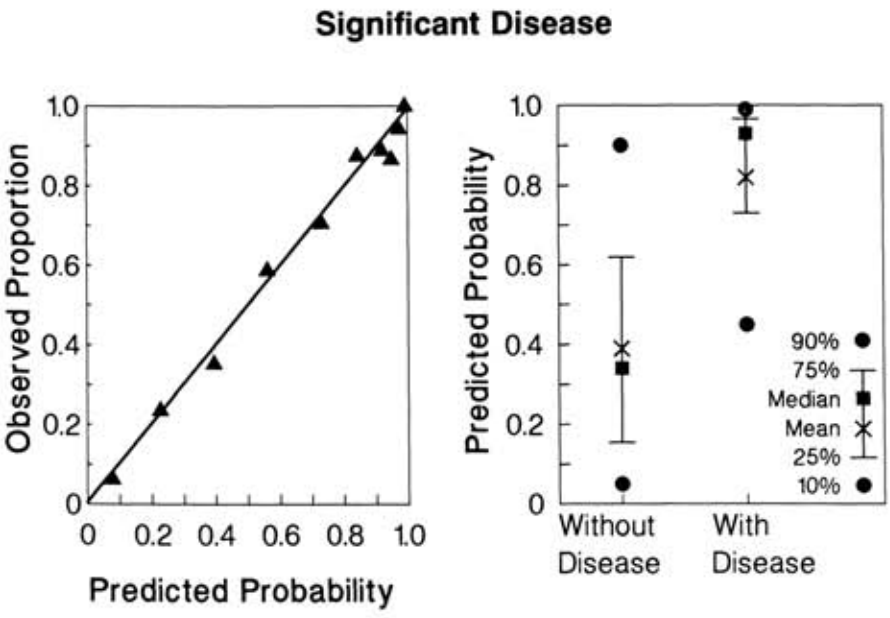

Figure 1. Reliability of the models for estimating the likelihood of significant coronary artery disease, with distributions of predicted probability. Significant disease was defined as 75% or greater narrowing of a major coronary artery. **Left panel.** The observed proportion with significant disease is shown as a function of the mean predicted likelihood of significant coronary artery disease for deciles of predicted risk. The solid line represents "perfect" reliability and is shown for reference. **Right panel.** The distribution of predicted likelihood of significant disease is shown for patients with and without significant disease.

Representative ROC curves for the information from the treadmill exercise test were developed using information stored in the Duke Database on 3104 consecutive symptomatic patients evaluated between 1969 and 1982 who underwent catheterization within 6 weeks of treadmill testing (17, 18). In a fashion similar to that described above, models for each of the outcome variables based on the treadmill test information were developed. Logistic models (for diagnostic outcomes) and Cox models (for survival) were used to build models relating treadmill variables to the outcome. Treadmill variables considered included the following: the interpretation (positive, indeterminate, or negative adequate), the maximum exercise heart rate, the treadmill exercise time, the development of angina during the test or that required stopping the test, and interaction terms reflecting the importance of "early positive" tests (the product of interpretation and the maximum heart rate or exercise time) or "symptomatically positive tests" (the product of interpretation and whether angina developed). Estimates of the likelihood of significant coronary artery disease, severe coronary artery disease, left main coronary artery disease, and 3-year survival from these models were generated for each of the 3104 patients, and corresponding ROC curves were calculated. For each outcome, the area under the ROC curve based on data from the initial assessment was statistically compared with that based on data from treadmill exercise testing (15).

The comparison of the ROC curves based on information from the physician's initial assessment with that from treadmill exercise testing may overestimate the value of the information from treadmill exercise testing because the sample from which the treadmill ROC curves were derived is the same as that in which the models were developed. In contrast, the ROC curves for initial assessment reflect the performance of models developed in one sample and applied to a different sample. Typically, model performance will degrade when applied to a new sample.

A fairer comparison might contrast the ROC curve based on information from the physician's initial assessment with the ROC curve based on the treadmill test data in the subgroup of outpatients who underwent treadmill testing. This subgroup of patients also undergoing catheterization included only 98 patients, of whom only 64, 28, and 7, respectively, had significant, severe, or left main coronary disease. Survival comparisons would have been based on 623 outpatients undergoing treadmill testing, of whom only 4 died within 3 years. Because this approach would have resulted in insufficient outcome events for comparison, the method described above was used.

Finally, we considered the potential effect on patterns of testing that could result if physicians used information from the initial assessment to better identify high-risk patients who would be referred for testing or low-risk patients who might be spared further testing. Two arbitrary definitions of high risk were considered: a predicted 3-year survival of less than 97% or a likelihood of left main disease of more than 2%; and a predicted 3-year survival of less than 95% or a likelihood of left main disease of more than 5%. To examine the effect of using these definitions, we calculated the likelihood of left main disease and of 3-year survival for each patient in our sample. We determined how many patients who had left main disease or who died within 3 years would have been "missed" (incorrectly labeled as low-risk) using these definitions. The proportion of patients at high or low risk in the outpatient study sample was then determined to provide an estimate of the percent of patients at low risk who were referred for testing but for whom greater reliance on information from the initial assessment might have made testing unnecessary. Total charges for testing were also determined based on current (1991) charges at Duke Hospital for tests the patients received. Charges for catheterization include only the laboratory and professional fees but do not include any additional hospital charges to reflect the increasing use of outpatient cardiac catheterization (19).

Results

Baseline characteristics for the 1030 outpatients and the subgroup of 168 patients subsequently referred for cardiac catheterization within 90 days are shown in Table 2. In general, the groups were similar and represented a broad spectrum of findings. Patients referred for catheterization were more likely to be male and to have a history of smoking, typical or progressive angina, and previous myocardial infarction.

Predicting Coronary Anatomy

The quality of the predictions for estimating diagnostic outcomes in the 168 patients referred for catheterization within 90 days is shown in Figures 1, 2, and 3. This subgroup included 109, 45, and 12 patients with

significant, severe, and left main disease, respectively. In Figure 1, *left*, the observed prevalence of significant coronary artery disease is shown as a function of the predicted likelihood of significant coronary artery disease for each decile of predicted risk. The 45-degree line of identity corresponds to perfect reliability and is shown for reference. The observed prevalence of significant disease is nearly identical to that predicted, indicating excellent reliability of the model predictions.

Figure 1, *right*, shows the distribution of predictions for patients with and without significant coronary artery disease. Ninety percent of patients with significant coronary artery disease had predictions of disease greater than 44%, whereas 62% of patients without significant coronary artery disease had predictions of the likelihood of significant coronary artery disease less than 44%. The c-index was equal to 0.87 (95% CI, 0.82 to 0.93), indicating that the model correctly rank-ordered pairs of patients with respect to their disease state 87% of the time.

Corresponding "reliability graphs" and presentations of discrimination are shown for severe disease (left main or three-vessel disease) in Figure 2 and for left main disease in Figure 3. Predictions are lower because anatomically severe disease is less prevalent than significant coronary artery disease. As was the case for significant disease, the predicted likelihood of severe disease and left main disease agreed well with observed outcomes. Although considerable overlap in predictions of the likelihood of severe disease and left main disease was present, the models were able to identify patients likely or unlikely to have these findings. Ninety percent of patients with severe disease had estimates greater than 20%, whereas 55% of patients without severe disease had estimates less than 20%. The c-index for severe disease estimates was 0.78 (95% CI, 0.71 to 0.85). All patients with left main disease had estimates greater than 2%, whereas 31% of patients without left main disease had estimates less than 2%. The c-index for left main disease estimates was 0.73 (95% CI, 0.59 to 0.87). Compared with the c-index for significant disease, the c-indexes for severe disease and left main disease were lower, indicating that these outcomes are harder to predict.

Predicting Survival

Follow-up information was obtained in 973 of the 1030 patients (94%). At the end of 3 years, 844 patients were alive (and had not undergone revascularization), 30 had died of cardiovascular causes, 19 had died of noncardiac causes, 18 had undergone angioplasty, and 62 had had coronary artery bypass graft surgery.

The ability to estimate survival is shown in Figures 4 and 5. The "reliability graph" in Figure 4, *left*, shows that predicted 3-year survival was virtually identical to that observed. The ability to distinguish patients who died from those who lived is shown in Figure 4, *right*. Ninety percent of the patients who died had estimates of 3-year survival less than 97%, whereas 41% of the patients who lived had estimates greater than 97%. The c-index was 0.82 (95% CI, 0.64 to 0.99), indicating that 82% of the time a patient who died was given a lower predicted 3-year survival probability than a patient who lived.

Because so few patients have died, it is difficult to show the discriminatory ability of the survival model. We divided the sample into two groups of equal size based on their model-predicted 3-year survival and calculated Kaplan-Meier survival curves for each group (Figure 5). If no discrimination was present, the two curves would be superimposed on each other. The ability to separate patients into high- and low-risk groups shows in a more familiar way the significant discriminatory ability of the model. The average 3-year predicted survivals for the two groups, 0.98 and 0.92, correspond closely to the observed Kaplan-Meier rates, 0.99 and 0.94.

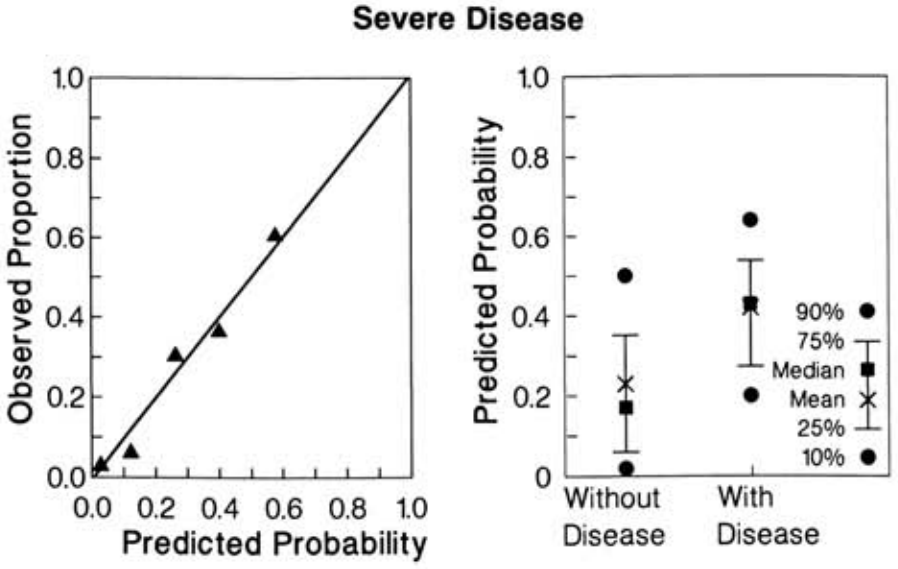

Figure 2. Reliability of the models for estimating the likelihood of severe coronary artery disease, with distributions of predicted probability. Severe disease was defined as left main or three-vessel disease. **Left panel.** The observed proportion with severe disease is shown as a function of the mean predicted likelihood of severe coronary artery disease for quintiles of predicted risk. The solid line represents "perfect" reliability and is shown for reference. **Right panel.** The distribution of predicted likelihood of severe disease is shown for patients with and without severe disease.

Comparing Predictions Based on Physician Assessment and Treadmill Exercise Test

To provide a perspective on the value of the information from the physician's initial assessment, we compared the discriminatory ability of the model estimates with that of the estimates made solely on the basis of information obtained from treadmill exercise testing. In Figure 6, ROC curves for each outcome derived from the information in the physician's initial assessment are compared with representative ROC curves derived from model estimates incorporating information from exercise treadmill testing (*see* Methods). In panel A of Figure 6, the ROC curve for the information from the initial assessment is shifted toward the upper left, indicating superior discrimination for the diagnosis of significant coronary artery disease ($P = 0.005$) compared with the ROC curve for treadmill exercise testing. In panel B, the two ROC curves are similar, indicating equivalent performance for discriminating among patients with or without severe coronary artery disease ($P > 0.2$). Panel C shows a mild trend for superiority of the treadmill exercise test in identifying patients with left main disease ($P > 0.2$), whereas panel D suggests a mild trend for superiority of the information from the physician's initial assessment in recognizing patients likely to die within 3 years ($P > 0.2$), although CIs are wide because so few deaths occurred in the outpatient sample. Thus, the discriminatory ability of a physician relying solely on information from the initial assessment, as opposed to information from treadmill exercise testing, would be superior for coronary artery disease, the same for severe coronary artery disease, slightly worse for left main disease, and slightly better for death within 3 years.

Approaches such as ours could be used to identify low-risk patients for whom additional testing is unnecessary. Among the 1030 study patients, 884 had a known 3-year survival or had left main disease identified at catheterization (the remaining 146 patients underwent revascularization procedures for non-left main disease, were missing follow-up data, or died of noncardiovascular causes). To show the potential value of the physician's initial assessment in risk stratification, consider two examples created by arbitrarily defining low risk as either 1) a predicted 3-year survival of 0.97 or greater and a likelihood of left main disease of 0.02 or less; or 2) a predicted 3-year survival of 0.95 or greater and a likelihood of left main disease of 0.05 or less. Based on the first definition (Table 3), 37% of the sample (324 of 884 patients) would have qualified as low risk and might be considered as not needing additional testing. In this low-risk group, fewer than 1% of patients actually either had left main disease or died within 3 years (the 146 patients not considered in the calculation included only 32 patients identified as being at low risk). Based on the second definition of low risk (*see* Table 3) 66% of the sample (582 of 884 patients) might be spared further testing. In this low-risk group, fewer than 1.5% of patients either had left main disease or died within 3 years.

Efficient use of data collected during the initial physician-patient encounter may also substantially lower the cost of care for patients with suspected ischemic heart disease. As shown in Table 3, the low-risk patients (when arbitrarily defined as having a risk for left main disease ≤2% and a 3-year survival ≥97%) were all referred for further testing (218 patients had treadmill exercise tests and 118 had radionuclide procedures [12 patients had multiple tests]). The additional outpatient testing charges at Duke University for these patients would total $112 054 at current prices. Further, these charges represent 35% of all outpatient charges for these tests among all patients evaluated ($112 054 of $323 224 total). This group also included 25% of patients who subsequently had cardiac catheterization within 90 days. Thus, of the total $556 754 charges (noninvasive testing and catheterization) for the patients considered, $169 954 (31%) occurred in the low-risk subgroup. Using a slightly more liberal definition of low risk, the low-risk subgroup accounted for 64% of out-

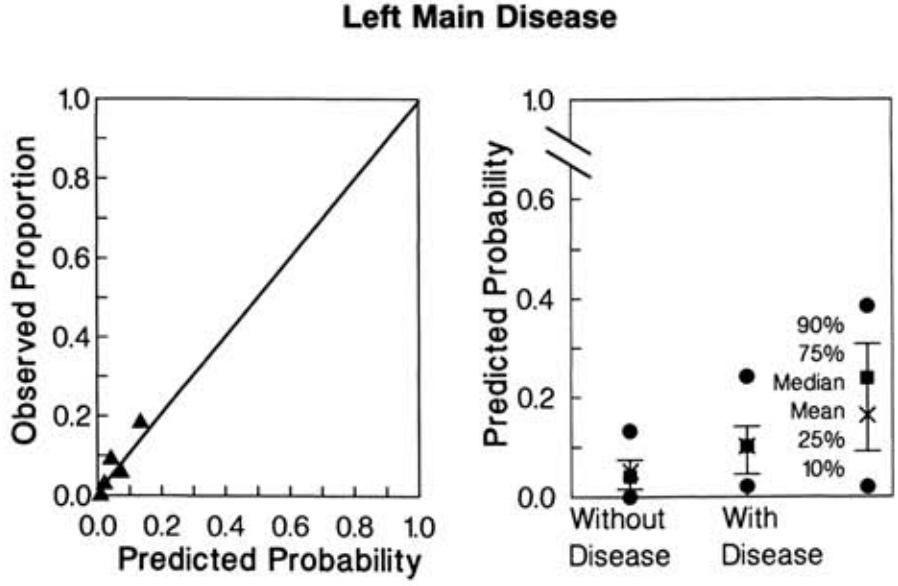

Figure 3. Reliability of the models for estimating the likelihood of left main coronary artery disease, with distributions of predicted probability. Left panel. The observed proportion with left main disease is shown as a function of the mean predicted likelihood of left main coronary artery disease for quintiles of predicted risk. The solid line represents "perfect" reliability and is shown for reference. **Right panel.** The distribution of predicted likelihood of left main disease is shown for patients with and without left main disease.

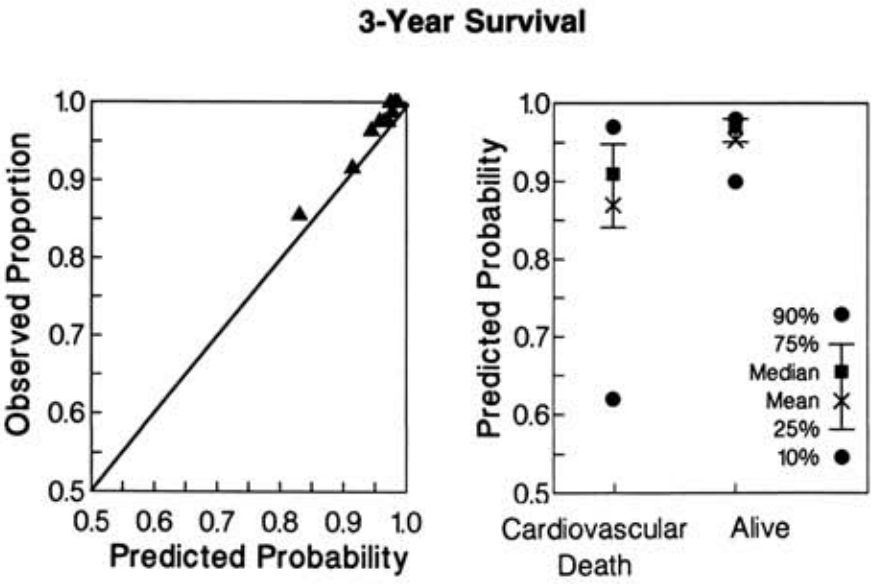

Figure 4. Reliability of the models for estimating the likelihood of 3-year survival, with distributions of predicted probability. Left panel. The observed proportion who survive 3 years is shown as a function of the mean predicted likelihood of 3-year survival for deciles of predicted risk. The solid line represents "perfect" reliability and is shown for reference. **Right panel.** The distribution of predicted 3-year survival is shown for patients who did and did not die within three years of cardiovascular causes.

patient charges ($205 570), 45% of catheterization charges ($106 150), and 56% of total testing charges ($311 720).

Discussion

Our study shows that important diagnostic and prognostic outcomes can be predicted from information collected by the physician as a part of the initial assessment. Ours is the first study to examine the value of the initial assessment prospectively and to apply the lessons learned from prognostic studies of inpatients to a large independent sample of outpatients.

Despite the fact that much of the clinical information collected by a physician is "soft" or subjective data, predictions of outcome based on the information from the initial evaluation are accurate and can be used to identify high- and low-risk patients. In this era of high technology, the physician can select from a wide variety of tests to evaluate patients with suspected coronary artery disease. It is tempting to eschew the information from the initial evaluation in favor of these "objective" tests. Such an approach, however, is inconsistent with the goal of providing cost-conscious quality care, because additional testing increases the risk and cost of evaluation. To restrain the growth of medical care costs while preserving the quality of care, physicians need to rely more on information from their initial evaluations.

The contribution the initial assessment can make to the evaluation process is substantial. The treadmill exercise test is often used to identify high- or low-risk patients. The comparisons between the discriminatory ability of the information from the initial assessment with that from the treadmill exercise test helps to place the potential importance of the initial evaluation in perspective. Aided by information from either the initial evaluation or the treadmill test, physicians will do as well or better by relying on the initial evaluation.

The examples shown in Table 3 suggest that substantial cost savings might accrue by placing greater reliance on the initial evaluation. It is tempting to speculate that 30% to 50% of the total charges related to diagnostic testing in patients with suspected coronary artery disease might be saved. Such a calculation, however, may overestimate the savings available because the tests provide other useful information (for example, functional assessments), and total costs reflect evaluation of patients in many different settings. Thus, we would urge caution in using our findings to support the concept that much of the cost of evaluation in patients with suspected coronary artery disease is unnecessary. Still, it is likely that substantial savings can be realized while maintaining or improving the quality of care if physicians place greater reliance on information obtained in the initial evaluation. We believe that a formal study that examines the effect of routinely available quantitative estimates (or guidelines based on them) on outcomes and cost is warranted.

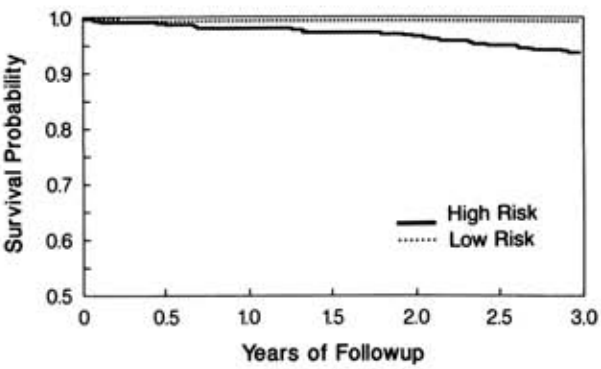

Figure 5. Survival in high- and low-risk patients. Using estimated 3-year survival rates, the sample was divided into two equal-sized groups of patients at "high" and "low" risk. Kaplan-Meier survival curves are shown.

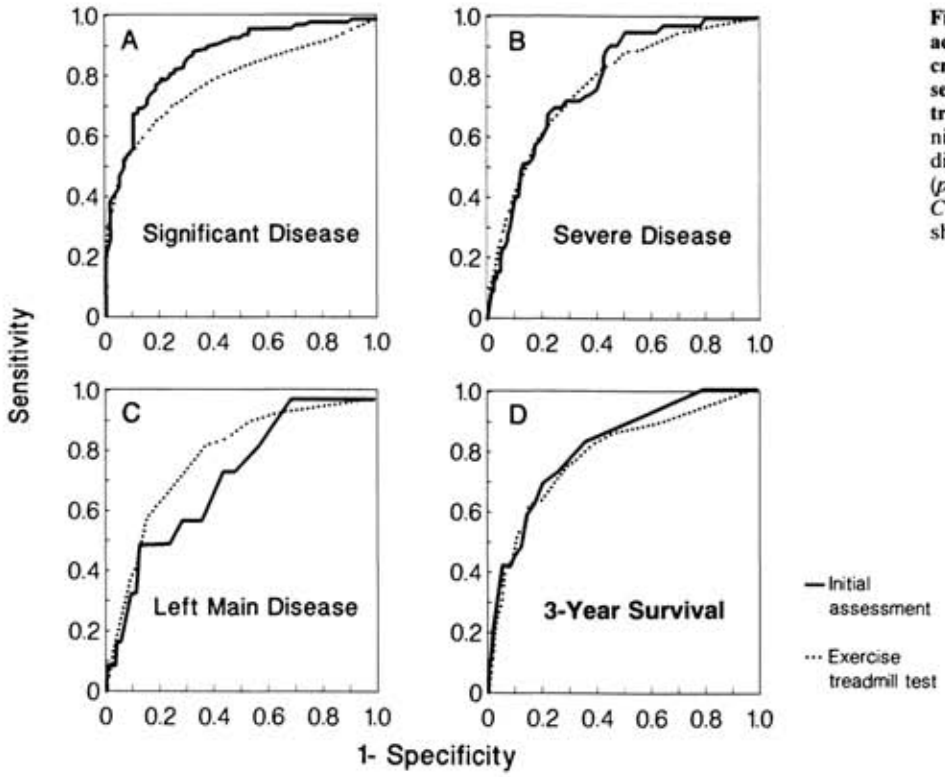

Figure 6. Receiver operating characteristic curves comparing the discriminatory ability of the initial assessment with that of the exercise treadmill test. Results for any significant disease (*panel A*); severe disease (left main or three-vessel) (*panel B*); left main disease (*panel C*); and survival (*panel D*) are shown.

We did a rigorous, prospective evaluation in patients seen at our institution. The models performed remarkably well even though the outpatients probably differed substantially from the inpatients in whom the models were developed. In our study, however, we did not consider whether our findings could be applied in other medical centers without a significant degradation in performance. The evaluation also considered sophisticated statistical model estimates based on a carefully collected database of information and not simple physician estimates. Although the decreasing cost and the availability of computers makes possible the application of such approaches, the findings cannot be applied in settings in which total reliance is placed on physician judgment. Further studies need to address whether the results can be applied in other institutions and to compare the ability of sophisticated models with that of clinician estimates. Previous studies (20-22) and preliminary results in our laboratory suggest that the models that we have developed based on our institutional experience can be applied in other institutions but that the predictive estimates made using such models are superior to those made by expert clinicians. Consequently, we urge caution in assuming that our results can be duplicated solely from physician judgment.

In summary, we examined a large consecutive series of outpatients referred for noninvasive evaluation because of suspected coronary artery disease. Although much of the information obtained by physicians during the initial assessment is subjective, our study confirms the importance of that information in identifying patients likely to benefit from further testing and supports the development of strategies that use the physician's initial assessment in the evaluation process.

Table 3. Distinguishing High-Risk Patients Who Have Left Main Disease or Will Die Within 3 Years

Definition	Left Main Disease or Death	Alive	Total
	← *n* →		
Definition A*			
High Risk	39	521	560
Low Risk	2	322	324
Definition B†			
High Risk	33	269	302
Low Risk	8	574	582

* High risk = likelihood of left main disease of more than 2% or estimated 3-year survival of less than 97%.
† High risk = likelihood of left main disease of more than 5% or estimated 3-year survival of less than 95%.

Appendix: Study Models

Significant Disease

The probability of significant coronary disease was calculated as

$$1/(1 + e^{-x})$$

where e = base of natural logarithm

$$x = a_1y_1 + a_2y_2 + \ldots + a_ky_k + B$$

where $y_1, y_2, \ldots, y_k$ are the characteristics,
$a_1, a_2, \ldots, a_k$ are the corresponding logistic regression coefficients, and
B is the intercept term (in this case, -7.376).

The predictive characteristics are listed below with their coefficients:

Characteristics	Coefficient
Age	0.1126
Sex (0 = male, 1 = female)	−0.328
Age * Sex (interaction)	−0.0301
Typical angina (1 if present)	2.581
Atypical angina (1 if present)	0.976
History of MI (1 if present)	1.093
ECG Q waves (1 if present)	1.213
History of MI * Q waves (interaction)	0.741
Smoking (1 if present)	2.596
Hyperlipidemia (1 if present)	1.845
Diabetes (1 if present)	0.694
ECG ST-T wave changes (1 if present)	0.637
Age * Smoking (interaction)	−0.0404
Age * Hyperlipidemia (interaction)	−0.0251
Sex * Smoking (interaction)	0.550

Severe Disease

The model used to estimate the likelihood of severe coronary artery disease is a conditional probability model. It is the product of two probabilities: the probability of significant disease (provided above) and the probability of severe disease, given the presence of significant disease (SEV|SIG); thus,

$$^{P}\text{SEV|SIG} = 1/(1 + e^{-x})$$

where e = base of natural logarithm

$x = a_1y_1 + a_2y_2 + \ldots + a_ky_k + B$

where $y_1, y_2, \ldots, y_k$ are the characteristics;

$a_1, a_2, \ldots, a_k$ are the corresponding logistic regression coefficients, and

B is the intercept term (in this case, −3.4732).

The predictive characteristics are listed below with their coefficients:

Characteristics	Coefficient
Log_{10} of duration of CAD + 1	0.3424
Type of pain (0 = nonanginal, 1 = atypical, 2 = typical)	0.3014
(Log_{10} of duration of CAD + 1) * type of pain (interaction)	0.1559
Age	0.0299
ECG Q waves (1 if present)	0.3513
Pain index (Typical angina * episodes weekly angina [maximum, 35]) * (1 * progressive pain + 4 * ST-T waves but no Q waves + 2 * presence of nocturnal angina)	0.0054
Sex (0 = male, 1 = female)	−0.3823
Risk factor index (hyperlipidemia + diabetes + hypertension)	0.1734
Vascular disease index (history of peripheral vascular disease + history of cerebrovascular disease + presence of carotid bruits)	0.2402

Left Main Disease

The probability of left main disease was calculated as

$$1/(1 + e^{-x})$$

where e = base of natural logarithm

$x = a_1y_1 + a_2y_2 + \ldots + a_ky_k + B$,

where $y_1, y_2, \ldots, y_k$ are the characteristics,

$a_1, a_2, \ldots$, and a_k are the corresponding logistic regression coefficients, and

B is the intercept term (in this case, −6.7271).

The predictive characteristics are listed below with their coefficients:

Characteristics	Coefficient
Typical angina	1.1252
Age (maximum, 65 years)	0.0483
Sex (0 = male, 1 = female)	−0.5770
Vascular disease index (history of peripheral vascular disease + history of cerebrovascular disease + presence of carotid bruits)	0.5923
Log_{10} of duration of CAD + 1	0.4027

Survival

$$S_t = S_0(t)^{\exp(x)}$$

The probability of being alive at any time (t) is equal to the underlying population survival curve S_0 at time t raised to the e^x power,

where e = base of the natural logarithm

$x = a_1y_1 + a_2y_2 + \ldots + a_ky_k$

where $y_1, y_2, \ldots$, and y_k are the characteristics, and

$a_1, a_2, \ldots$, and a_k are the corresponding Cox proportional hazard regression coefficients.

The predictive characteristics are listed below with their coefficients:

Characteristics	Coefficient
Prognostic pain index (*Episodes of daily angina * (6 * unstable angina + 2 * progressive angina + nocturnal angina + 3 * presence of ST-T wave changes * [1-0.5 * (ECG Q waves)]))	0.0364
Myocardial index (History of CHF + 2 * class IV CHF + cardiomegaly + ECG PVCs + ventricular gallop + 2 * [history of previous MI or Q waves])	0.4506
Vascular disease index (History of peripheral vascular disease + history of cerebrovascular disease + presence of carotid bruits)	0.5333
Conduction index (4 * ECG LBBB + 1 * ECG RBBB + 2 * ECG LAD + 4 * ECG IVCD)	0.08975
Age	0.02260
Sex (0 = male, 1 = female)	−0.6732
Typical angina	0.2952
Centering constant	−1.8833

S_0 = 0.9782 at 1 year
0.9610 at 2 years
0.9422 at 3 years
0.9243 at 4 years
0.9097 at 5 years

Appendix Abbreviations: CAD = coronary artery disease; CHF = congestive heart failure; ECG = electrocardiographic; IVCD = intraventricular conduction defect; LAD = left axis deviation; LBBB = left bundle-branch block; MI = myocardial infarction; PVCs = premature ventricular contractions; and RBBB = right bundle-branch block.

Grant Support: In part by research grants HS-04873, HS-06503, and HS-05635 from the Agency for Health Care Policy and Research; research grant HL-17670 from the National Heart, Lung and Blood Institute; and research grant LM04613 from the National Library of Medicine.

Requests for Reprints: David B. Pryor, MD, Duke University Medical Center, Box 3531, Durham, NC 27710.

Current Author Addresses: Dr. Pryor: Duke University Medical Center, Box 3531, Durham, NC 27710.
Ms. Shaw: Duke University Medical Center, Box 3850, Durham, NC 27710.
Mr. McCants: Duke University Medical Center, Box 3817, Durham, NC 27710.
Dr. Lee: Duke University Medical Center, Box 3363, Durham, NC 27710.

Dr. Mark: Duke University Medical Center, Box 3485, Durham, NC 27710.
Dr. Harrell: Duke University Medical Center, Box 3363, Durham, NC 27710.
Dr. Muhlbaier: Duke University Medical Center, Box 3865, Durham, NC 27710.
Dr. Califf: Duke University Medical Center, Box 31123, Durham, NC 27710.

References

1. **Pryor DB, Harrell FE Jr, Lee KL, Califf RM, Rosati RA.** Estimating the likelihood of significant coronary artery disease. Am J Med. 1983;75:771-80.
2. **Pryor DB, Harrell FE Jr, Rankin JS, Lee KL, Muhlbaier LH, Oldham HN Jr, et al.** The changing survival benefits of coronary revascularization over time. Circulation. 1987;76(Suppl):V13-21.
3. **Califf RM, Harrell FE Jr, Lee KL, Rankin JS, Hlatky MA, Mark DB, et al.** The evolution of medical and surgical therapy for coronary artery disease: a 15-year perspective. JAMA. 1989;261:2077-86.
4. **Harrell FE Jr, Lee KL, Califf RM, Pryor DB, Rosati RA.** Regression modelling strategies for improved prognostic prediction. Stat Med. 1984;3:143-52.
5. **Harrell FE Jr, Califf RM, Pryor DB, Lee KL, Rosati RA.** Evaluating the yield of medical tests. JAMA. 1982;247:2543-6.
6. **Harris PJ, Harrell FE Jr, Lee KL, Behar VS, Rosati RA.** Survival in medically treated coronary artery disease. Circulation. 1979;60: 1259-69.
7. **Pryor DB, Harrell FE Jr, Rankin JS, Lee KL, Muhlbaier LH, Hlatky MA, et al.** Trends in the presentation, management and survival of patients with coronary artery disease—the Duke Database for Cardiovascular Disease. In: Higgins MW, Luepker RV, eds. Trends in Coronary Heart Disease Mortality: The Influence of Medical Care. New York: Oxford University Press; 1988:76-87.
8. **Sox HC Jr, Hickam DH, Marton KI, Moses L, Skeff KM, Sox CH, et al.** Using the patient's history to estimate the probability of coronary artery disease: a comparison of primary care and referral practices. Am J Med. 1990;89:7-14.
9. **Rosati RA, McNeer JF, Starmer CF, Mittler BS, Morris JJ Jr, Wallace AG.** A new information system for medical practice. Arch Intern Med. 1975;135:1017-24.
10. **Harris PJ, Lee KL, Harrell FE Jr, Behar VS, Rosati RA.** Outcome in medically treated coronary artery disease. Ischemic events: nonfatal infarction and death. Circulation. 1980;62:718-26.
11. **Blunden PB, Ross CC, McCants CB, Hammond WE, Muhlbaier LH, Pryor DB.** Protocol-driven automated followup for a large clinical cardiology database. In: American Association for Medical Systems and Informatics. Proceedings, AAMSI Congress. Bethesda, Maryland: The American Association for Medical Systems and Informatics. 1989:135-43.
12. **Harrell FE Jr, Peterson BL.** The LOGIST procedure. In: SUGI Supplemental Library User's Guide. Cary, North Carolina: SAS Institute; 1988;1:15-20.
13. **Cox DR.** Regression models and life-tables (with discussion). J Royal Statist Soc. 1972;34:187-220.
14. **Harrell FE Jr.** The PHGLM Procedure. In: SUGI Supplemental Library Users Guide. Cary, North Carolina: SAS Institute; 1986: 437-66.
15. **Hanley JA, McNeil BJ.** The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982;143: 29-36.
16. **Kaplan EL, Meier P.** Nonparametric estimation from incomplete observations. J Am Statist Assn. 1958;53:457-81.
17. **Mark DB, Hlatky MA, Harrell FE Jr, Lee KL, Califf RM, Pryor DB.** Exercise treadmill score for predicting prognosis in coronary artery disease. Ann Intern Med. 1987;106:793-800.
18. **Mark DB, Hlatky MA, Lee KL, Harrell FE Jr, Califf RM, Pryor DB.** Localizing coronary artery obstructions with the exercise treadmill test. Ann Intern Med. 1987;106:53-5.
19. **Lee JC, Bengtson JR, Lipscomb J, Bashore TM, Mark DB, Califf RM, et al.** Feasibility and cost saving potential of outpatient cardiac catheterization. J Am Coll Cardiol. 1990;15:378-84.
20. **Lee KL, Pryor DB, Harrell FE Jr, Califf RM, Behar VS, Floyd WL, et al.** Predicting outcome in coronary disease. Statistical models versus expert clinicians. Am J Med. 1986;80:553-60.
21. **Boswick JM, Lee KL, Aronson LG, O'Connor CM, Honan MB, Califf RM, et al.** Use of computerized statistical models to educate medical housestaff in estimating clinical outcomes. In: Proceedings: 11th Annual Symposium of Computer Applications in Medical Care. New York: Institute of Electrical and Electronic Engineers; 1987: 162-5.
22. **Kong DF, Lee KL, Harrell FE Jr, Boswick JM, Mark DB, Hlatky MA, et al.** Clinical experience and predicting survival in coronary disease. Arch Intern Med. 1989;149:1177-81.