Independent and Incremental Prognostic Value of Tests Performed in Hierarchical Order to Evaluate Patients With Suspected Coronary Artery Disease

Validation of Models Based on These Tests

Stewart G. Pollock, MD; Robert D. Abbott, PhD; Charles A. Boucher, MD; George A. Beller, MD; and Sanjiv Kaul, MD

Background. The additive prognostic value of tests done in a hierarchical order for the detection of coronary artery disease (CAD) is not always known. The principal goal of this study, therefore, was to assess the incremental prognostic value of data obtained in succession (clinical, exercise stress testing, 201Tl imaging, and coronary angiography) in patients with suspected CAD. A second goal was to develop models for determining prognosis based on results of these tests and to test the clinical validity of these models in unrelated patients.

Methods and Results. Data from two groups of patients who had undergone such evaluation and had been followed for a mean of 4.4 years were analyzed. There were 204 patients from Massachusetts General Hospital (MGH) and 299 from the University of Virginia (UVA). There were 20 deaths and 21 nonfatal infarctions in the MGH group and 41 deaths and nine infarctions in the UVA group. Both univariate and multivariate Cox regression analyses were performed to assess the individual and incremental prognostic value of these tests. In both groups, 201Tl imaging provided significant additional prognostic information compared with clinical and exercise stress test data (p<0.05). At MGH, where the lung/heart 201Tl ratio had been analyzed, coronary angiography did not provide additional prognostic information. In this group of patients, the combination of clinical and exercise 201Tl variables provided greater prognostic information than the combination of clinical and angiographic data (p<0.001). In the UVA cohort, in which the lung/heart ratio had not been analyzed, coronary angiography provided incremental prognostic information compared with clinical and exercise 201Tl data alone (p<0.05). When models developed using data from either sample were applied to the other unrelated sample, there was often close agreement between the overall observed rates and those predicted by the models. This was also true for the low-risk and high-risk subgroups. Some models, however, did not perform as well as other models, which suggests that models that do well in one sample may not always be generalized to other groups.

Conclusions. Tests performed in hierarchical order for the evaluation of suspected CAD provide additional prognostic information. Models developed using clinically relevant combinations of test results obtained from different patient populations are frequently able to predict absolute levels of survival in unrelated but similar samples. (Circulation 1992;85:237–248)

Exercise stress testing, exercise 201Tl imaging, and coronary angiography are established methods of evaluating patients with suspected coronary artery disease (CAD). In addition to providing diagnostic information, these tests have been shown to yield important prognostic information from the National Institutes of Health, Bethesda, Md. Dr. Kaul is the recipient of Clinical Investigator Award KO8-HL-01833 and FIRST Award R29-HL-38345 of the National Institutes of Health.

Address for correspondence: Sanjiv Kaul, MD, Division of Cardiology, Box 158, University of Virginia, Charlottesville, VA 22908.

Received February 26, 1991; revision accepted September 10, 1991.
tion. Many patients undergo multiple tests for the evaluation of suspected CAD. In such patients, whereas the additive diagnostic value of sequential testing has been elucidated, the additive prognostic information obtained from sequentially performed tests is unclear. For instance, does \(^{201}\text{Tl}\) imaging add to the prognostic information already available from clinical and exercise stress test data, and if so, does coronary angiography provide further information in this regard?

The principal goal of this study, therefore, was to assess the prognostic value of clinical, exercise stress test, \(^{201}\text{Tl}\) imaging, and coronary angiographic data in these four individual categories and in various clinically relevant combinations in patients with suspected CAD. A second goal was to develop models for determining prognosis in such patients based on results of these tests and to test the clinical validity of such models in unrelated patient cohorts.

**Methods**

The clinical, exercise stress test, \(^{201}\text{Tl}\) imaging, and coronary angiographic data from both the Massachusetts General Hospital (MGH) and the University of Virginia (UVA) patients have been described in detail in previous reports. A limited summary of patient characteristics and methods is provided here. Data from both patient cohorts were used to develop models for predicting prognosis. To determine goodness-of-fit, these models were evaluated in the same patients from whom they were developed. To determine clinical use, these models were tested in the unrelated patient sample from the other institution undergoing similar testing.

**Patient Populations**

Both groups of subjects (325 from MGH and 383 from UVA) were referred for evaluation of chest pain and underwent both exercise \(^{201}\text{Tl}\) imaging and cardiac catheterization within 3 months of each other between 1978 and 1981. During the same period, approximately 2,800 patients at MGH and 3,800 patients at UVA underwent exercise \(^{201}\text{Tl}\) imaging for known or suspected CAD. The study groups, therefore, constituted 10–12% of all patients undergoing exercise \(^{201}\text{Tl}\) imaging at these institutions during this period. Of those entered in the study, 38% at MGH and 22% at UVA had \(^{201}\text{Tl}\) imaging after cardiac catheterization.

Ten percent of the patients at MGH were lost to follow-up. Their clinical, exercise \(^{201}\text{Tl}\) imaging, and angiographic characteristics were not different from those not lost to follow-up. Only one patient from UVA was lost to follow-up. Because the decision to offer revascularization to patients within 3 months of testing was probably influenced by the results of the tests, thereby potentially avoiding an end point, such patients were excluded from analysis. The patients included in this study, therefore, comprised 204 from MGH and 299 from UVA (characteristics are listed in Table 1). Although MGH patients were younger, the percent of male patients and those with multivessel CAD was similar in the two groups.

**Clinical Data**

The clinical variables analyzed in this study include age, history of infarction, \(\beta\)-blocker therapy, and type of chest pain (typical/atypical). Because most patients in both groups were men, gender was not analyzed.

**Exercise Stress Test**

All patients underwent symptom-limited exercise testing that was terminated for fatigue, claudication, angina, dyspnea, hypotension, or ventricular tachycardia. Heart rate, blood pressure, and a 12-lead electrocardiogram were recorded every minute during baseline, exercise, and a 5-minute recovery period. The duration of exercise was also recorded for each patient. All tests were classified as being either positive or negative. When the baseline electrocardiogram was normal, a test was considered positive if either >1 mm of flat or downsloping or >2 mm of upsloping ST depression was noted 0.08 seconds after the J point. In the presence of baseline ST segment depression but in the absence of left bundle branch block, left ventricular hypertrophy, or digitalis use, the test was considered positive if there was at least an additional 2 mm of ST depression. For purposes of this study, any other test results were considered negative.

**\(^{201}\text{Tl}\) Imaging**

At peak exercise, 1.5–2.0 mCi \(^{201}\text{Tl}\) was injected intravenously and the patient was encouraged to exercise for an additional 30–60 seconds. Planar

### Table 1. Characteristics and Event Rates in Patient Populations

<table>
<thead>
<tr>
<th>Category</th>
<th>Massachusetts General Hospital</th>
<th>University of Virginia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original cohort</td>
<td>325</td>
<td>383</td>
</tr>
<tr>
<td>Lost to follow-up</td>
<td>32</td>
<td>1</td>
</tr>
<tr>
<td>CABG within 3 months of testing</td>
<td>89</td>
<td>83</td>
</tr>
<tr>
<td>Patients entered in present study</td>
<td>204</td>
<td>299</td>
</tr>
<tr>
<td>Men</td>
<td>166 (81%)</td>
<td>244 (82%)</td>
</tr>
<tr>
<td>Women</td>
<td>38 (19%)</td>
<td>55 (18%)</td>
</tr>
<tr>
<td>Mean age (years)</td>
<td>52±10</td>
<td>57±10</td>
</tr>
<tr>
<td>Coronary anatomy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No CAD</td>
<td>35 (17%)</td>
<td>93 (31%)</td>
</tr>
<tr>
<td>One-vessel CAD</td>
<td>61 (30%)</td>
<td>61 (20%)</td>
</tr>
<tr>
<td>Multivessel CAD</td>
<td>108 (53%)</td>
<td>145 (49%)</td>
</tr>
<tr>
<td>Events</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>41 (20%)</td>
<td>50 (17%)</td>
</tr>
<tr>
<td>Deaths</td>
<td>20 (10%)</td>
<td>41 (14%)</td>
</tr>
<tr>
<td>Nonfatal MI</td>
<td>21 (10%)</td>
<td>9 (3%)</td>
</tr>
</tbody>
</table>

CABG, coronary artery bypass graft surgery; CAD, coronary artery disease; MI, myocardial infarction.
images were acquired in the anterior and 45° and 70° left anterior oblique views. In UVA patients, quantitative analysis was performed in seven segments in the anterior and 45° left anterior oblique views while these data were collected. Although quantitative analysis in the MGH cohort was performed in 15 segments from all three views, only seven segments analogous to those analyzed in the UVA group were analyzed for this study. Each segment was examined for the presence of a persistent or redistributing defect. Quantitative lung/heart ratio of 201T1 was analyzed only in the MGH cohort because it was not performed at UVA when the data for this study were collected.

Cardiac Catheterization

Coronary angiographic data were analyzed using calipers and a ruler by a single observer at UVA and by a consensus of three observers at MGH. All angiograms were interpreted with blending to clinical, exercise stress test, and 201T1 imaging data. A significant stenosis was defined as >50% luminal diameter narrowing of one or more of the three coronary arteries or their major branches. The major branch of the left anterior descending artery was the diagonal artery. When present, an intermedia artery was included as a branch of the left anterior descending artery. The major branch of the left circumflex artery was the obtuse marginal artery and the major branch of the right coronary artery was the posterior descending artery. In left dominant systems, the posterior descending artery was considered a major branch of the left circumflex artery.

Follow-up

Follow-up was conducted for up to 8 years (mean, 4.2±2.2 years) in MGH patients and for up to 9 years (4.6±2.6 years) in UVA patients. Follow-up was obtained by contacting the patient or the patient’s family or physician. The end points for the study included cardiac or presumed cardiac-related death or nonfatal myocardial infarction. The latter was documented by a history confirmed by elevation of cardiac enzymes and/or new Q waves on the electrocardiogram. All events were confirmed by the records at the study institutions or the records at the institution where they occurred.

Statistical Methods

For each of the variables analyzed in this report, effects on a cardiac event (death or nonfatal myocardial infarction) were evaluated using Cox proportional hazards regression models. Patients with coronary artery bypass surgery later than 3 months after testing and those suffering noncardiac death were censored at the time of these events. Cox regression models were derived separately in both groups of patients for each of the four categories of risk factors: clinical, exercise stress test, 201T1 imaging, and coronary angiography. Estimates of relative risk corresponding to selected differences between significant variables were calculated. Evaluation of the additional value of each category added sequentially (clinical, clinical and exercise stress test, etc.) was based on increases in the overall likelihood ratio statistic.

To evaluate models for predicting events, the variables identified as significant predictors of events

TABLE 2. Comparison of Significant Predictors of Events by Using Univariate Cox Regression Analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Massachusetts General Hospital</th>
<th>University of Virginia</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Events (n=41)</td>
<td>No events (n=163)</td>
</tr>
<tr>
<td>Age in years (mean)</td>
<td>53 (9)</td>
<td>52 (10)</td>
</tr>
<tr>
<td>β-Blocker (%)</td>
<td>61 (24)</td>
<td>58 (24)</td>
</tr>
<tr>
<td>Previous infarction (%)</td>
<td>61 (24)*</td>
<td>37 (23)</td>
</tr>
<tr>
<td>Typical angina (%)</td>
<td>61 (24)</td>
<td>56 (25)</td>
</tr>
<tr>
<td>Exercise stress test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increase in HR (beats/min)</td>
<td>63 (26)*</td>
<td>69 (25)</td>
</tr>
<tr>
<td>Duration (min)</td>
<td>6.4 (2.8)</td>
<td>7.0 (3.0)</td>
</tr>
<tr>
<td>ST depression (positive test) (%)</td>
<td>56 (25)</td>
<td>52 (25)</td>
</tr>
<tr>
<td>Abnormal blood pressure (%)</td>
<td>10 (9)</td>
<td>13 (11)</td>
</tr>
<tr>
<td>201T1 imaging</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lung/heart ratio (mean)</td>
<td>0.59 (0.15)*</td>
<td>0.48 (0.13)</td>
</tr>
<tr>
<td>Redistributing defects (mean)</td>
<td>1.3 (1.1)*</td>
<td>0.8 (1.1)</td>
</tr>
<tr>
<td>Persistent defects (mean)</td>
<td>3.5 (2.3)*</td>
<td>2.5 (2.0)</td>
</tr>
<tr>
<td>Coronary angiography</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diseased vessels (mean)</td>
<td>2.1 (1.1)*</td>
<td>1.6 (1.1)</td>
</tr>
</tbody>
</table>

HR, heart rate; NA, data not available.

*p<0.05, †p<0.01, ‡p<0.001.
TABLE 3. Estimated Relative Risk for Events by Using Multivariate Cox Regression Analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Compared with</th>
<th>Massachusetts General Hospital</th>
<th>University of Virginia</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Relative risk</td>
<td>95% CI</td>
</tr>
<tr>
<td>Clinical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior MI</td>
<td>No prior MI</td>
<td>2.6</td>
<td>1.4–4.9</td>
</tr>
<tr>
<td>β-Blockers</td>
<td>No β-blockers</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Exercise stress test</td>
<td></td>
<td>1.3</td>
<td>1.0–1.6</td>
</tr>
<tr>
<td>HR response</td>
<td>Compared with HR of 20 beats/min greater</td>
<td></td>
<td></td>
</tr>
<tr>
<td>201TI imaging</td>
<td></td>
<td>2.4</td>
<td>1.8–3.2</td>
</tr>
<tr>
<td>Lung/heart ratio</td>
<td>Compared with ratio of 0.15 less</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segments with redistribution</td>
<td>Compared with two segments less</td>
<td>2.0</td>
<td>1.3–3.2</td>
</tr>
<tr>
<td>Coronary angiography</td>
<td></td>
<td>1.7</td>
<td>1.2–2.2</td>
</tr>
<tr>
<td>Diseased vessels</td>
<td>Compared with one vessel less</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CI, confidence interval; HR, heart rate; MI, myocardial infarction; NA, data not available.

from the Cox regression from one sample were used to predict outcomes in the other sample. Models with and without 201TI data were derived and tested. Predictions were based on logistic regression models adapted for survival data analysis with features that are similar to the Cox model but permit simple derivation of survival curves for a variety of risk factor levels.36,37 Within a study sample, Kaplan-Meier estimates of survival were used to determine actual survival;38 predicted survival was based on regression models derived from the other sample.

In this instance, we were interested in predicting survival based on only the risk factors from the unrelated sample. We were not interested in predictions that became confounded with differences in underlying survival rates between the two cohorts. As a result, predictions in one sample are based on the underlying survival curve for that sample, whereas for the unrelated sample, coefficients from important risk factor relations observed in that sample were used. Any departures from the Kaplan-Meier estimates in one group were attributed to poor prediction from risk factors found to be important elsewhere (see “Appendix” for further details).

The predicted rates were calculated by averaging the individual survival probabilities for each patient within each year. The predicted and observed survival rates were then compared using a Kolmogorov-Smirnov two-sample test.39 This analysis was performed both for the patient cohorts as a whole as well as for low-risk and high-risk subgroups within the cohorts.

Results

Individual Value of Risk Factors for Determining Prognosis

Event rates for both patient samples are depicted in Table 1. Of the four clinical variables analyzed, history of myocardial infarction was the only significant predictor of cardiac events in the MGH cohort (Table 2). Patients with prior infarction suffered an event at a rate more than twice the rate of those without such a history after adjusting for age, β-blocker therapy, and type of chest pain (Table 3). In comparison, the use of β-blockers was the only significant predictor of cardiac events in patients from UVA (Table 2). In this cohort, patients treated with β-blockers had a near doubling of risk compared to those not on these medications after adjusting for age, history of infarction, and type of chest pain (Table 3).

A suboptimal heart rate response during the exercise stress test portended a worse prognosis in both patient samples (Table 2). A patient whose heart rate rose 20 beats per minute less than an otherwise similar patient had a 30% excess risk of an event at MGH and a nearly twofold excess at UVA. The inability to exercise for a longer duration was a significant predictor of events in the UVA sample on univariate analysis but not on multivariate analysis (Table 2). The occurrence of ST segment depression or abnormal systolic blood pressure response (fall or failure to rise during exercise) were not predictors of events in either patient cohort.

All 201TI variables (lung/heart 201TI ratio and the number of persistent and redistributing defects) were significant univariate predictors of events in the MGH sample (Table 2). The lung/heart 201TI ratio superceded the other variables on multivariate analysis (Table 3). A difference in this ratio of 0.15 translated into a greater-than-twofold increase in the risk of events. When this ratio was excluded from analysis, the number of segments demonstrating redistribution was the most powerful 201TI predictor of events. A two-segment difference in the number of segments with redistribution increased the event rate twofold (Table 3). Although the number of persistent
defects was a significant univariate predictor of events (Table 2), it did not predict events on multivariate analysis.

In the UVA cohort, in which lung/heart $^{201}$TI ratio had not been quantitated, the number of segments with redistribution was the best $^{201}$TI predictor of cardiac-related events. A two-segment difference in the number of redistributing segments increased the risk of an event greater than twofold (Table 3). Unlike the results from MGH, the number of persistent defects was not a significant predictor of events in this patient cohort on univariate analysis. The number of diseased coronary arteries was a significant predictor of events at both medical centers. A twofold increase in the event rate for each additional diseased artery was noted in both samples (Table 3). Of note, most of the variables that were not significant in one sample at least had a trend in the same direction as that observed in the sample for which they were significant.

**Additional Value of Sequentially Performed Tests**

At MGH, exercise stress test data were not of significant additional prognostic value to clinical data (Figure 1A), whereas at UVA they provided significa-
number of segments with redistribution). The third (model C) used clinical and coronary angiographic data. Exercise stress test data were not included in these models because they had not been found to provide significant independent prognostic information in the MGH sample. Two models to determine prognosis were also formulated based on the results obtained from the UVA sample. The first (model D) used clinical, exercise stress test, and $^{201}$Ti data, whereas the second (model E) used clinical, exercise stress test, and coronary angiographic data. All five models are described in the "Appendix."

Model A, using clinical information and the lung/heart $^{201}$Ti ratio derived from the exercise $^{201}$Ti data at MGH, was compared with the observed survival in the same cohort. Because the model was applied to the same sample it was derived from, the overall predicted event rate closely approximated the overall actual event rate (Figure 2). Two subgroups were defined to further examine the performance of this model: a high-risk subgroup of 46 patients who had a history of a myocardial infarction and a lung/heart $^{201}$Ti ratio of $>$0.51, which is $>$2SD above the mean normal value, and a low-risk subgroup of 73 patients without a history of myocardial infarction and a lung/heart $^{201}$Ti ratio of $\leq$0.51. In these two situations, as in the entire patient cohort, there were no significant differences between the observed and predicted event rates (Figure 2).

Because the lung/heart $^{201}$Ti ratio had not been derived in this sample, model A could not be validated in the UVA cohort. A model using the number of segments demonstrating redistribution (model B) was applied to this group instead. Before testing the validity of the model, however, it was assessed in the MGH sample from which it was derived. The overall observed and predicted event rates were very close for the entire cohort (Figure 3A). Low-risk (67 patients without a history of infarction and no segments demonstrating redistribution) and high-risk (33 patients with a history of infarction and more than one segment showing redistribution) subgroups were also identified in this patient sample. Similar to the entire cohort, the model provided reasonable predictions of events in these subgroups (Figure 3A). When this model was used in the 299 patients from UVA, it appeared to perform poorly, although there was no significant difference between the overall observed and predicted event rates (Figure 3B). Here, the MGH model overpredicted actual survival in the UVA sample. The results were similar in the low-risk and high-risk subgroups as well (Figure 3B).

The overall predicted event rate using model C, which used clinical data along with the number of diseased vessels on coronary angiography, closely agreed with the overall observed rate in the MGH cohort from which it was derived (Figure 4A). In this sample, a high-risk subgroup of 60 patients with a history of infarction and multivessel CAD and a low-risk subgroup of 70 patients with no prior infarction and either one-vessel CAD or no CAD were identified. As in the entire cohort, the observed and predicted survival curves were also in close agreement for these two subgroups (Figure 4A). When the validity of this model was tested in the UVA cohort, it also provided reasonable prediction of the survival rates of cardiac events in the entire patient cohort as well as in the low-risk and high-risk subgroups (Figure 4B).

A similar result was obtained when model D (developed at UVA) was tested in the MGH cohort. When this model (which was derived using clinical and $^{201}$Ti imaging data) was tested for goodness-of-fit in the UVA sample, it accurately predicted the observed survival in the entire group (Figure 5A). This model also remained reliable in the 60 low-risk patients who were not on β-blockers and who had no evidence of redistribution, as well as in the 48 high-risk patients who were on β-blockers and who had one or more segments showing redistribution (Figure 5A). When the validity of this model was tested in the MGH cohort, it provided reasonable predictions of survival rates of the entire 204 patients and those of the 49 low-risk patients (Figure 5B). The predicted survival was also similar to the observed rates in the 39 high-risk patients for the first 4 years of follow-up. After this period, there was a divergence between observed and predicted rates because
no events were observed between 4 and 8 years (Figure 5B). The number of patients followed after 4 years, however, was small: 18 at 5 years, 11 at 6 years, six at 7 years, and three at 8 years, respectively.

Model E was generated using clinical and coronary angiographic data from the UVA cohort. When tested for goodness-of-fit in the same sample, it closely predicted the survival rates of the entire patient cohort as well as for the 63 low-risk patients (who were not on β-blockers and had one-vessel CAD or no CAD) and the 86 high-risk patients (who were on β-blockers and had multivessel CAD) (Figure 6A). This model also closely predicted event rates when it was tested in the entire MGH cohort (Figure 6B). The model, however, appeared to perform poorly in the low-risk and high-risk subgroups, although the differences between the observed and predicted event rates were not significant. The model overpredicted survival in the 49 low-risk patients (95% versus 82% at 8 years) and underpredicted it in the 69 high-risk patients (50% versus 63%) in this patient group.

Discussion

To our knowledge, this is the first study that has determined the incremental prognostic value of exercise stress testing, exercise 201TI imaging, and coronary angiography to known clinical information in patients undergoing these tests in a hierarchical order for evaluation of suspected CAD. In addition, this is the first time in which the prognostic value of a number of variables has been tested in a patient cohort and validated in an unrelated sample undergoing similar testing. Our study demonstrates the incremental prognostic value of tests performed in succession in patients with suspected CAD and confirms the individual and combined prognostic value of these tests. Our study also provides the unusual opportunity of not only evaluating model performance but also of evaluating risk factor effects on disease outcomes in unrelated samples that are subject to the control of the same group of investigators.

Additive Value of Tests Done in Succession

Several clinical, exercise 201TI imaging, exercise stress testing, and cardiac catheterization variables have been demonstrated to be of prognostic value in patients with known or suspected CAD. The few studies that have compared these variables have either identified the most important predictor of events or have defined predictors in order of their importance. This approach has been
death and nonfatal myocardial infarction. Adding the exercise stress test results to the known clinical variables added to the prognostic information at UVA but not at MGH.

When $^{201}$Tl imaging data were analyzed, the lung/heart $^{201}$Tl ratio and number of defects with redistribution were found to be significant predictors of events. Addition of $^{201}$Tl variables added significantly to the prognostic value of clinical and exercise stress test data, particularly if the lung/heart $^{201}$Tl ratio was known. When the lung/heart $^{201}$Tl ratio was included in the analysis, the number of diseased vessels on coronary angiography did not add to the prognostic information available from clinical, exercise stress test, and $^{201}$Tl data. If, however, this variable was not included in the analysis and only the number of segments with redistribution were counted, the number of diseased vessels on coronary angiography added significantly to the clinical, exercise stress test, and $^{201}$Tl data. The overall prognostic importance of clinical, exercise stress test, and angiographic data was inferior to clinical, exercise stress test, and $^{201}$Tl data when the lung/heart $^{201}$Tl ratio was analyzed.

**Development of Models to Predict Prognosis**

Investigators have used logistic regression analysis for determining prognosis. This approach provides the capability of estimating absolute risk in a particular patient but ignores variability in the follow-up interval of patients from whom the model is derived unless they are properly adapted for the analysis of survival data. Because it is more likely for a patient with the same risk factors to develop an event if he or she is followed for a longer period compared with another followed for a shorter period, models developed using ordinary logistic regression are not valid unless the follow-up period is the same in all patients. To overcome this limitation, Cox proportional hazards analysis or an adapted logistic model is used in which the follow-up interval is factored into the analysis.

In the present study, we used a time-sensitive adaption of the logistic regression analysis to develop models from which the probability of an event for an individual patient could be reported and calculated by the clinician. These models yield results that are very similar to the Cox models. The variables included in our models were selected from two or more of the four categories that were analyzed in this study. The models were developed to answer clinically relevant questions in different scenarios. They were examined in the cohort from which they were derived, whereas their validity was tested in an unrelated sample.

Our results indicate that lung/heart ratio provides the greatest incremental information to clinical and exercise stress test data. When this variable is known, knowledge of coronary anatomy does not provide any additional prognostic information. Furthermore, the combined prognostic information from clinical, exercise stress test, and $^{201}$Tl lung/heart ratio is greater
than that from the combination of clinical, exercise stress test, and angiographic data. This finding should not be surprising. Increased lung/heart $^{201}$TI ratio implies exercise-induced pulmonary edema and has been correlated with the number of diseased vessels on angiography, resting left ventricular function, low double product during exercise, and presence of clinical heart failure.\textsuperscript{29} Using a model based on clinical information and lung/heart $^{201}$TI ratio in the patient cohort from which it was derived, we found that not only was the overall event rate correctly predicted (as would be expected, because the model was derived from this population), but also that low-risk and high-risk subgroups were correctly identified. Because quantitation of the lung/heart $^{201}$TI ratio was not being performed routinely at the time these data were collected, we could not test the validity of the model incorporating this variable in another unrelated sample.

We were able to test models based on clinical information and number of defects with redistribution in unrelated patient samples. The presence or absence of redistribution is routinely assessed on $^{201}$TI imaging and implies the presence of significant stenosis proximal to viable myocardium.\textsuperscript{28} The model developed at UVA performed well in the sample at MGH, but the model developed at MGH did not perform as well in the UVA cohort. Nevertheless, although the latter model was not accurate in determining absolute risk in an unrelated patient cohort, it was very effective in stratifying patients in this sample into low-risk and high-risk subgroups. That is, the relation between redistribution and a cardiac event is duplicated in an unrelated sample, implying that whereas absolute risk in different samples might be different, associations may be generalized.

In many situations, patients undergo cardiac catheterization after exercise testing without also undergoing $^{201}$TI imaging. Models were, therefore, developed to include clinical (and exercise stress test) data along with the number of diseased vessels to predict adverse cardiac outcome. In general, these models also performed well in identifying overall prognosis in an unrelated sample. It is important to note, however, that clinical (and exercise stress test) data added significant information to the angiographic data and substantially enhanced the power of coronary anatomy alone in predicting events. Furthermore, although both models were good at predicting overall survival, the one developed at UVA was not as good in stratifying between low-risk and high-risk subgroups at MGH. Because of the small sample sizes, differences between predicted and observed survival in these subgroups were not statistically significant; however, these differences may be clinically meaningful.

**Study Limitations**

The models were generated from referral populations in tertiary care centers, and although they often seemed useful for separating low-risk and high-risk subgroups, their use in community hospital practice needs to be tested. Selection bias is evident from the fact that the patients studied both at MGH and UVA constituted approximately 10–12% of patients undergoing exercise $^{201}$TI imaging at these institutions at the time these data were collected. Selection bias in terms of gender is also shown in the relatively younger age groups for both patient cohorts in which, as evident from the data, CAD is more prevalent in men. The importance of variables such as lung/heart ratio and redistribution on $^{201}$TI imaging has, however, also been documented in unselected patients not undergoing concomitant coronary angiography.\textsuperscript{2,5,7–9}

Because patients undergoing early bypass surgery (within 3 months of testing) were excluded from analysis, these results pertain to medically treated patients. At the time these data were collected (1978–1981), the decision to offer bypass surgery was based largely on results of the exercise stress test and coronary angiography. Had bypass surgery been included as an end point, the prognostic value of both the exercise stress test and cardiac catheterization would have been greater. When these data were collected, it was not common practice to send patients to surgery based on the results of exercise $^{201}$TI imaging. Consequently, the true independent and additive prognostic value of this test is probably well represented in this study. That a significant minority of patients from both cohorts had $^{201}$TI imaging after coronary angiography was probably because of the need to determine the physiological significance of particular coronary stenoses in these patients rather than to assess their prognosis.

Because cineangiography was not performed in all patients, we did not include left ventricular ejection fraction in the analysis. It is possible that had this variable been included, the prognostic power of cardiac catheterization would have been greater.\textsuperscript{14} The measurement of left ventricular ejection fraction, however, can also be obtained using noninvasive techniques such as blood pool imaging and two-dimensional echocardiography (which can also be acquired during exercise). Consequently, it can be argued that had either of these tests been obtained, the prognostic value of noninvasive testing also could have been higher.\textsuperscript{43,44}

A useful means for validating the models is in observing how well they perform in various subsets of patients. In our study, no significant differences between the observed and modeled survival curves in the low-risk and high-risk groups were noted, and none of the probability values that were derived from such comparisons came close to being noteworthy. Unfortunately, when evaluating model performance in the low-risk and high-risk groups, sample sizes became excessively small and the statistical power to detect failure in the models became limited. In this regard, it is especially interesting to us to try to identify factors that might explain the underprediction and overprediction of a model in the low-risk and high-risk subgroups and the factors that we
failed to observe that can differentiate between the survival experience in rural (UVA) and urban (MGH) communities.

Clinical Implications

There are several practical issues that relate to the results of these studies. First, it would seem that quantitative exercise $^{201}$Tl imaging with measurement of the lung/heart $^{201}$Tl ratio could be advocated as an optimal noninvasive means for determining prognosis in ambulatory patients with suspected CAD. Second, based on our results, it could be argued that the exercise stress test alone (without cardiac imaging) may not provide useful prognostic information in excess of that already provided by the clinical data. Third, it should be emphasized that these results pertain to two institutions with considerable experience with quantitative planar $^{201}$Tl imaging. Whether identical information can be obtained from visual assessment of planar or reconstructed tomographic images needs to be seen. Although there are ample data suggesting that important prognostic information can be derived from planar images, the long-term prognostic value of tomographic imaging has not been yet documented.

Although not tested in this study, it could be argued that similar prognostic information could also be obtained in laboratories with expertise in other forms of cardiac imaging such as exercise radionuclide angiography or exercise echocardiography. Even though a single test may provide powerful prognostic information, it may not reveal the entire prognostic profile of a given patient. For instance, although the levels of risk in patients with three-vessel and one-vessel CAD can be discriminated using coronary angiography, patients with two-vessel disease and no other risk factors such as prior infarction may need to undergo additional testing (such as $^{201}$Tl imaging) to determine their prognosis and, hence, their optimal management. Such issues need further testing to develop cost-effective and efficient clinical algorithms for evaluating patients with suspected CAD.

Appendix

All models are of the form

$$P_I = \frac{1}{1 + \exp(-x)}$$

where $x = \ln\left[ \frac{a}{1 + a} \right]$ (ratio) in model $A$; $-3.81 + 0.81 \times (\text{MI}) + 5.76 \times (\text{segment with redistribution})$ in model $B$; $-4.30 + 0.71 \times (\text{MI}) + 0.46 \times (\text{number of diseased vessels})$ in model $C$; $-2.57 + 0.52 \times (\beta\text{-blocker therapy}) - 0.03 \times (\text{change in heart rate}) + 0.41 \times (\text{segment with redistribution})$ in model $D$; and $-3.57 + 0.52 \times (\beta\text{-blocker therapy}) - 0.04 \times (\text{change in heart rate}) + 0.71 \times (\text{number of diseased vessels})$ in model $E$.

MI (prior myocardial infarction) is coded as 1 for MI and 0 for no MI. $\beta$-Blocker therapy is coded as 1 for such therapy and 0 as not for such therapy.

$P_I$ denotes the probability of an event during the first year of follow-up, conditioned on survival at the beginning of that interval. There were no significant changes in the intercept term with time implying that, for a given set of risk factors, the probability of an event is the same for each year of follow-up, provided that the patient has survived to the beginning of that interval. The estimated probability of survival through the first year of follow-up, $S(1)$, is given by

$$S(I) = \frac{1}{1 - P_1}$$

To test models $B$ and $C$ in UVA patients, the intercepts were used from the UVA population and the coefficients were used from the MGH cohort. For models $B$ and $C$, therefore, the intercepts were $-4.02$ and $-4.65$, respectively, instead of $-3.81$ and $-4.30$, respectively. Similarly, to test models $D$ and $E$ in the MGH patients, the intercepts were used from the MGH population and the coefficients were used from the UVA cohort. These intercepts were, therefore, $-2.55$ and $-3.23$ for models $D$ and $E$, respectively, instead of $-2.57$ and $-3.57$.

Although our use of intercepts from a sample being predicted can improve the performance of the prediction by an unrelated group, we were interested in determining whether failures in prediction could still be observed. If so, the failure of one model to make predictions in an unrelated sample could be better attributed to factors other than underlying forces of morbidity and mortality. There were, in fact, instances when models failed, and in such cases, we believe and expect that unobserved factors exist that might explain differences in absolute survival between unrelated cohorts. The comparisons we provide become especially useful when predictions are made in low-risk and high-risk groups. Of course, the fact that there were no significant failures in prediction could always be attributed to limited sample sizes, although from a clinical point of view, discrimination by an unrelated sample into low-risk and high-risk subgroups generally seems to perform well.

Acknowledgments

Dr. Robert H. Jones, MD, and Dr. Raymond J. Gibbons, MD, first suggested to the authors to assess the incremental value of tests done in hierarchical order to determine prognosis in their patient cohorts. This work has also been influenced by the writings of Dr. George C. Diamond, MD.

References


44. Applegate RJ, Dell’italia JJ, Crawford MH: Usefulness of two-dimensional echocardiography during low-level exercise testing early after uncomplicated acute myocardial infarction. Am J Cardiol 1987;60:10–14


KEY WORDS • coronary artery disease • 201TI • exercise • stress • testing • angiography
Independent and incremental prognostic value of tests performed in hierarchical order to evaluate patients with suspected coronary artery disease. Validation of models based on these tests.

S G Pollock, R D Abbott, C A Boucher, G A Beller and S Kaul

_Circulation._ 1992;85:237-248
doi: 10.1161/01.CIR.85.1.237

_Circulation_ is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 1992 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/85/1/237

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

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

Subscriptions: Information about subscribing to _Circulation_ is online at:
http://circ.ahajournals.org//subscriptions/