Patients with chronic illnesses, such as coronary artery disease, have long periods of time between their disease onset, intervention, and eventual outcome. The physician's inability to couple long term outcomes to the process of patient care confuses the evaluation of therapeutic options and testing procedures. At the Duke University Medical Center, the computer has been used to capture the experience of patients with coronary artery disease and establish the feedback loop necessary for using experience to improve clinical decision making. Methods currently being used for quantitating the diagnostic and prognostic information added by a noninvasive test for an individual patient suspected of having coronary artery disease are described.

THE PROBLEM

The care of the patient with any chronic disease is hampered by the lack of a feedback loop that couples process and outcome. In an acute illness the time intervals between the disease onset, management intervention and outcome are all brief. The physician can improve the process of patient care and learn from previous experience by observing the response to therapy. In contrast, chronic illnesses such as coronary artery disease have long periods of time between onset, intervention, and eventual outcome. The physician is unable to learn from previous experience since therapy cannot be related to long-term outcome. The experience of colleagues is not helpful because they too have no way of relating the effects of their therapy to the outcomes of their patients.

The lack of a feedback loop confuses the evaluation not only of therapeutic options but of testing procedures as well. Diagnostic, management, and prognostic efficacies of different testing procedures are difficult to evaluate since the value of the procedure must often be linked to a temporally distant outcome.

The purpose of this manuscript is to describe the methods currently used at the Duke University Medical Center to quantify the diagnostic and prognostic information added by a test for an individual patient with suspected coronary artery disease and to illustrate the method used in practice to communicate this information to the physician.

*This work was supported by Research Grant HS-03834 from the National Center for Health Services Research, and Research Grant HL-17670 from the National Heart, Lung, and Blood Institute, Training Grant LM-07003 and Grants LM-03373 and LM-00042 from the National Library of Medicine and Grants from the Prudential Insurance Company of America, the Kaiser Family Foundation and the Andrew W. Mellon Foundation.
DETERMINING THE ADDED DIAGNOSTIC AND PROGNOSTIC VALUE OF A TEST

Since the added diagnostic or prognostic value of a test can only be considered in light of all information known previously about the patient, a large experience that characterizes the patient prior to the test, characterizes the results of the test, and links the pre-test characterization and test results with diagnostic and prognostic outcomes is required. The Duke databank of cardiovascular disease represents such a collected clinical experience. The appropriate use of the clinical experience requires an accurate method that incorporates the independent contribution of all characteristics into the expected diagnostic and prognostic outcome for an individual patient.

The Duke Databank For Cardiovascular Disease

The Duke databank for cardiovascular disease began in 1969 to collect the clinical experience of all patients referred for cardiac catheterization because of suspected coronary artery disease. Complete baseline information including patients' histories, physical examinations, electrocardiograms, chest x-rays, and the results of all noninvasive tests were collected prospectively at the time of cardiac catheterization. Over 90% of this patient population has been prospectively followed at six months, one year, and annually thereafter to ascertain symptomatic and functional status and to document nonfatal myocardial infarction and death. Over 8,000 patients are included. By using the computer to cluster and retrieve the experience of similar types of patients, the process of patient care can be coupled to outcome despite the long periods of time separating the disease presentation, intervention, and outcome.

Early in the databank experience, the importance of integrating the process of patient care into the data collection process was recognized. Since 1974, the cardiac catheterization report has been generated from the databank record of the patient. Rather than dictating a catheterization report, the physician completes standardized forms that record the results of the patient's history, physical examination, electrocardiogram, chest x-ray, noninvasive tests, and cardiac catheterization. This information is then entered by the appropriate secretary, data technician, or testing personnel into the databank using coded data entry algorithms. The user-friendly algorithms are designed to require a minimum number of keystrokes and yet ensure the accuracy of data entered with prescribed choice formats and error checking. A report is then automatically generated and reviewed for accuracy by the attending physician before being transmitted to the referring physician. The attending physician has improved the quality of the data collected by ensuring the accuracy of the report. The automatic report generators provide accurate data collection at a minimal cost.

Senior databank personnel include clinical cardiologists, biostatisticians, and computer science experts. Interaction among the team has been fostered by locating all personnel in one area. The biostatisticians have been responsible for developing the methodology used to predict the diagnostic and prognostic outcomes that reestablish the feedback loop necessary to use our previous experience to improve clinical practice.

Predicting Outcomes For Individual Patients

Many different methods can be used to make diagnostic and prognostic predictions for an individual patient. Perhaps the simplest method is subrouping. In this approach the current patient is matched for appropriate prognostic descriptors to a similar group of patients in the databank file. By examining the experience of similar previous patients, an estimate can be made for the current patient. One drawback of this approach is the limited number of patients available for matching, particularly for unusual patients. As the subgroup size or the number of characteristics used for matching is reduced, the accuracy of the prognostic estimates for the current patient will correspondingly diminish. In addition, since long periods of time are required to develop a sufficiently large database, changes in prognosis that occur over time may not be considered when estimating the prognosis for the current patient. Subgrouping methods cannot systematically adjust for the time trends that often are present.

Another method that can be used to estimate the prognosis of current patients is multivariable regression techniques. Multivariable regression models that relate the independent contribution of each important baseline characteristic to prognosis can be developed based on the previous databank experience. Depending upon the method selected, several assumptions are made. Discriminant models typically assume an underlying normal distribution of each of the characteristics. This assumption is violated for many of the baseline characteristics of a population of patients with coronary artery disease. For example, dichotomous variables such as sex will not be normally distributed.

The multivariable regression techniques we prefer are logistic regression for discrete outcomes such as the presence of significant or severe disease, and the Cox proportional hazards model for time dependent outcomes such as survival or infarct free survival. In the latter approach the dependent variable is the time to the event of interest rather than simply whether or not an event occurs at a fixed point during follow-up. Assumptions such as the proportional hazards relationship of the characteristics that are made can be formally tested. The models require efficient computer programming since iterative computational methods must be used to maximize the likelihood functions. With any regression technique, particularly when few outcome events occur in relation to the number of baseline characteristics considered, modeling may result in overfitting the data and spurious associations may be identified. By combining variables into summary indices that describe similar pathophysiologic
processes, overfitting can be minimized and the accuracy of predictions for new patients can be improved.

Recently, interest has centered on applying artificial intelligence techniques to making predictions. Many of these methods are based on recognizing important patterns of characteristics or deriving partitions of patient features that appear to discriminate between patients with different outcomes. At the present time, we prefer the regression models because of their effective use of multiple and continuous patient characteristics.

There is no generally accepted way to determine the best method for making predictions for the current patient. Since the goal of any predictive method is to make accurate predictions, the best method for validating a predictive strategy is to determine the accuracy of the predictions in an independent population. Problems in assessing accuracy occur because there are two components to accurate predictions. The first is reliability: in all patients for whom a 60% chance of coronary disease was predicted, 60% of the patients should have coronary artery disease. The second is the ability to discriminate among patients with different outcomes; what is the average prediction for all patients with disease compared with all patients without disease. The problem expressed in other words is: if the prevalence of disease in a population was 60% and a 60% prediction was given to all patients, the reliability would be perfect but there would be no discrimination ability.

The most successful methods for making predictions in our experience are multivariable regression techniques. Figure 1 illustrates the reliability of a logistic regression model using characteristics from the history, physical examination, chest x-ray, and electrocardiogram to predict probabilities of significant disease in an independent prospective sample of 1,811 patients.

The observed prevalence of disease for subgroups with similar predicted probabilities (for each 0.5 interval) is shown. The median predicted probability for a patient with significant disease was 94% while the median predicted probability for a patient without significant disease was 33%.

While patients whose probability of a given outcome is either very high or very low may not require further work up for a particular outcome, many patients have intermediate outcome probabilities. The physician interested in a particular outcome will not be content with the uncertainty of an intermediate probability. In this setting the physician may wish to consider additional tests to reduce the uncertainty in a predicted outcome. The added value of a test in this setting will be its ability to reliably alter an intermediate pre-test probability. For example, the treadmill test would not be considered useful if a patients pre-test probability of disease was 60% and following the test, the post-test probability remained 60%. Conversely, if the post-test probability of disease for a patient with a pre-test probability of 60% rose to 95%, the treadmill test might be considered useful in establishing the diagnosis.

The Duke database for cardiovascular disease has been used to develop models for predicting the likelihood of significant disease, severe disease (presence of left main or three vessel disease—subgroups thought by many to have an improved prognosis with surgery), survival, and infarct free survival for patients with suspected coronary artery disease at each step of their evaluation. The physician managing an individual patient with coronary disease can be provided with estimates of the likelihood of these outcomes that are based on all information known about a patient at any given point in the patient’s evaluation. In practice, we have incorporated the interpretation of the different tests and have automatically coupled it to the process of patient care.

To illustrate how the system works in practice we can follow a patient through the course of an evaluation. The patient presents to Duke and is referred to a physician for evaluation of suspected coronary artery disease. A history, physical examination, electrocardiogram and chest x-ray are performed. This information is collected and entered using coded data entry algorithms into the databank. A note of the initial assessment is automatically generated from the databank files and includes a brief description of the patient's history, physical examination, pertinent chest x-ray and electrocardiographic results, as well as a computed profile that provides the physician with an estimate of the likelihood of significant disease, severe disease, left main disease, and 1, 3, and 5 year survival probabilities. This patient might be referred for treadmill exercise electrocardiography. The results of that test and the interpretation of the electrocardiographic changes are entered into the databank. The report of the treadmill exercise test is then generated and automatically coupled with a revised profile that incorporates the information from the initial assessment as well as that from the treadmill tests. Figure 2 is an example of a completed treadmill test report. The physician receives a brief summary of the test results as well as the profile that describes the expected outcomes for the patient. The difference between the pre and
post test profiles is the added value of the treadmill test for that particular patient. Similar report profiles are generated (or are in the process of being developed) following radionuclide angiography and cardiac catheterization.

THE FUTURE

Quantifying the added information of a particular test for specific disease outcomes is a significant step in improving the efficiency of evaluation for a patient with suspected coronary artery disease. The increased integrative capabilities of the computer are used to couple the process of patient care with the expected outcome of a chronic disease. In the future we hope to reduce the number of tests providing largely redundant information by demonstrating to the physician that estimates based on the predictive models are at least as accurate as their own. Future analyses will attempt to quantitate the likely value of a contemplated test before it is obtained. The goal in these analyses is to use all information about a patient known prior to a test to estimate the likely yield of a contemplated test procedure. For example, a 50 year old male with atypical angina, who smokes, has a previous history of myocardial infarction but without electrocardiographic evidence of infarction, and has no ST-T wave changes or history of hyperlipidemia or diabetes has a 73% likelihood of significant disease. We hope to quantitate for this patient the likely post-test profile should he receive the treadmill test. This might be expressed, for example, as 95% confidence intervals that would be interpreted as saying "if the treadmill test could be performed in 100 patients like this one, 95% of the time post test probabilities would range between 50 and 98%." If the physician can provide the certainty needed to rule in or out a specific disease outcome, and based on the risk and cost of the test, how likely the test has to be to provide this information, we will be able to reliably estimate whether or not a contemplated test is useful for a particular patient.

CARDIAC DIAGNOSTIC UNIT
DUKE UNIVERSITY MEDICAL CENTER

Apr 25, 1983

TREADMILL EXERCISE REPORT

Exercise Duration: 8 Minutes 8 Seconds
(2 min. 8 sec into Stage 3)

Treadmill Protocol: Bruce protocol

Treadmill Interpretation: NEGATIVE (TARGET HEART RATE REACHED)

Maximum Heart Rate Reached: 180

There was a dysrhythmia (Occasional PVC’s). The test was terminated because of shortness of breath. The patient experienced chest pain thought to represent angina during the test. Max BP noted during exercise was 220/90 and BP at peak exercise was 220/90.

Interpreted by

J. HINES, MD

Reviewed by

D. PRYOR, MD

POST TREADMILL PROGNOSTIC PROFILE:

There are approximately 225 patients in the databank that are like this patient.

Chance of left main disease 0%
Chance of 3-vessel disease (Excluding Left Main) 4%
Chance of any significant disease 23%

Of 100 medically treated patients like this one we would estimate that:
at 1-Year: 99% would be alive
at 3-Years: 98% would be alive
at 5-Years: 97% would be alive

Figure 2.