Applying the Acute Ischemic Heart Disease Predictive Instrument

George A. Corey, MD, and Joel H. Merenstein, MD
Gainesville, Florida, and Pittsburgh, Pennsylvania

A predictive instrument (or index) previously reported to be of value in reducing unnecessary coronary care unit admissions was tested in a randomized study. Acceptability to the physician was then measured by monitoring utilization in a subsequent nonrandomized phase and by debriefing.

The predictive instrument retained predictive accuracy in the new setting with good correlation between predicted and actual risk of acute cardiac ischemia (r = 0.925). False-positive diagnosis rate decreased from 71 percent to 0.0 percent (P = .0096) in a subgroup admitted to the intermediate care unit, consistent with previously reported usefulness in low-risk patients. Acceptability was poor, however, with utilization rate of only 2.8 percent of eligible patients. Debriefing revealed low perceived usefulness. This problem will need to be addressed if widespread utilization is to occur. The criteria of predictive accuracy, usefulness, and acceptability are suggested as a standard panel for testing new predictive instruments.

A recently developed predictive index for acute cardiac ischemia has been reported to be safe and effective in reducing unnecessary coronary care unit (CCU) admissions.[1,2] The authors of these reports proposed the general use of this instrument in hospital emergency departments, with an expected resultant decrease of 250,000 unnecessary CCU admissions each year. Before accepting such an extrapolation, the instrument must be subjected to testing in community hospitals where the resources and visible presence of a research team are absent. A test of the acute ischemic heart disease predictive instrument (AIHD-PI) in such a community hospital setting is reported here.

Three test criteria were defined before beginning the study:

1. Predictive accuracy, ability of the AIHD-PI to estimate risk;
2. Usefulness, ability of this estimate of risk to improve diagnostic accuracy of the physician;
3. Acceptability, ability of the AIHD-PI to be perceived as worthwhile by the emergency department staff who use it.

Submitted, revised, February 3, 1987.

From the Department of Family Practice, San Bernardino County Medical Center, San Bernardino, California, and the Department of Family Medicine, St. Margaret Memorial Hospital, Pittsburgh, Pennsylvania. Requests for reprints should be addressed to Dr. George A. Corey, University of Florida, Alachua General Hospital Family Practice Residency Program, 625 SW 4th Avenue, Gainesville, FL 32601.

The first two criteria were chosen on the basis of published editorial commentary on the subject[3] and the validation difficulties encountered by another predictive index.[4,5] A recent review included these two criteria as important tests for clinical prediction rules.[6] The third criterion was felt to be crucial in that, no matter how accurate or useful a predictive instrument is, it can be of no value if physicians decline to use it.

It should be noted that most published studies of predictive tools utilize the first criterion (predictive accuracy) only.[7–21] Fewer studies address the second criterion (usefulness).[1,2,22,23] Assessment of the third criterion (acceptability) has not been previously addressed.

METHODS

This study had two phases, a randomized controlled use of the predictive index followed by an assessment of actual use combined with a debriefing of physicians. Approval of the hospital Investigational Review Committee was granted for both phases of the study.

Study Setting

The study was conducted at St. Margaret Memorial Hospital in Pittsburgh, Pennsylvania, a 250-bed privately funded community hospital that also supports a non-uni-

versity–affiliated family practice residency. The emergency department is staffed by full-time emergency physicians, part-time private physicians, and moonlighting family practice residents.

Physician and nursing staffs were instructed in the details of obtaining consent and completing their sections of the form. Reprints of the original studies were provided. Emphasis was given to the previously reported value of the predictive index.

Calculation of the Predictive Index Probability

The AIHD-PI utilizes a logistic regression function to evaluate the probability of acute cardiac ischemia based on four historical and three electrocardiographic variables. In the original studies, a research assistant used a programmable calculator to solve the function and provide the emergency department physician with the calculated predictive instrument probability (PIP), the probability of acute cardiac ischemia in a given patient.

Several methods of calculating the PIP have been suggested.[24,25] This study utilized a simple worksheet, allowing rapid calculation of the PIP as well as recording of appropriate data. The PIP was represented as percent chance of acute cardiac ischemia, ranging from 4 percent to 90 percent.

Study Subjects

All men aged 30 years and older and all women aged 40 years or older presenting to the emergency department from October 25, 1984, to March 25, 1985, with a chief complaint of chest pain, shortness of breath, or changed pattern of angina pectoris were eligible. Entry criteria were the same as those used in the multicenter AIHD-PI study.[2] Either the physician or the nursing staff could initiate enrollment of eligible patients. Written consent was obtained from each patient.

An equal number of experimental and control worksheets were placed in closed envelopes and arranged in random order through use of a random number generator. Patients were enrolled when consent was obtained and then were assigned a random form.

Administering the AIHD-PI

The emergency department nurse assigned to each enrolled patient was responsible for obtaining answers to the four historical questions on the form. The physicians interpreted the electrocardiogram and completed the three questions pertaining to it. The physicians then used a table at the bottom of the worksheet to compute the probability of acute ischemia. The same data input and electrocardiogram interpretation were required on the control worksheet, but it did not have a table with which to compute the PIP.

As in the multicenter AIHD-PI study, physicians were instructed that they could incorporate the predictive index probability into their decision making but were not informed on how to do so and were free to ignore it altogether.

Data Collection

Initial diagnosis was determined from the emergency department record. Diagnosis was coded as "acute cardiac ischemia" (acute myocardial infarction or unstable angina) or "not acute cardiac ischemia." When initial diagnosis was not clearly specified on the record, admission to a monitored bed was taken as evidence of suspected acute ischemia.

Follow-up of patients who were not admitted to the hospital was accomplished by telephone interview with the patient and by telephone or mail correspondence with the patient's primary physician. Follow-up of patients who were hospitalized was obtained by reviewing the hospital record and by communication with the primary physician, who was not aware of the experimental or control status of the case. All reviews of hospital records, electrocardiograms, and all communication with patients and their physicians were conducted by one of the authors who was blinded to experimental or control status of each case. The final diagnosis indicated by the primary physician, after inpatient or outpatient investigation, was taken as final. Diagnosis was coded as either "acute cardiac ischemia" or "not acute cardiac ischemia," as above. Follow-up was obtained on 100 percent of the patients. Interval between enrollment and follow-up was between one and five months.

Data Analysis

The false-positive diagnosis rate was calculated as false-positives/(true positives + false positives), where positive applied to any patient whose initial diagnosis was acute cardiac ischemia. True positive thus indicated that the initial impression was correct, and false positive indicated that the initial impression was incorrect. These definitions were identical to those used in the multicenter AIHD-PI study.

Nonrandomized Phase

The goal of this nonrandomized phase was to observe patterns of use by physicians in an uncontaminated model. The nonrandomized phase eliminated the ran-

domization protocol and consent form, items that were believed to be nuisances to a busy emergency department staff. Physicians were instructed to use the AIHD-PI at their discretion. The physician needed only to obtain an AIHD-PI data form from a convenient dispenser and spend from 30 to 60 seconds to calculate the PIP. The forms were numbered to determine how many the physicians were using. Presence of the author in the emergency department was also limited to a weekly visit to monitor utilization of the data forms.

Debriefing of Emergency Department Staff

A structured debriefing interview was designed for both the physicians and the nurses who were present in the emergency department while the trial was under way. The interview was administered at the completion of the study. The six physicians who had utilized the AIHD-PI most frequently were interviewed. Full-time emergency department staff, part-time staff, and moonlighting residents were represented in this group, accounting for 64 percent of enrolled patient encounters. Included in the structured physician debriefing were three questions addressing acceptability-related issues.

Statistical Methods

Actual numbers of patients with acute ischemia were compared with predicted numbers, at each PIP, using a linear regression model. This comparison was utilized to test calibration of the AIHD-PI. Significance testing was based on Pearson's correlation coefficient. The receiver operating characteristic curve was used to assess the ability of the AIHD-PI to distinguish ischemia from no ischemia. Significance testing of this ability was based on a two-sample *t* test of the PIPs of patients with ischemia vs no ischemia. These procedures for testing calibration and discriminant ability are fully described elsewhere.[24] Linear regression and *t* tests were performed using BMDP6D and BMDP3D programs, respectively. Comparison of false-positive diagnosis rates and utilization rates were conducted using a chi-square test corrected for continuity, or Fisher's exact test if so stated, and two-tailed significance levels were reported.

RESULTS

Two patients were excluded from the study after randomization. One patient (experimental group) did not have an electrocardiogram (at the discretion of the emergency department physician) and was discharged from the emergency department with a diagnosis of bronchitis. An electrocardiogram is necessary for calculation of the PIP.

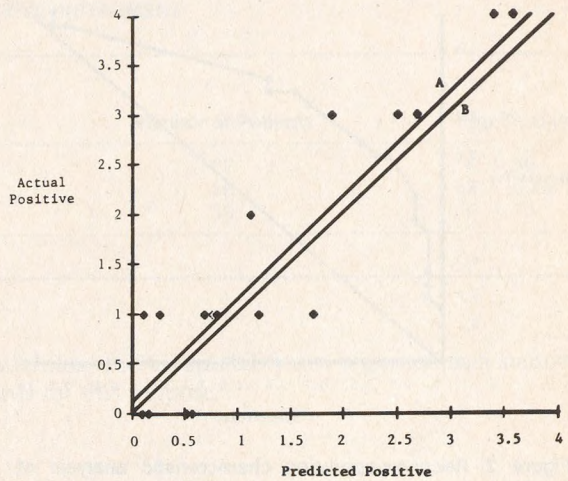

Figure 1. Number of patients who actually had ischemia plotted against number predicted to have ischemia based on calculations of their predictive index probability. Regression line A shows positive correlation (r = 0.925). Perfect calibration is represented by line B

A second patient (experimental or control status unknown) was excluded because the data form was lost in the emergency department.

Predictive Accuracy

Calibration, the match of predicted probabilities of cardiac ischemia with observed probabilities, is examined in Figure 1. Good correlation is seen (r = 0.925, P = .0001) between the number of patients predicted to have ischemia and the number who actually did have ischemia at each PIP level.

Discrimination, the ability of the AIHD-PI to distinguish disease from nondisease, is demonstrated graphically by the receiver-operating characteristic (Figure 2). Patients with acute cardiac ischemia had a higher mean PIP than those without acute ischemia (52.5 percent vs 28.1 percent, P = .0001), demonstrating statistically significant discrimination.

Usefulness

The ability of the AIHD-PI to improve physician diagnostic accuracy is depicted in Table 1. Fewer false positives were admitted to the intermediate care unit (IMC) when the AIHD-PI was used (P = .0096, Fisher's exact test). A trend toward fewer false positives in all monitored beds was seen, but this trend did not reach statistical signifi-

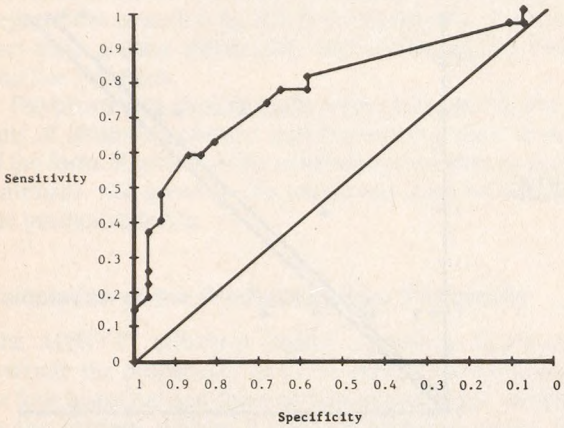

Figure 2. Receiver-operating characteristic analysis of predictive index as though it were a diagnostic test. Hypothetically assigning each predictive index probability as upper limit of normal, sensitivity and specificity relationships are determined. The straight diagonal line represents a test with no discriminant ability

cance. An unexpected trend toward more false-positive diagnoses was observed in the CCU, though this trend also did not reach statistical significance. No false-negative diagnoses were observed in this study.

Acceptability

During the randomized phase of the study, 462 eligible patients were seen in the emergency department and 58 were enrolled. During the subsequent nonrandomized phase, 71 eligible patients were seen in the emergency department, but the AIHD-PI was used in only two cases. Utilization of the AIHD-PI decreased during the nonrandomized phase from 12.6 percent to 2.8 percent ($P = .027$).

Structured debriefing yielded the responses displayed in Table 2. Unsolicited criticisms of the AIHD-PI, offered by more than one physician, addressed specific weaknesses of some of the variables. As an example, a patient with electrocardiographic acute myocardial infarction may have a PIP identical to that of a patient with old right bundle branch block. Credibility was lost in subsequent cases once the physician witnessed such an obvious error.

DISCUSSION

The CCU and the IMC were treated separately because the philosophy of admitting to these units differs at the study institution. Discussion with physician staff revealed that patients with lower likelihood of ischemia were admitted to the IMC rather than the CCU. The positive findings of this study in the IMC were consistent with previous findings, that the AIHD-PI was most helpful in patients with less than definitive or borderline presentations.[1,2] Patients admitted to the CCU have a higher probability of acute ischemia and do not benefit from the use of the AIHD-PI.

The 2.8 percent utilization rate of the AIHD-PI during the real-life phase was clearly disappointing. The support and visible presence of a research effort during the randomized phase appeared to enhance utilization of the AIHD-PI.

Despite the beneficial effect on false-positive diagnosis rate in the IMC, the AIHD-PI was not perceived as being helpful by the physicians who used it. A lack of concordance here is the important finding. While the PIP may supply enough information to influence triage decisions, it may not provide a dramatic or memorable result frequently enough to be perceived as worthwhile. This contention is consistent with the magnitude of effect seen in the large multicenter study of the AIHD-PI. The 30 percent decrease in unnecessary admissions noted in that study would translate to an altered decision in only one in 16 patients presenting to the emergency department with chest pain.[2] If a physician's decision to utilize the AIHD-PI is based on anecdotal recollection, this rate would provide little stimulus for its use.

Cautions are in order regarding the positive findings of this study. Changes in admission behavior may have resulted merely from the awareness that a study intervention was present, a Hawthorne effect.[26] A carefully designed control group was used to minimize this possibility.

Low enrollment rates hinder generalization of the positive findings regarding predictive accuracy and usefulness. The single setting design with a limited number of physicians limits generalizability regarding acceptability results.

CONCLUSIONS

Predictive Accuracy. The AIHD-PI retained predictive accuracy in the new setting. Inspection of the receiver-operating characteristic curve reveals that high sensitivity was available only at loss of considerable specificity. Ability to discriminate disease from nondisease is thus only moderate.

Usefulness. The AIHD-PI retained an ability to alter the triage behavior of physicians in this setting. As in prior studies, the AIHD-PI was shown to reduce the number

TABLE 1. USEFULNESS OF ACUTE ISCHEMIC HEART DISEASE PREDICTIVE INSTRUMENT IN DECREASING FALSE-POSITIVE RATE

Disposition	False-Positive Diagnosis (Percent)		Number of Patients	Significance
	Control	Experiment		
Coronary care unit	18	45	22	P = .36
Intermediate care unit	71	0.0	16	P = .0096*
All monitored beds	39	25	33	P = .74

* *Fisher's exact test*

TABLE 2. RESULTS OF STRUCTURED DEBRIEFING OF PHYSICIANS AFTER USING ACUTE ISCHEMIC HEART DISEASE PREDICTIVE INSTRUMENT (AIHD-PI)

Question	Yes	No
Does the predictive instrument probability (PIP) give you information that you do not already have?	0	6
Does the time required to calculate the PIP present a problem?	1	5
Should the use of the AIHD-PI be continued now that the study has ended?	0	6

of false positives in patients less likely to have acute ischemia. An overall beneficial effect cannot be claimed from these study results, however.

Acceptability. The AIHD-PI was poorly accepted by the physician staff. Readily available to physicians trained to use it, the AIHD-PI simply was not used. Physicians perceived lack of usefulness and variables that could lead to misleading PIP scores in some cases.

Practical use of the AIHD-PI in its current form would certainly fail at the study institution owing to poor acceptability to the physician staff. A list of strategies to improve acceptability should be directed at the deficiencies perceived by the physicians who use it. Improvement of the receiver-operating characteristic to achieve more perceptible information gains is a direct solution, though difficult. Modification of variables that are intuitively disturbing to physicians may also be important. Right bundle branch block, for example, must not influence PIP scores as greatly as acute myocardial infarction.

A three-step approach was needed to characterize fully the performance of the AIHD-PI. Use of fewer criteria would have been quite misleading. The format used in testing the AIHD-PI should be of general value in testing future predictive instruments in both academic and nonacademic settings. The criteria of predictive accuracy, usefulness, and acceptability are suggested as a standard panel for this purpose.

Acknowledgment

John H. Kurata, PhD, MPH, Department of Family Practice, San Bernardino County Medical Center, and Howard E. Rockette, PhD, University of Pittsburgh Graduate School of Public Health, assisted with the statistical analysis.

References

1. Pozen MW, D'Agostino RB, Mitchell JB, et al: The usefulness of a predictive instrument to reduce inappropriate admissions to the coronary care unit. Ann Intern Med 1980; 92(part 1):238–242
2. Pozen MW, D'Agostino RB, Selker HP, et al: A predictive instrument to improve coronary-care-unit admission practices in acute ischemic heart disease. N Engl J Med 1984; 310:1273–1278
3. McNeil BJ, Hanley JA: Statistical approaches to clinical predictions. N Engl J Med 1981; 304:1292–1294
4. Rose CC, Murphy JG, Schwartz JS: Performance of an index predicting the response of patients with acute bronchial asthma to intensive emergency treatment. N Engl J Med 1984; 310:573–577
5. Centor RM, Yarbrough B, Wood JP: Inability to predict relapse in acute asthma. N Engl J Med 1984; 310:577–579
6. Wasson JH, Sox HC, Neff RK, Goldman L: Clinical prediction rules—Applications and methodological standards. N Engl J Med 1985; 313:793–799
7. Goldman L, Weinberg M, Weisberg M, et al: A computer-derived protocol to aid in the diagnosis of emergency room patients with acute chest pain. N Engl J Med 1982; 307:588–596
8. Fischl MA, Pitchenik A, Gardner LB: An index predicting relapse and need for hospitalization in patients with acute bronchial asthma. N Engl J Med 1981; 305:783–789
9. Slap GB, Brooks JSJ, Schwartz JS: When to perform biopsies of enlarged peripheral lymph nodes in young patients. JAMA 1984; 252:1321–1326
10. Lee TH, Cook EF, Weisberg M, et al: Acute chest pain in the emergency room. Arch Intern Med 1985; 145:65–69
11. Mulley AG, Thibault GE, Hughest RA, et al: The course of patients with suspected myocardial infarction. N Engl J Med 1980; 302: 943–948
12. Fuchs R, Scheidt S: Improved criteria for admission to cardiac care units. JAMA 1981; 246:2037–2041

13. Sawe U: Early diagnosis of acute myocardial infarction with special reference to the diagnosis of the intermediate coronary syndrome: A clinical study. Acta Med Scand (suppl) 1972; 545:1–76
14. Nattel S, Warnick JW, Ogilvie RI: Indications for admission to a coronary care unit in patients with unstable angina. Can Med Assoc J 1980; 122:180–184
15. de Dombal FT, Leaper DJ, Staniland JR, et al: Computer-aided diagnosis of acute abdominal pain. Br Med J 1972; 2:9–13
16. Dolgin SM, Schwartz JS, Kressel HY, et al: Identification of patients with cholesterol or pigment gallstones by discriminant analysis of radiographic features. N Engl J Med 1981; 304:808–811
17. Diamond GA, Forrester, JS: Analysis of probability as an aid in the clinical diagnosis of coronary artery disease. N Engl J Med 1979; 300:1350–1358
18. Brush JE, Brand DA, Acampora D, et al: Use of the initial electrocardiogram to predict in-hospital complications of acute myocardial infarction. N Engl J Med 1985; 312:1137–1141
19. DeSmet AA, Fryback DG, Thornbury JR: A second look at the utility of radiographic skull examination for trauma. Am J Radiol 1979; 132:95–97
20. Detrano R, Yiannikas J, Salcedo EE, et al: Bayesian probability analysis: A prospective demonstration of its clinical utility in diagnosing coronary disease. Circulation 1984; 69(3):541–547
21. Gershel JC, Goldman HS, Stein REK, et al: The usefulness of chest radiographs in first asthma attacks. N Engl J Med 1983; 309:336–339
22. Brand DA, Frazier WH, Kohlhepp WC, et al: A protocol for selecting patients with injured extremities who need x-rays. N Engl J Med 1982; 306:333–339
23. de Dombal FT, Leaper DJ, Horrocks JC, et al: Computer-assisted diagnosis of abdominal pain using "estimates" provided by clinicians. Br Med J 1972; 2:350–354
24. Berwick DM, Thibodeau LA: Receiver operating characteristic analysis of diagnostic skill. Med Care 1983; 21:876–885
25. Lo JW: Letter. N Engl J Med 1984; 311:1255
26. Stern R: Letter. N Engl J Med 1984; 311:1255

Commentary

Stephen H. Gehlbach, MD
Tuscaloosa, Alabama

It should surprise no one that physicians are preoccupied with being right. Proper choices on the MCAT examination are essential to medical school admission, and correct responses to the queries of the National Board of Medical Examiners must follow that. The right answers to attending physicians about the pathophysiology of congestive heart failure or the latest enzyme levels of the coronary care patient with a possible heart attack are required to get the proper letters of recommendation to the right residency program. By the time students are ready to assume patient care responsibilities, they are well steeped in the culture of being right, or perhaps, more accurately, of not making mistakes. To be sure, this obsession has its virtues. When the physician makes mistakes, the patient suffers, and most of us would argue quite passionately that the drive to make correct decisions is motivated by concern for the patient's welfare.

It is puzzling, then, that we appear reluctant to accept a growing number of schemes to improve the accuracy of our decision making. The foregoing article by Corey and Merenstein illustrates this disturbing paradox. Evaluating a previously validated index for predicting acute cardiac ischemia in their 250-bed community hospital, these authors confirm the findings of previous investigators[1,2] that this readily computed scoring system, based on a logistic regression model, accurately discriminates between patients with and without acute ischemia. It can reduce the number of patients unnecessarily admitted to the coronary care unit by 25 percent and still properly identify 95 percent of patients with true cardiac ischemia.[2] When Corey and Merenstein evaluated the acceptability of the instrument, however, they found little enthusiasm. During a nonrandomized, follow-up phase of their trial, the index was used on only two of 71 eligible patients. A debriefing session with the emergency department physicians uncovered uniform lack of interest.

Why was this instrument not perceived as useful? According to Corey and Merenstein, the index lost credibility with physicians when they discovered a possible error in the scale. But is this sufficient grounds for rejecting this statistical tool out of hand? Every technology has its limitations, and there is no diagnostic methodology we know of that is error free. Similar resistance to decision-making technology was documented some years ago when Cummins[3] reported on the failure of clinicians to utilize a list of high-yield indications in determining which patients with head trauma required skull radiography. The list of high-yield criteria had been developed from careful clinical studies and shown to be effective at distinguishing head-trauma patients for whom skull radiographs might be useful from those for whom it was a low-yield procedure. Cummins found that almost 80 percent of skull radiographs were requested for patients who had none of the indications on the well-publicized, high-yield list.

When Cummins interviewed the first- and second-year residents responsible for ordering these emergency room radiographs, he found a number of explanations for not adhering to guidelines: disagreement with the criteria, pa-

tient or peer pressure to order examinations, and a feeling that ordering was a routine part of patient evaluation and needed no scrutiny. Fear of malpractice was not an issue. While Cummins found these explanations to be neither "perverse nor irrational," they do not seem compelling reasons for rejecting a simple, effective, decision-making technology.

Is something else occurring? Are there other reasons why we physicians so eagerly embrace new diagnostic imaging techniques and laboratory tests for which efficacy is often poorly demonstrated, yet resist a tested technology to improve decision making? Several observations are worth consideration.

All errors are not alike. In our efforts to make correct diagnostic decisions, we constantly balance the risks of two types of error: We may fail to recognize that a patient has a condition such as a myocardial infarction or skull fracture, or we may decide that the patient has a disease or condition when he does not. In the current decision-making parlance, these are type I and type II errors, respectively, and relate to the often described properties of the sensitivity and the specificity of a diagnostic procedure. Clinicians are much more concerned with type I than with type II errors. We would far prefer to order excessive numbers of examinations or hospitalize too many patients than to miss the diagnosis. In our minds the medical consequences of failing to diagnose disease easily outweigh the cost considerations and iatrogenic risks of overdiagnosis and treatment. This attitude endures despite evidence that the medical benefits of admission to coronary care units or identification of linear skull fractures may be less dramatic than we suppose.[4,5,6] We are activists and prefer doing something for our patients to doing nothing at all. Unfortunately most of the decision-making devices suggest we do less rather than more. Decision-making tools, in short, go against our grain as interventionists.

Overvalued intuition. A second part of the problem may be overestimating our capacity as diagnosticians. The physicians queried by Corey and Merenstein and Cummins found exceptions to the rules offered by the decision-making instruments and felt they could do better by themselves. The literature does not support this notion. The studies on which the acute ischemia and high-yield skull radiograph criteria instruments were based show that the instruments are more accurate than physicians operating on their own. Studies by deDombal et al[7] on computer-assisted diagnosis of abdominal pain are elegant demonstrations that systematic application of Bayesian analysis outperforms the clinician's inconsistent diagnostic reasoning. Scriven[8] argues that a large body of literature supports the superiority of even the simplest statistical models over clinicians' routine performances. The models and computers have no more information than clinicians but process the information more completely and consistently. Human decision makers may form premature hypotheses that bias data gathering or allow recent experience and social or emotional features of the case to distort their estimates of the true probability of disease.

Matter over mind. The idea that statisticians and computers are superior diagnosticians to physicians is the straw that breaks the predictive instrument's back. None of the rest of our evolving new technology is as personally threatening to physicians as decision-making tools. We can accept the superiority of new magnetic resonance imaging and immunofluorescent antibody techniques. They do not compete with our mental processes. But when technology betters the way we think, watch out! No one wants to be replaced by a computer. What were all those years of medical school and residency training good for if a regression equation can do it better? Did we waste all those hours of learning to be right?

Corey and Merenstein suggest that a predictive instrument must be measured on its acceptability to physicians. Obviously these instruments are of little value if they are not utilized. Whether the root of this unacceptability lies in the inadequacies of the instrument or the frailties of human nature, however, requires further thought. Can we as a profession successfully develop and adopt statistical decision-making models into our clinical activities, or are the threats too great? How willing are we to challenge our treasured intuition to really improve our chances of being right?

References

1. Pozen MW, D'Agostino RB, Mitchell JB, et al: The usefulness of a predictive instrument to reduce inappropriate admissions to the coronary care unit. Ann Intern Med 1980; 92:238–242
2. Pozen MW, D'Agostino RB, Selker HP, et al: A predictive instrument to improve coronary-care-unit admission practices in acute ischemic heart disease. N Engl J Med 1984; 310:1273–1278
3. Cummins RO: Clinicians' reasons for overuse of skull radiographs. AJR 1980; 135:549–552
4. Mather HG, Morgan DC, Pearson NG, et al: Myocardial infarction: A comparison between home and hospital care for patients. Br Med J 1976; 1:925–929
5. Hill JD, Hampton JR, Mitchell JRA: A randomized trial of home-versus-hospital management for patients with suspected myocardial infarction. Lancet 1978; 1:837–841
6. Bell RS, Loop JW: The utility and futility of radiographic skull examination for trauma. N Engl J Med 1971; 284:236–239
7. deDombal FT, Leaper DJ, Horrocks JC, et al: Human and computer-aided diagnosis of abdominal pain: Further report with emphasis on performance of clinicians. Br Med J 1974; 1:376–380
8. Scriven M: Clinical judgment. In Engelhardt HT Jr, Spicker SF, Towers B (eds): Clinical Judgment: A Critical Appraisal. Dordrecht, Holland, D Reidel, 1979, pp 3–16

Dr. Gehlbach is Director, Division of Public Health, University of Massachusetts, Amherst, Massachusetts.