SPECIAL SECTION

The Potential of Training to Increase Acceptance and Use of Computerized Decision Support Systems for Medical Diagnosis

Fuji Lai and **Jean Macmillan,** Aptima, Inc., Woburn, Massachusetts, and **Denise H. Daudelin** and **David M. Kent,** Tufts-New England Medical Center, Boston, Massachusetts

Objective: The goals of this study were to understand the reasons underlying the limited use of medical decision-support tools and to explore the potential of a computer-based tutorial to mitigate barriers to use. **Background:** Medical decision-support tools such the Acute Cardiac Ischemia Time-Insensitive Predictive Instrument (ACI-TIPI) have demonstrated statistical validity and clinical impact for patient safety but have seen limited adoption and use. **Methods:** The study developed a brief Web-based "demystifying" ACI-TIPI tutorial employing case-based training and evaluated the effectiveness of that tutorial in changing self-reported attitudes and behaviors. **Results:** Clinicians using the tutorial reported greater understanding of how to use the ACI-TIPI score appropriately and increased confidence in the score. Case studies in the tutorial that provided examples of how to use the score for actual cases were rated as especially helpful. **Conclusion:** This study suggests that a primary barrier to the use of statistical decision support tools for patient diagnosis is lack of training or experience in combining a population-based numerical risk score with other diagnostic information about the individual patient's case that is not considered in that score. The results of this study indicate that there is a potential for a relatively brief tutorial to increase acceptance and use of decision support tools for medical diagnosis. **Application:** These findings have the potential for the identification of methods to help clinicians learn how to use statistical and probabilistic information to better assess risk and to promote integration of decision support tools into medical decision making for improvement of patient safety.

INTRODUCTION

The accuracy of physician diagnosis and treatment choice decisions is a fundamental element of patient safety. Decision support tools, based on statistical analysis of medical data on patient characteristics and case outcomes, seem to have the potential to increase this decision accuracy by drawing on the wealth of available medical data. The importance of promoting the use of such medical informatics tools for the improvement of patient safety has recently come to the forefront in reports such as *Crossing the Quality Chasm: A New Health System for the 21st Century* (Institute of Medicine, 2001).

The ability of statistical decision support tools to increase patient safety by improving the quality of diagnosis and treatment choice decisions can be realized only if these tools are actually used by physicians, however. Despite their demonstrated potential, the tools are not in widespread clinical use. This paper explores some of the barriers to the acceptance and use of these tools and evaluates the potential of a brief computer-based tutorial to overcome these barriers. The goal is to assess whether a small additional investment in providing tutorial information to physicians about the tool might lead to more widespread use.

The study focuses on one decision support tool, which was designed to increase the accuracy of

Address correspondence to Fuji Lai, Aptima, Inc., 12 Gill St., Suite 1400, Woburn, MA 01801; fujilai@aptima.com. *HUMAN FACTORS,* Vol. 48, No. 1, Spring 2006, pp. 95–108.

decisions in hospital emergency departments regarding the treatment of patients who may have acute cardiac ischemia (ACI). The accuracy of these decisions is a critical patient safety issue. An incorrect decision in an ACI case could result in a patient with a severe health problem being sent home, yet the admission of patients who do not prove to have ACI problems consumes valuable time and resources that are needed elsewhere.

Statistical decision support tools can calculate the potential benefits and risks of alternative treatment choices for an individual patient based on the patient's demographic information, such as age and gender, as well as clinical information from the patient's past medical history, physical examination, and laboratory or radiological tests. For some types of decisions, the benefits for a procedure or treatment might apply broadly to a relatively large group of patients, permitting a simple clinical guideline – for example, individuals over the age of 50 should have a colonoscopy every 10 years. For other types of decisions, in which benefits and risks are more evenly balanced or individual patient characteristics are more influential, no simple decision rule can be developed. For such decisions, the statistical risks for an individual patient can be calculated and presented to the clinician through a computer-based decision support system (CDSS). The risks calculated by the CDSS are to be considered along with other information not taken into account by the CDSS in making a final decision about diagnosis and treatment.

A variety of CDSSs have been developed to synthesize and integrate patient-specific information, perform complex evaluations, and then present the results to the physician in a timely manner, typically in the form of a judgment or recommendation, (Hunt, Haynes, Hanna, & Smith, 1998; Wyatt & Speigelhalter, 1991). Evaluation of these tools suggests that they can improve the quality of medical decision making. A 1998 review identified 68 controlled trials of CDSSs (Hunt et al., 1998) that met predefined criteria. Of these, fully two thirds demonstrated benefit in terms of physician performance. However, despite this strong evidence that some CDSSs can improve physician performance, these tools have not, on the whole, taken root in clinical medicine.

The barriers that prevent more widespread adoption and use of the recommendations provided by CDSSs in medicine are not well understood. A variety of explanatory factors have been suggested for the failure to use CDSSs in general, including (a) information fitness – relevance of the processed information in the work context; (b) user trust – trust in automation and transparency of the underlying model; and (c) ease of use – learnability and usability (Zachary, 1988). Research has also shown that successful decision support systems are those that provide a fit between the thought processes of the expert users and the content, form, and timeliness of the information being presented to those users (Lee & Moray, 1994; Masalonis & Parasuraman, 1999).

In this study, we investigated barriers to acceptance and use of a specific CDSS designed to support diagnosis of ACI in emergency departments: the Acute Cardiac Ischemia Time-Insensitive Predictive Instrument (ACI-TIPI). This instrument has been shown in clinical trials to be effective in reducing the number of unnecessary admissions for patients with suspected ACI (Selker et al., 1998). The goals of the study were to develop hypotheses about the factors that prevent the ACI-TIPI from being more broadly used, to identify which of these potential barriers to ACI-TIPI use might be addressed through training, and to evaluate whether a brief computer-based tutorial showed any evidence of being able to remove or reduce these potential barriers. The results of the study, although focused specifically on medical diagnosis, offer insights into the potential of training, even brief training, to increase the acceptance and effective use of CDSSs in other domains.

The Problem of Diagnosing Acute Cardiac Ischemia in the Emergency Department

Each year in the United States more than 6 million patients with symptoms suggestive of ACI present to emergency departments (Selker et al., 1998). The task of identifying the approximately 25% of these patients who actually have ACI (ACI includes patients with either acute myocardial infarction [AMI] or unstable angina pectoris), without filling all available beds with patients who do not have ACI, remains a challenge even for seasoned clinicians (Pope & Selker, 1999). No single clinical symptom, clinical sign, laboratory result, or electrocardiogram (ECG) finding can reliably distinguish most patients with ACI from those with other conditions. Decisions about whether patients with chest pain need urgent admittance

to the hospital must be made quickly with limited information, and physicians have systematic biases in their estimation of risk in these patients (McNutt & Selker, 1988; Tierney, Fitzgerald, & McHenry, 1986). On one hand, there can be significant hazards to avoiding or delaying a needed admission, as the early mortality of AMI and unstable angina is quite high. On the other hand, hospital beds in general, and coronary care unit beds in particular, can be scarce and expensive.

As the number of acute interventions with demonstrated efficacy has increased, and as concern over medicolegal liability has also increased, clinicians have tended to admit any patient with even a low suspicion of ACI. About 60% of patients without ACI, but who present with suggestive symptoms, are admitted (Pope & Selker, 1999). Typically, only about one third (18%–42%) of the 1.5 million patients admitted to coronary care units actually "rule in" for AMI, and only about 50% to 60% have ACI (Pozen, D'Agostino, & Mitchell, 1980; Pozen, D'Agostino, Selker, Sytkowski, & Hood, 1984; Selker, Pozen, & D'Agostino, 1985). The prevalence of ACI for those admitted for possible cardiac ischemia to "step-down" units and the wards is, not surprisingly, much lower still. Even with this relatively low specificity, approximately 5% of patients (i.e., roughly 1 in 20) presenting with actual AMI, and about 10% of patients presenting with actual unstable angina pectoris, are erroneously discharged (Van de Does, Lubsen, & Pool, 1976; Pozen et al., 1980, 1984), and missed AMI diagnosis accounts for about 25% of all malpractice damages paid by emergency physicians (Karcz et al., 1996), far more than any other single diagnosis.

Given the prevalence of ACI, even relatively small improvements in the accuracy and efficiency of diagnosis and facilitating therapy could have enormous patient safety consequences. Thus, the dilemma of whether to admit or discharge a chest pain patient seems well suited for a decision aid that could rapidly process relevant clinical and ECG information about the patient and assist the physician who is making these difficult triage and treatment decisions. Several such decision aids have been developed. These include the original ACI (Pozen et al., 1984) and the ACI-TIPI (Selker, Griffith, & D'Agostino, 1991) tools, the latter of which predicts the probability of ACI as a 0% to 100% probability. The Thrombolytic Predictive Instrument (Selker et al., 1992) provides a prediction of the probability of mortality and other complications of AMI with and without thrombolytic therapy, thus supporting the critical decision of whether to administer thrombolytic therapy. Other predictive models include Goldman's chest pain protocol (Goldman, Cook, & Brand, 1988), which was developed from a large database using recursive partitioning; Baxt's neural network model (Baxt & Skora, 1996; and ACORN, a nonstatistical program based on symbolic rules (Wyatt, 1989). Of these, only the ACI and the ACI-TIPI tools have been demonstrated to have clinical impact in prospective clinical trials.

The ACI-TIPI

The ACI-TIPI is a decision aid developed by the Center for Cardiovascular Health Services Research at Tufts-New England Medical Center (Tufts-NEMC). The instrument predicts the probability of ACI as a 0% to 100% score. The instrument is embedded into commercially available ECG machines (Figure 1) and is printed in real time on the ECG header as a decision aid for the clinician. This is called "time insensitive" because the variables on which the predictions are based can be reliably obtained both during real-time clinical use and by retrospective medical record review, so that the tool can be used for risk adjusting for quality assurance purposes. It is based on a logistic regression equation taking into account seven factors: presence or absence of chest pain, whether or not the chest pain is the primary symptom, age, gender, presence of significant Q waves, presence of ST segment elevation or depression, and presence of T wave inversion or elevation.

APPROACH

We chose to investigate the barriers to acceptance and use of the ACI-TIPI by collecting data from several different perspectives in order to "triangulate" on a better understanding of why clinicians fail to use the tool and what might be done to change that behavior. The study had three phases:

1. *Hypothesis development.* We developed hypotheses about why the ACI-TIPI was not more widely used, based on the general literature on CDSS acceptance and on exploratory interviews with a small sample of expert physicians familiar with the ACI-TIPI.

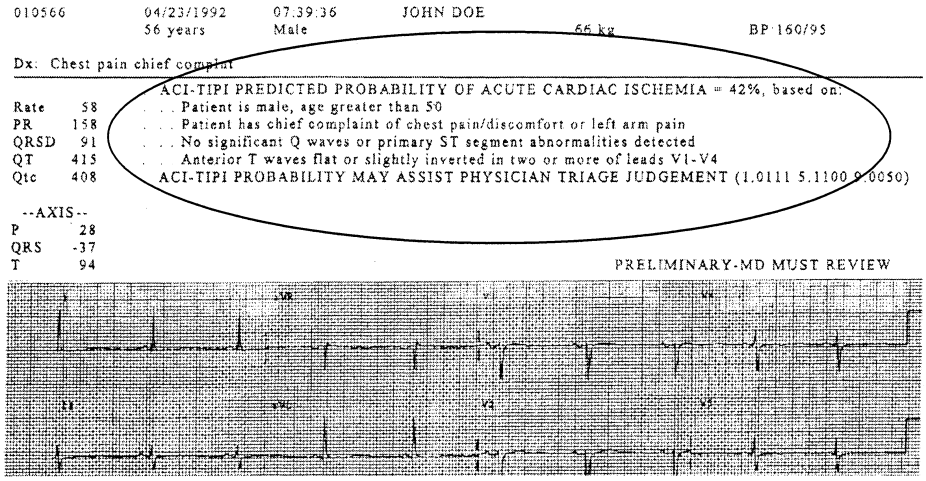

Figure 1. ACI-TIPI printout on header output of ECG machines.

2. *Tutorial development.* We examined these hypotheses to determine which of the potential barriers to use might be affected by training, and we developed a brief Web-based tutorial designed to address these factors.
3. *Tutorial evaluation.* We evaluated whether the tutorial had an effect on ACI-TIPI acceptance and use through (a) observation of diagnostic decision making in standardized ACI cases and (b) a questionnaire that rated attitudes toward the ACI-TIPI and reported use of the instrument. This study was conducted with 16 Tufts-NEMC physicians (internal medicine residents) who were randomly assigned to either receive or not receive the tutorial.

Study Limitations

It should be noted that this effort was a preliminary study with limited time and resources and was designed to assess the possible benefits of further work. The sample size is relatively small, with 5 participants in the preliminary interviews and 16 participants in the main study. The goal was to explore whether a tutorial had the potential to affect the use of the ACI-TIPI tool. The first step was to determine whether the tutorial had a substantial effect (of a magnitude such that it could be detected with a small sample size) on attitudes toward the ACI-TIPI. Given the detection of such an effect on attitudes with a small sample, the planned next step was to assess the effects of the tutorial on actual use using a larger sample.

Development of Hypotheses Regarding Barriers to ACI-TIPI Use

We began our study with a series of hypotheses about possible barriers to CDSS acceptance and use (see Figure 2) based on the general literature, including (a) lack of awareness on the part of decision makers of the existence or purpose of the CDSS; (b) lack of trust in and understanding of how numerical CDSS scores are computed; (c) difficulty in reading or interpreting the information presented; (d) underestimation of the validity and value of the CDSS recommendations; (e) lack of understanding of how to incorporate the score into work flow and decision making; and (f) lack of understanding of how to apply statistical recommendations to specific cases. We believed that some or all of these factors might be affecting the use of the ACI-TIPI.

In order to identify potential barriers to use that are specific to the ACI-TIPI, we conducted unstructured exploratory interviews with expert clinicians who were familiar with the ACI-TIPI tool. (The questions that guided the interview are outlined in Appendix A.) We first assessed their perception of the value of the tool and then addressed specific questions about the use and nonuse of the tool, focusing on concerns about the tool and potential barriers to use. Finally, we invited interviewees to give us their ideas on whether and how a tutorial might increase acceptance and use of the ACI-TIPI.

Our clinician interviewees were 5 expert physicians affiliated with hospitals in the greater Boston area as well as Springfield, Massachusetts, Providence, Rhode Island, and Detroit, Michigan. All of these physicians had substantial clinical experience

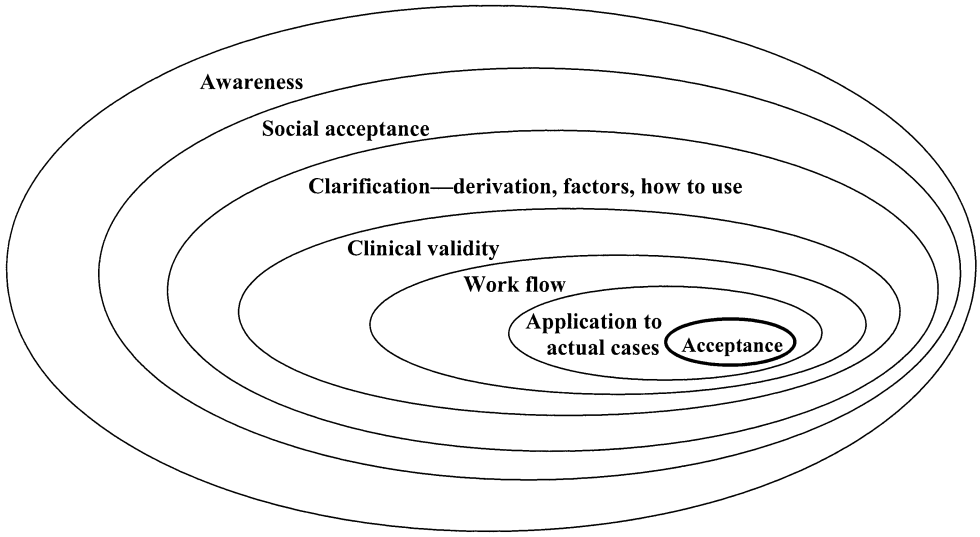

Figure 2. Barriers to adoption of computer-based decision support tools.

evaluating chest pain patients and also had thorough familiarity with the ACI-TIPI through their involvement in prior research studies.

Our interviews confirmed that all of the general barriers to CDSS use that are shown in Figure 2 are present for the ACI-TIPI tool, and they also suggested other barriers we had not anticipated. The issues that our sample of clinicians believed were primary in preventing widespread adoption of the tool included the following:

- lack of awareness of the ACI-TIPI
- lack of understanding of the ACI-TIPI score and how it is derived
- concern that the score does not take all relevant factors into account
- skepticism regarding the validity of the score and whether the score has been endorsed by medical organizations such as the American Heart Association
- social barriers to use, including the perception that it is "bad form" to use the score and "only novices use it"
- uncertainty as to how to use the score appropriately to make decisions for individual patient cases
- medicolegal concerns regarding the role that the ACI-TIPI score might play in justifying decisions if legal issues arise, especially in cases of missed AMIs

Tutorial to Address Barriers to Use

Based on these hypotheses about barriers to use for the ACI-TIPI, we developed a Web-based tutorial designed to address those factors that appear to be amenable to a training intervention. The tutorial was deliberately designed to be extremely brief, requiring not more than 10 min to complete. This time limit was suggested by our sample of expert clinicians, who believed that 10 min was the maximum amount of time that it was realistic to expect clinicians to spend on a tutorial for a single CDSS. Our intent was to test whether such a brief tutorial could have any demonstrable impact on potential barriers to use of the ACI-TIPI.

The tutorial was designed to accomplish the following:

Increase background knowledge about the ACI-TIPI. This includes the factors upon which the probability output is based, so that the user has a mental model of the instrument and underlying processes and can thus arrive at the most comfortable method for incorporating the information from the instrument into his or her own mental decision model.

Increase awareness of the accuracy of the instrument and of its usefulness. This can be done by exposure to studies that have established the validity of the information presented by the instrument and its impact on clinical care. It has been found that physicians were more likely to use an automated diagnostic aid when they felt the aid provided high-quality information and when their own diagnostic confidence was lower (Berner & Maisiak, 1999). The tutorial emphasized that the ACI-TIPI score is based on a database of cases much bigger than any individual clinician would

encounter and also that the actual outcomes of these cases are known.

Increase awareness of the method for clinical use of the instrument. This was done by using three representative case studies that illustrated instances in which the score might have the greatest impact (see example in Figure 3). Physicians were asked to make a judgment regarding each case. They were then given the outcome of the case and told how the ACI-TIPI score could have assisted them in decision making. Key points of emphasis were that the score is designed to supplement, not substitute for, the physician's judgment and that it is important to interpret the ACI-TIPI score in context and to recognize the situations in which it is potentially most useful. A statement also briefly addressed possible medicolegal concerns that had been brought up by the expert clinicians.

Design of Case Studies for the Tutorial

The initial interviews with expert clinicians suggested that case studies were likely to be a valuable teaching mechanism. Case studies have been found to be an effective teaching tool in many domains (Christensen & Hansen, 1987; McNair & Hersum, 1954). They not only drive home points in a more applied manner, they also provide an interactive way for the clinician to learn, as compared with simply reading text.

Based on the interview results and the findings from randomized clinical trials that use of the ACI-TIPI decreases the false positive rate for ACI admissions (Selker et al., 1998) we developed hypotheses (Table 1) on how the ACI-TIPI may influence decisions as a function of a physician's planned disposition of the case without the ACI-TIPI. We used these hypotheses to guide our design of a tutorial to teach the clinician how to best incorporate the ACI-TIPI output into his or her decision-making process.

To develop cases for the tutorial, we identified instances in which we believed it would be beneficial to highlight appropriate use of the ACI-TIPI score in the decision. These were based on our hypotheses from Table 1 concerning the types of situations in which the score may have the greatest impact. The situations that were addressed using the three case studies in the tutorial are described in the following paragraphs. Our tutorial focuses on the upper left and upper center cells of Table 1 (in italics).

Case 1: Provides alert or second opinion. An unexpectedly high score may alert a clinician to revisit a case in further detail. This might be especially useful in flagging cases that could have been missed because the patient did not fit into the stereotypical ACI demographics. For example, studies have shown that the failure of a physician to hospitalize patients presenting with valid ACI symptoms is greater if the patient is not white or if the patient is a woman younger than 55 (Selker et al., 1985).

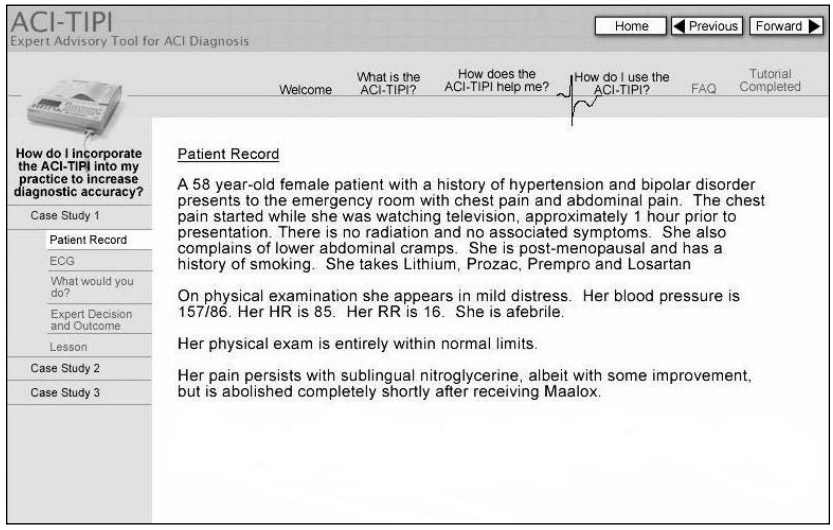

Figure 3. Sample tutorial page.

TABLE 1: Hypothesized Value of ACI-TIPI as a Function of Physician's Planned Disposition About Case Without ACI-TIPI

Physician's Risk Assessment and Planned Disposition Without ACI-TIPI Score	ACI-TIPI Score		
	Low	Medium	High
Low: Can send home with high confidence	*Confirms decision to send home. Increases confidence that nothing was missed in the ECG.*	*Alerts physician to double-check ECG: did I miss something?* *Was another explanation found for symptoms?*	Alerts physician to double-check ECG: did I miss something? Was another explanation found for symptoms?
Medium: Indeterminate – the "gray zone"	*Increases confidence that nothing was missed in the ECG.* *Is the physician aware risk factors not of available to ACI-TIPI?* *Is there another explanation for the symptoms?*	Is the physician aware of risk factors not available to ACI-TIPI? Is there another explanation for the symptoms?	Alerts physician to double-check ECG: did I miss something?
High: Should definitely admit	Is the physician aware of risk factors not available to ACI-TIPI? Is there another explanation for the symptoms?		Confirms admit decision

Note. Cells in italics indicate the hypothesis that these are the cases in which the ACI-TIPI has the greatest impact (based on the finding of a reduced rate of false positives when ACI-TIPI scores were provided).

Case 2: Increases confidence in decision. The score can serve to increase confidence that risk factors have not been overlooked in the ECG and that it is appropriate to send the patient home. In the particular example in the tutorial, it became apparent that there was a reasonable hypothesis/explanation of a non-ACI-related source of the chest pain (e.g., musculoskeletal pain).

Case 3: Caution – Consider risk factors not taken into account by the ACI-TIPI score. In this example, the treating physician in his or her face-to-face encounter with the patient may pick up on important additional information that a machine cannot see, such as the "look of the patient" or the presence of other explanations for the chest pain. Thus it is important to understand that the diagnosis consists of looking at a constellation of factors, of which the ACI-TIPI can "see" only a subset (Figure 4).

Evaluation of the Effects of the Tutorial

Sixteen physicians (internal medicine residents) from Tufts-NEMC served as participants for the tutorial evaluation. The level of experience of these residents varied from approximately 1 to 3 years of postgraduate training, and they all had possible exposure to the ACI-TIPI when admitting patients from the emergency department to the hospital. These residents were block randomized to receive or not receive the tutorial in order to ensure equal numbers in both treatment groups. Participants in the tutorial group were asked to complete the tutorial within 2 weeks of agreeing to participate. Because the tutorial was Web-based and required the participants to log onto a Web site, we were able to check to ensure that all participants in the tutorial group completed the tutorial. The control group did not receive the tutorial. After completing the tutorial, the tutorial group spent a short period (1–4 weeks) when the ACI-TIPI was available to them under usual care conditions in the emergency department. At the end of this period, all participants were called back to complete questionnaires and hour-long interviews. The interviewer was blind to the group assignment of the physicians being interviewed during the first two parts of the interview. The

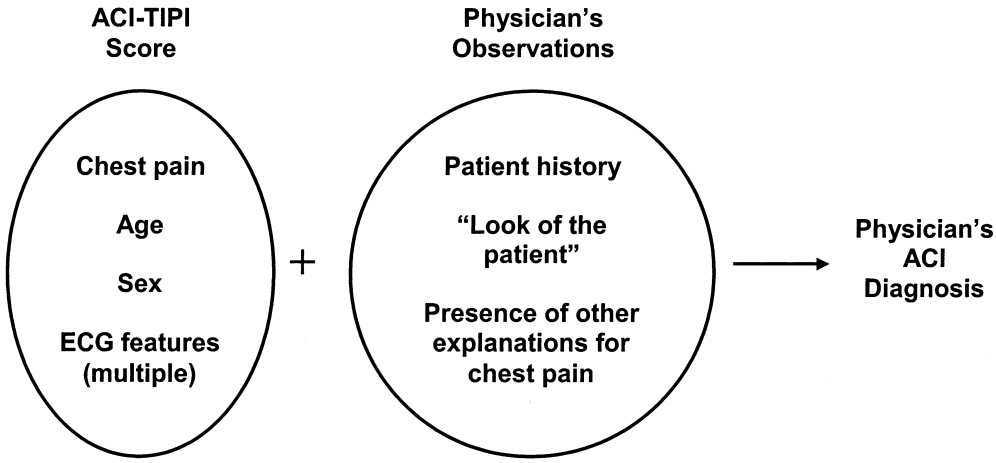

Figure 4. ACI diagnosis constellation of factors.

evaluation interview had the following three components:

Exploration of the diagnostic process for "standardized" patients who were at moderate risk for ACI. A member of the research team acted as a mock patient and provided prescripted answers to questions about history and symptoms. The physician was provided with ECG results that contained the ACI-TIPI score, but no attempt was made to point out or discuss that score during this part of the interview. Physicians were asked to "think aloud" regarding their diagnostic process. The goal was to see if the tutorial had any influence on the physician's use of the ACI-TIPI score to diagnose the standardized cases. In addition, we hoped to gain further insight into how resident physicians obtain information and reason about suspected ACI cases.

A questionnaire with structured questions and rating scales (see Appendix B) concerning attitudes toward and use of the ACI-TIPI. In this part of the interview, physicians were asked about their awareness of the ACI-TIPI, their understanding of the score, and their use of the score in their clinical practice. This part of the interview was intended to assess the level of knowledge of and attitude toward the ACI-TIPI in both groups and, more specifically, whether the tutorial group was more informed about the ACI-TIPI and whether their perceptions of the value of the score or their reported use of the score were affected by the tutorial. We also asked the participants if they could recall the ACI-TIPI score for the two standardized cases they had just seen as an indicator of whether the tutorial had increased the salience of the score for the standardized cases.

A questionnaire with structured questions and rating scales on the usability and value of the tutorial (see Appendix C). At this point the interviewers were "unblinded" as to whether the interviewee had seen the tutorial. Physicians in the tutorial group were asked directly about their experience with the tutorial and their opinion of its value. Members of the control group were free to leave. The purpose of this part of the interview was to determine whether some parts of the tutorial were perceived as more valuable than others and to gain insight for future tutorial improvements.

Development and Use of Standardized Cases for the Evaluation Interview

We used think-aloud protocol techniques along with standardized cases in the evaluation interview in order to gain further insight into the process by which physicians make ACI diagnostic decisions and the role of the ACI-TIPI score in those decisions. Our approach combined in-depth interview techniques for eliciting information with the repeated use of predesigned representative cases, as one might do in a controlled experiment. The resulting methodology provided a means of collecting and representing expert decision processes and information usage patterns in a way that provided additional rigor and internal validation of collected results through replication. The resulting data

supported formal comparison of protocols collected under similar conditions, allowing us to determine the ways that ACI-TIPI information does or might influence diagnostic decision making.

The interviews with representative cases were conducted by two trained interviewers: one to observe, collect data, and push the expert to explain his or her reasoning; and the other to act as a mock patient, responding to the physician's requests for diagnostic symptoms. The mock patient read from predefined cases (lists of all symptoms exhibited by the case patient) in response to queries to ensure that each participant was exposed to the same data each time that case was repeated. Four cases were designed to investigate weights of diagnostic factors. Two patient histories (History A = sharp pain in left chest for 8 hr; History B = dull pressure in center of chest for 24 hr) were developed from the extensive databases available at Tufts-NEMC. These histories were specifically designed to present patients with symptoms that do not clearly rule in or rule out ACI. Two different electrocardiographs (ECG 1 = ACI-TIPI score of 62%; ECG 2 = ACI-TIPI score of 30%), with substantially different ACI-TIPI scores printed on their top margins, were selected to be paired with the clinical histories. The combination of a history and ECG was what we called a "case." The four cases are summarized in Table 2.

Each physician participant was presented with each of the two clinical histories, paired with the ACI-TIPI electrocardiograms, such that the two combinations of clinical histories and electrocardiograms (i.e., A1/B2 and A2/B1) were balanced across participants and treatment groups. Only the first line of the case presentation was given, and the physician participant was then invited to ask questions of the patient as he or she would in an actual clinical setting. They were instructed that their goal was to arrive at a clinical judgment and decision concerning the proper course of action for that patient, given the evidence provided. The participants were asked for a recommended action, the information this decision was based on, the perceived risks of this decision, a confidence rating that this is the best way to manage the patient, and an estimation of the actual probability of ACI. Participants were asked to think aloud whenever possible, so that the interview monitor could capture the experts' reasoning as they considered the evidence. At the end of each case, the interviewer quickly reviewed the evidence considered and asked clarifying questions needed to disambiguate protocol contents.

TABLE 2: Four Cases Used for Interviews

	ECG 1 62%	ECG 2 30%
History A: Sharp pain	Case 1	Case 2
History B: Dull pressure	Case 3	Case 4

Note. Each case is a combination of a patient history and an ECG (with ACI-TIPI score).

RESULTS

The tutorial evaluation interviews produced three types of results: (1) protocol results concerning the use of information to diagnose the standardized cases by the tutorial and non-tutorial groups; (2) questionnaire results concerning understanding of, attitudes toward, and use of the ACI-TIPI in the tutorial and non-tutorial groups; and (3) questionnaire results regarding the usefulness of the tutorial for the tutorial group only.

Results for Diagnosis of Standardized Cases

We did not see systematic differences in the use of or reference to the ACI-TIPI scores between the tutorial and nontutorial groups of physicians. We did, however, see a marked difference among the four standardized cases presented. Recall that the two case histories (A and B) used to create the four cases were quite similar. Both were males of the same age, with some history of prior coronary problems. The pattern of pain reported for the two cases was different, however, with History A (used to create Cases 1 and 2) reporting a sharp pain in the left chest for the previous 8 hr and History B (used for Cases 3 and 4) reporting dull pressure in the center of the chest for the previous 24 hr. ECG 1 (used to create Cases 1 and 3) showed ST segment changes consistent with ischemic heart disease (ACI-TIPI score of 62%), whereas ECG 2 (used to create Cases 2 and 4) was relatively, although not completely, normal (ACI-TIPI score of 30%). The result was that physicians perceived Case 3 (which combined dull pressure for 24 hr with an ACI-TIPI score of 62%) as considerably more serious than Cases 1, 2, and 4. The impressions of the clinicians are summarized in Table 3.

TABLE 3: Reactions to Four Cases

	ECG 1 (ACI-TIPI 62%)	ECG 2 (ACI-TIPI 30%)
History A: Sharp pain	Not too concerned that it is ACI, even though ECG is abnormal	Not too concerned that it is ACI; also, ECG is not abnormal
History B: Dull pressure	Concerned; also, ECG is abnormal	A little concerned that it is ACI, but ECG is not abnormal

These results provide insight into the relative weights these physicians placed on the abnormalities present in the ECG and the patient's reported history of pain (the only factor that differed significantly between Histories A and B) in evaluating the cases. It required both the abnormal ECG and the more worrisome pain history to substantially raise the physician's level of concern about the case. In the absence of an abnormal ECG, the pain history was not considered likely to be serious. The abnormal ECG did not, by itself, raise the level of concern, however. The less worrisome pain history and the relatively normal ECG was, as might be expected, the case with which physicians seemed least concerned, with 2 physicians suggesting that they would send the patient home and the remainder suggesting admission to the observation unit or observation in the emergency department.

These results confirm the reports by expert physicians during the initial interviews (used to design the tutorial) that their decision process places considerable weight on factors that are difficult to measure precisely and are not considered in the ACI-TIPI score, such as the nature of the pain reported by the patient. The context in which a numerical risk score is to be applied – including factors in the individual case that are not included in the score, such as pain history – is clearly critical in the acceptance and use of the score. Our initial interviews indicated that clinicians receive no training or guidance in how to interpret and use the ACI-TIPI score in the context of actual cases, and the results for the standardized cases confirm our belief that such training has the potential to increase ACI-TIPI use.

Results for Evaluation Questionnaire

We had a number of hypotheses about why physicians may fail to use the ACI-TIPI score, and we used these hypotheses to guide the design of a questionnaire (see Appendix B) about the value and use of the ACI-TIPI score that was administered to physicians who had received the tutorial as well as those who had not. The questionnaire included questions designed to assess awareness and understanding of the ACI-TIPI, the perceived value of the ACI-TIPI score, confidence in the score, use of the score, and concerns about the score. These questions were designed to clarify physicians' opinions of and interactions with the ACI-TIPI. We hypothesized that we would see an effect of the tutorial on physicians' reported understanding of the ACI-TIPI and on their perceptions about its value. The results to be reported indicate that even though the tutorial was extremely brief, it did change both reported understanding and perceived usefulness of the ACI-TIPI. Note also that the interviewers administering the questionnaire were "blind" to whether the physicians had received the tutorial, so their expectations about the value of the tutorial should not have affected the interviewees' responses.

Significance testing for differences between the tutorial and nontutorial groups' responses was done by two-tailed t tests. Because of the small size of the two groups interviewed (8 physicians in each of the two conditions), effects of a substantial size may fall short of statistical significance at the traditional $p < .05$ level. We therefore report results as marginally significant at $p < .10$, indicating promising effects that are worth investigating with a larger sample (Rosenthal & Rosnow, 1984).

Recall of score. At the beginning of the questionnaire, participants were asked to recall the ACI-TIPI scores for the two cases they had just diagnosed. Tutorial participants had a better recall of the scores in the cases presented (overall recall rate of 75% for the tutorial group and 31% for the nontutorial group, respectively).

Influence of score. There was some evidence that the tutorial affected how often physicians reported taking the ACI-TIPI score into consideration

and the extent to which they believed the score influenced their decisions. Differences were at a marginal level of significance ($p < .10$) but were considered promising given the small sample size ($n = 16$) in this initial study. Both the frequency of consideration of the score (tutorial: mean = 3.1, $SD = 2.23$; nontutorial: mean = 1.4, $SD = 0.52$; $p = .064$) and extent of influence of the score on the decision (tutorial: mean = 2.4, $SD = 1.55$; nontutorial: mean = 1.2, $SD = 0.46$; $p = .070$) were found to be higher for the tutorial group. This finding supports our hypothesis that a tutorial may increase the "use" or extent to which the ACI-TIPI score is taken into consideration when making an ACI diagnosis. Also, the tutorial may have had the effect of permitting clinicians to acknowledge the influence of the ACI-TIPI score on their decision making.

Understanding of score. The tutorial significantly ($p < .05$) improved physicians' reported understanding of the ACI-TIPI score. There was a significant difference between the tutorial and nontutorial groups' ratings of their level of understanding both about how the score is derived (tutorial: mean = 4.2, $SD = 1.58$; nontutorial: mean = 2.6, $SD = 1.40$; $p = .041$) and how the score is to be used (tutorial: mean = 4.6, $SD = 1.92$; nontutorial: mean = 2.4, $SD = 1.76$; $p = .033$).

Perceived usefulness of score. The mean rating of the perceived usefulness of the score was significantly higher for the tutorial group (tutorial: mean = 3.9, $SD = 1.55$; nontutorial: mean = 2.2, $SD = 0.96$; $p = .023$). To further explore this result, we performed a correlation between the ratings of reported understanding and those of perceived usefulness. The correlation between understanding how the score is derived and the perceived usefulness of the score across all participants was found to be .82, whereas the correlation between understanding how to use the score and perceived usefulness was found to be .76. Thus the significant effect of the tutorial on understanding can be related to the perceived usefulness of the score. This supports the hypothesis that understanding is critical to perceived usefulness.

Confidence in score. Ratings of confidence in the score, defined as the perception of whether or not the score reflects the patient's condition correctly and reliably, was higher in the tutorial group (tutorial: mean = 3.7, $SD = 1.42$; nontutorial: mean = 2.5, $SD = 1.20$; $p = .098$).

Opinion of the Tutorial

The final part of the interview, just for the tutorial group, involved questions about the physicians' experience with the tutorial (Appendix C). Tutorial participants perceived the tutorial as increasing their level of understanding of the score (mean rating of 5.4 out of 7, $SD = 1.51$), confidence in the score (mean rating of 4.1 out of 7, $SD = 1.81$), and perceived usefulness of the score (mean rating of 4.9 out of 7, $SD = 1.13$). Ratings were also high for informativeness of content (mean rating of 6 out of 7, $SD = 1.07$) and clarity of presentation (mean rating of 6 out of 7, $SD = 1.07$). The most interesting results from this part of the interview were the high ratings and favorable comments given by physicians to the case study portion of the tutorial, which taught them by example how to incorporate the score into their decision-making process. The case studies were perceived as the most enjoyable and useful part of the tutorial.

DISCUSSION

The results of this study indicate that there is a potential for a relatively brief tutorial to increase acceptance and use of CDSSs for medical diagnosis. Many of the identified barriers to acceptance and use of the ACI-TIPI involved factors that could be addressed through training. What we have shown, as a first step, is that a very brief tutorial can change reported understanding and attitudes. If we had found that the tutorial had no effect on reported attitudes and understanding, then we would have concluded that training, at least in the brief form that is considered feasible for busy clinicians, is not likely to affect ACI-TPI use. Instead, we conclude that the short-term effects of brief training are promising enough to justify a more intensive, larger, longer term study that would measure the effects of training on actual long-term use.

Implications for Patient Safety

Informatics tools that are well designed and effectively used have tremendous potential for improving the quality of health care and enhancing patient safety. However, despite the strong evidence that some CDSSs can improve physician performance, these tools are not, on the whole, in

widespread use in clinical medicine. The ACI-TIPI, if widely used, could result in greater decision accuracy for possible ACI cases, if *accuracy* is defined as fewer admissions of patients who do not have ACI, along with a decrease (or no increase) in the number of genuine ACI patients who are not admitted. To achieve this patient safety benefit, however, the tool must be accepted and used.

This study showed that a very brief (10-min) tutorial produced a significant short-term change in physicians' reported attitudes toward the ACI-TIPI from 1 to 4 weeks after taking the tutorial. Whether this change in reported understanding and attitudes would result in a longer term increase in use of the ACI-TIPI remains to be demonstrated. It is especially encouraging, however, that the tutorial group had a more accurate recall of the ACI-TIPI scores in the standardized cases. This provides evidence, not based on self-report, that the tutorial made this group more aware of the ACI-TIPI score and therefore possibly more likely to take it into account in their decisions. These results indicate only that a tutorial may be effective in increasing acceptance and use of the ACI-TIPI; they do not establish that effectiveness. However, the fact that significant differences in attitude were detected up to 4 weeks after a 10-min exposure to a tutorial indicates that a very small investment in training might have a meaningful benefit in acceptance and use. This is especially important for a CDSS such as the ACI-TIPI, as in this case a small change in decision making can have substantial effects on patient safety and public health.

Design of Effective Training to Increase CDSS Use

Several different aspects of the study point to the importance of context, in the form of individual characteristics and experiences of the patient, on diagnostic decisions. Physicians in early interviews commented that they suspected clinicians did not know how to integrate the ACI-TIPI scores with information they were obtaining directly from the patient, such as reported history of pain or pressure, and that physicians received no training in this integration. Three case studies were incorporated into the tutorial specifically to address this concern, and these cases proved to be the part of the tutorial that was perceived as most valuable by participants. Finally, the diagnosis of standardized cases in which case information regarding reported symptoms, ECG results, and ACI-TIPI scores were systematically manipulated showed that physicians put a great deal of weight on the combination of ECG/score and symptoms together rather than on either one independently. The study suggests that the use of in-context case-based training in the interpretation and use of statistical risk scores has a high potential value for increasing CDSS use.

This study focused specifically on the ACI-TIPI CDSS for medical diagnosis of acute cardiac ischemia. It seems likely, however, that the results would generalize to other medical CDSSs as well. Furthermore, the barriers to use of the ACI-TIPI that were identified by physicians correspond quite closely to those identified across the board for CDSSs in a variety of domains.

These findings have the potential for far-reaching contributions in the identification of methods to help clinicians learn how to use statistical and probabilistic information to better assess risk. This will ultimately promote integration of CDSSs into medical decision making for improvement of patient safety and healthcare quality.

APPENDIX A: INITIAL INTERVIEW QUESTIONS FOR EXPERT CLINICIANS

1. Do you use the ACI-TIPI score in arriving at your treatment decision?
2. What do you perceive to be the value of the ACI-TIPI tool?
3. What are some concerns about the ACI-TIPI tool? (What may cause physicians to use or not use the tool?)
4. What would you recommend as strategies for increasing use of the tool?
5. What do you think of the role of a tutorial/educational approach for increasing effective use?
6. Do you have any ideas for the tutorial (format, content)?

   a. What do you think of the following ideas for a tutorial: review of ACI diagnosis, development and testing of tool, interactive calculator for the score?

   b. Where would the tutorial best fit in the clinical work flow, and how should it be administered?

APPENDIX B: ACI-TIPI EVALUATION QUESTIONNAIRE

1. a. Do you look at the ACI-TIPI score on the ECG?
   1 2 3 4 5 6 7
   never sometimes always

   b. Can you recall the ACI-TIPI scores on the two cases just presented?

2. a. Does the ACI-TIPI score typically agree with your independent assessment of a patient's condition?
   1 2 3 4 5 6 7
   never sometimes always

   b. If it differs, how does the ACI-TIPI score usually differ from your independent assessment? ACI-TIPI score is usually:
   Lower Higher

3. This three-part question asks about the influence of the ACI-TIPI score on your decisions when evaluating patients with chest pain in whether you think a patient needs to be admitted:

   a. How often do you take the ACI-TIPI score into consideration?
   1 2 3 4 5 6 7
   never sometimes always

   b. When you do take the ACI-TIPI score into consideration, what is the extent to which the score influences your decision?
   1 2 3 4 5 6 7
   not at all somewhat a great deal

   c. Have you ever reconsidered or even changed a decision because of an ACI-TIPI score – if so, how often?
   1 2 3 4 5 6 7
   never occasionally often

4. How, and under what circumstances or types of cases might you incorporate the ACI-TIPI score into your decision-making process? (Describe a specific example of where you might be more influenced by the ACI-TIPI score.)

5. What factors affect the influence the ACI-TIPI score has on your decisions?

6. Do you have any specific concerns about the ACI-TIPI score? If so, explain:

7. How would you rate your understanding of the ACI-TIPI score:

   a. Do you understand how the score is derived?
   1 2 3 4 5 6 7
   don't understand at all understand somewhat very much understand

   b. Do you understand how to use the score?
   1 2 3 4 5 6 7
   don't understand at all understand somewhat very much understand

   c. Is there anything that you would like to know about the ACI-TIPI score?

8. a. How would you rate your confidence in the ACI-TIPI score, i.e., how well do you think it reflects the patient's condition?
   1 2 3 4 5 6 7
   do not trust at all trust somewhat trust great deal

   b. Why do you select that confidence rating?

9. a. Overall, how would you rate the usefulness of the ACI-TIPI score?
   1 2 3 4 5 6 7
   not at all useful somewhat useful very useful

   b. Is there anything that might increase the usefulness of the ACI-TIPI score to you?

APPENDIX C: TUTORIAL SURVEY

1. How clear was the information in the tutorial?
   1 2 3 4 5 6 7
   not at all clear moderately clear very clear

2. Was the tutorial easy to navigate?
   1 2 3 4 5 6 7
   very confusing moderately easy very easy

3. a. Did you find the tutorial informative?
   1 2 3 4 5 6 7
   not at all somewhat very

   b. What part of the tutorial was most informative, i.e., what was the most significant thing you learned from the tutorial?

4. Did the tutorial address any concerns about the ACI-TIPI score? If so, which concerns?

5. Did the tutorial help in understanding how the ACI-TIPI score is derived?
   1 2 3 4 5 6 7
   not at all somewhat a great deal

6. Did the tutorial help in understanding how to use the ACI-TIPI score?
   1 2 3 4 5 6 7
   not at all somewhat a great deal

7. Did the tutorial increase your confidence in the ACI-TIPI score, i.e., how well you think it reflects the patient's condition?

1	2	3	4	5	6	7
not at all			somewhat			a great deal

8. Did the tutorial increase your perception of the usefulness of the ACI-TIPI score?

1	2	3	4	5	6	7
not at all			somewhat			a great deal

9. Do you have suggestions for improving the tutorial, and is there anything else you would like to see in the tutorial?

ACKNOWLEDGMENTS

This research was performed under a grant from the Agency for Healthcare Research and Quality (AHRQ), Contract 290-01-0009.

REFERENCES

Baxt, W. G., & Skora, J. (1996). Prospective validation of artificial neural network trained to identify acute myocardial infarction. *Lancet, 347*(8993), 12–15.

Berner, E. S., & Maisiak, R. S. (1999). Influence of case and physician characteristics on perceptions of decision support systems. *Journal of the American Medical Informatics Association, 6,* 428–434.

Christensen, C. R., & Hansen, A. J. (1987). Teaching and the case method. Boston, MA: Harvard Business School.

Goldman, L., Cook, F., & Brand, D. A. (1988). A computer protocol to predict myocardial infarction in emergency department patients with chest pain. *New England Journal of Medicine, 318,* 797–803.

Hunt, D. L., Haynes, R. B., Hanna, S. E., & Smith, K. (1998). Effects of computer-based clinical decision support systems on physicians' performance and patient outcomes: A systematic review. *Journal of the American Medical Association, 280,* 1339–1346.

Institute of Medicine, Committee on the Quality of Health Care. (2001). *Crossing the quality chasm: A new health system for the 21st century.* Washington, DC: National Academies Press. Available from National Academies Press Web site, http://www.nap.edu/catalog/10027.html

Karcz, A., Korn, R., Burke, M. C., Caggiano, R., Doyle, M. J., Erdos, M. J., et al. (1996). Malpractice claims against emergency physicians in Massachusetts: 1975–1993. *American Journal of Emergency Medicine, 14,* 341–345.

Lee, J. D., & Moray, N. (1994). Trust, self-confidence, and operators' adaptation to automation. *International Journal of Human-Computer Studies, 40,* 153–184.

Masalonis, A. J., & Parasuraman, R. (1999). Trust as a construct for evaluation of automated aids: Past and future theory and research. In *Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting* (pp. 184–188). Santa Monica, CA: Human Factors and Ergonomics Society.

McNair, M. P., & Hersum, A. C. (1954). *The case method at the Harvard Business School.* New York: McGraw-Hill.

McNutt, R. A., & Selker, H. P. (1988). How did the acute ischemic heart disease predictive instrument reduce unnecessary coronary care unit admissions? *Medical Decision Making, 8,* 90–94.

Pope, J. H., & Selker, H. P. (1999). Identifying acute cardiac ischemia in the emergency department. In C. P. Cannon (Ed.), *Contemporary cardiology: Management of acute coronary syndromes* (pp. 111–133). Totowa, NJ: Humana Press.

Pozen, M. W., D'Agostino, R. B., & Mitchell, J. B. (1980). The usefulness of a predictive instrument to reduce inappropriate admissions to the coronary care unit. *Annals of Internal Medicine, 92,* 238–242.

Pozen, M. W., D'Agostino, R. B., Selker, H. P., Sytkowski, P. A., & Hood, W. B. Jr. (1984). A predictive instrument to improve coronary care unit admission practices in acute ischemic heart disease: A prospective multicenter clinical trial. *New England Journal of Medicine, 310,* 1273-1278.

Rosenthal, R. R., & Rosnow, R. L. (1984). *Essentials of behavioral research.* New York: McGraw-Hill.

Selker, H. P., Beshansky, J. R., Griffith, J. L., Aufderheide, T. P., Ballin, D. S., Bernard, S. A., et al. (1998). The use of the Acute Cardiac Ischemia Time-Insensitive Predictive Instrument (ACI-TIPI) to assist emergency department triage of patients with chest pain or other symptoms suggestive of acute cardiac ischemia: A multicenter controlled clinical trial. *Annals of Internal Medicine, 129,* 845–855.

Selker, H. P., Griffith J. L., Beshansky, J. R., Califf, R. M., D'Agostino, R. B., Laks, M. M., et al. (1992). *The Thrombolytic Predictive Instrument (TPI) project: Combining clinical study databases to take medical effectiveness research to the streets.* Washington, DC: Department of Health and Human Services, Agency for Health Care Policy Research.

Selker, H. P., Griffith, J. L., &, D'Agostino, R. B. (1991). A tool for judging coronary care unit admission appropriateness valid for both real-time and retrospective use: A time-insensitive predictive instrument (TIPI) for acute cardiac ischemia: A multicenter study. *Medical Care, 29,* 610–627.

Selker, H. P., Pozen, M. W., & D'Agostino, R. B. (1985). Optimal identification of the patient with acute myocardial ischemia in the emergency room. In R. M. Califf & G. S. Wagner (Eds.), *Acute coronary care: Principles and practice* (pp. 289–298). Boston, MA: Martinus Nijhoff.

Tierney, W. M., Fitzgerald, J., & McHenry, R. (1986). Physician's estimate of the probability of myocardial infarction in emergency room patients with chest pain. *Medical Decision Making, 6,* 12–17.

Van de Does, E., Lubsen, J., & Pool, J. (1976). Acute coronary events in a general practice: Objectives and design of the Imminent Myocardial Infarction Rotterdam Study. *Heart Bulletin, 7,* 91–98.

Wyatt, J. (1989). Lessons learned from the field trial of ACORN, an expert system to advise on chest pain. In: B. Barber, D. Cao, & D. Qin, (Eds.), *Proceedings of the Sixth World Conference on Medical Informatics*, (pp. 111–115). Amsterdam: North Holland.

Wyatt, J., & Spiegelhalter, D. (1991). Field trials of medical decision-aids: Potential problems and solutions. In P. Clayton (Ed.), *Proceedings of the 15th Symposium on Computer Applications in Medical Care* (pp. 3–7). New York: McGraw Hill.

Zachary, W. W. (1988). Decision support systems: Designing to extend the cognitive limits. In M. Helander (Ed.), *Handbook of human-computer interaction* (pp. 997–1030). New York: Elsevier Science.

Fuji Lai is business area leader for Medical Systems at Aptima, Inc. She received an M.S. in biomedical and mechanical engineering from Harvard University in 1999.

Jean MacMillan is chief scientist at Aptima, Inc. She received a Ph.D. in cognitive psychology from Harvard University in 1987.

Denise H. Daudelin is project director and instructor in medicine at the Institute for Clinical Research and Health Policy Studies, Tufts-New England Medical Center. She received an M.P.H. in health services/health law from Boston University in 1989.

David M. Kent, a practicing general internist/hospitalist, is an assistant professor of medicine, clinical investigator, and associate director of Clinical Research Training at the Sackler Graduate School of Biomedical Studies, Tufts University. He received an M.D. in 1991 from the McGill University Faculty of Medicine.

Date received: January 23, 2004
Date accepted: June 13, 2005