

INNOVATIONS IN EDUCATION AND CLINICAL PRACTICE

A Web Exercise in Evidence-based Medicine Using Cognitive Theory

Farrell J. Lloyd, MD, MPH, Valerie F. Reyna, PhD

Our aim was to improve clinical reasoning skills by applying an established theory of memory, cognition, and decision making (fuzzy-trace theory) to instruction in evidence-based medicine. Decision-making tasks concerning chest pain evaluation in women were developed for medical students and internal medicine residents. The fuzzy-trace theory guided the selection of online sources (e.g., target articles) and decision-making tasks. Twelve students and 22 internal medicine residents attended didactic conferences emphasizing search, evaluation, and clinical application of relevant evidence. A 17-item Likert scale questionnaire assessed participants' evaluation of the instruction. Ratings for each of the 17 items differed significantly from chance in favor of this alternative approach to instruction. We conclude that fuzzy-trace theory may be a useful guide for developing learning exercises in evidence-based medicine.

KEY WORDS: evidence-based medicine; medical informatics; evaluation of chest pain in women; fuzzy-trace theory; sensitivity; specificity; positive predictive value; negative predictive value.

J GEN INTERN MED 2001;16:94-99.

In order to apply evidence-based medicine, clinicians must learn to access the research literature. Internet-based search engines such as MEDLINE, clinical databases such as the Cochrane Library Database,¹ and services such as MD Consult, as well as online textbooks and journals can augment the physician's "fund of knowledge." Therefore, advances in evidence-based research, together with unprecedented access to original sources of information, provide a unique challenge to medical decision-makers that may improve the quality of care provided to their patients—if they can find and correctly interpret this readily available information. These issues have been a focus of recent innovations in teaching evidence-based medicine at the University of Arizona Health Sciences Center.

OBJECTIVES

The aim of our curricular innovation project was to improve clinical reasoning skills by applying an established theory of memory, cognition, and decision making to instruction in evidence-based clinical medicine. (The theory—fuzzy-trace theory—is supported by scientific evidence, and its explanation of clinical reasoning in cardiovascular disease, as applied in this instructional intervention, has also been supported).^{2,3} The method of instruction incorporates elements of most evidence-based medicine workshops; learners use a computer search engine to access data, evaluate the evidence they find, and apply this reasoning to clinical cases. However, the facilitator's interventions differ from those used in traditional instruction, and rely on fuzzy-trace theory's explanation of pitfalls in clinical reasoning, especially difficulties in updating judgments of risk of disease once the results of diagnostic tests are known. That is, most interventions designed to improve higher order reasoning skills assume the computer metaphor of mind and accordingly emphasize precise, quantitative thinking (e.g., calculation of probabilities). In contrast, research on fuzzy-trace theory provides strategies for reasoning that avoid quantitative calculation while preserving quantitative accuracy (see below). For example, learners are taught to use specific graphical strategies to represent pretest and posttest probabilities (risk of disease before and after diagnostic testing) rather than to simply perform Bayesian calculations or use online Bayesian calculators, both of which are likely to increase calculation errors in practice. Bayesian calculation and use of online calculators are taught in addition to more intuitive methods of probability estimation. Thus, the instructor can anticipate, recognize, and manage difficulties in clinical reasoning by discouraging specific kinds of quantitative thinking (i.e., avoiding calculation of posttest probability of disease) and encouraging alternative qualitative strategies.

CONCEPTUAL FRAMEWORK

Instruction is designed for didactic ambulatory case conferences and internal medicine teaching rounds (i.e., for third-year clinical clerkships, fourth-year subinterns, medical interns, and residents) and focuses on the burgeoning literature on evaluation of women at risk for

Received from the Departments of Medicine (FJL, VFR) and Surgery (VFR), University of Arizona, Tucson, Ariz.

Address correspondence to Dr. Lloyd: Department of Medicine, University of Arizona, PO Box 245031, Tucson, AZ 85724-5031 (e-mail: lloyd@ahsc.arizona.edu).

ischemic heart disease.⁴⁻⁷ Much of this instruction emphasizes online searching to gather data (e.g., data about evidence-based clinical guidelines and subsequent commentaries on those guidelines) and application of these data to clinical scenarios. However, advances in medical knowledge and improved access to information are not sufficient to improve clinical decision making.^{8,9}

Recent developments in the cognitive sciences demonstrate that the process humans use to incorporate information into a "fund of knowledge" is not analogous to the information processing of a computer.^{10,11} Computers store information in a verbatim, or rote, form that can be most easily accessed by searching for exact input terms. If a computer has stored the term "acute cardiac ischemia," the most efficient means of retrieval will be to search for the verbatim term "acute cardiac ischemia." Humans, however, are likely to remember the imprecise gist of information that they have stored, such as recalling "acute cardiac ischemia" as "coronary artery disease" or some other related term. Naturally, expertise influences how information is remembered, but for both experts and novices, research has shown that humans remember the gist of information, and judgment and decision making are based on that gist representation.^{2,3,10}

These differences between human and computer information processing lead to predictable frustrations that learners have with reconciling computer functions and human reasoning. Although human-computer mismatches have been widely recognized, the traditional instructional response is to assist learners in becoming more sophisticated computational thinkers, i.e., to attempt to make the learners think like a computer. According to research on fuzzy-trace theory, this approach, paradoxically, promotes errors in specific judgment and decision-making tasks (such as posttest updating of probability judgments).^{2,3,11} Fuzzy-trace theory bridges the gap between humans and technology by anticipating how human cognitive function is assisted or hindered by the steps required to access information online and to apply it to a clinical scenario (and by providing guidelines for alternative strategies that better suit human thinking). A simple example is that learners can anticipate that search engines (e.g., MEDLINE) deliver very different results given superficially different search procedures or terms, and can learn to write down particularly arbitrary and counterintuitive procedures. For example, when using the terms "woman" and "chest pain" in a MEDLINE search, only one article will be located unless the box marked "map terms" is checked. A more complex example is that learners can anticipate that they will misinterpret results of Bayesian online calculators (and alternative visual information display tools should be used that allow learners to think in terms of the gist of information rather than exact quantities). Our instructional approach addresses these and other examples of human-computer mismatches and pitfalls in applications to clinical decision making that are anticipated by fuzzy-trace theory.

PROGRAM DESCRIPTION

This medical informatics and clinical decision making exercise is currently conducted with small groups of 6 to 12 medical residents and medical students. Participants meet in a learning technology classroom at the University of Arizona Health Sciences Center and each has access to a computer workstation with a direct network, Internet connection to the Arizona Health Sciences Library Web page. The library Web page has many electronic sources of information including MEDLINE access, electronic journals, textbooks, and the Science Citation Index. The instructor is at a computer workstation with a projected screen. Each of the tasks in the exercise is reviewed after completion. The instructor completes tasks with the students on a real-time basis, assisting individuals until all students have finished, and then reviewing correct procedures with the entire group. The instructor places many of the Web sites in the browser's history file or favorites/bookmarks. Participants are given a series of information accessing and related clinical decision-making tasks (e.g., to evaluate chest pain in a 49-year-old woman who is described in a detailed clinical scenario). Questions concerning Bayesian reasoning about the interpretation of diagnostic tests are also asked to elicit high-level clinical decision-making processes (e.g., what is the posttest probability of disease if the diagnostic test is positive or negative?).

The following describes a specific exercise that demonstrates decision-making processes used by clinicians and how they may be conveyed to learners within the evidence-based medicine framework while incorporating fuzzy-trace theory principles.

Sample Exercise

You have been asked to discuss the evaluation of chest pain in women for Morning Report. As you are preparing for the discussion, you remember a review article concerning evaluating chest pain in women that was published in the *New England Journal of Medicine* in the last few years.

1. Find the citation of the review article.
2. What other papers have cited this article and how many times has it been cited since it was published? Are there any updates or potential controversies?
3. Find the complete article and read the discussion concerning diagnostic testing. Do you agree with the recommendations for testing? Why do you agree or not agree?
4. Table 3 recommends a stress test for women who are at high risk of coronary artery disease. Consider a 49-year-old woman who presents to the emergency room after 1 hour of moderately severe shortness of breath and substernal chest pressure that began while she was painting the outside of her house. Significant past medical history is insulin-dependent diabetes for 12 years with evidence of

early retinopathy and nephropathy. At the time of her evaluation in the emergency room she is pain-free. Vital signs, physical exam, and electrocardiogram are all normal. An exercise treadmill test is performed. What is her pretest probability of disease? If the treadmill test has a sensitivity of 70% and specificity of 70%, what is the probability of clinically significant coronary artery disease if the test is:

- ◆ positive?
- ◆ negative?

5. Find a Web site that calculates this for you.

In addition, after participants find the online clinical calculator, the instructor uses a decision-making tool that visually represents the pretest probability (determined by the learner with the help of the review article if necessary), as well as the sensitivity and specificity of the exercise treadmill test (the test recommended by the referenced article). This visual representation was presented via an overhead projector using a 10-by-10 grid with 100 squares where each square represents a woman with potential ischemic heart disease (see Fig. 1). The students are each given a sheet of paper with 100 squares and are instructed to fill in the boxes according to their estimated pretest probability for a given case. They are also shown

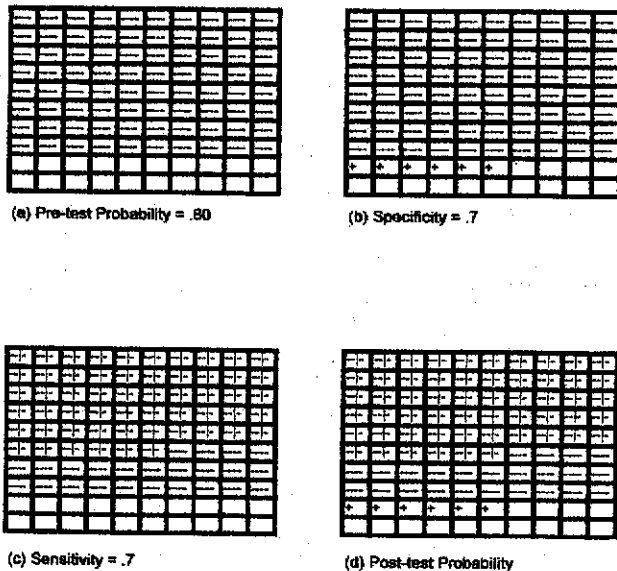


FIGURE 1. A visual representation for a patient at risk of ischemic heart disease. Women with disease and positive test (true positives): ---; women with disease and negative tests (false-negatives): ----; women without disease and a positive test (false-positives): +; women without disease and a negative test (true negatives): blank. (A) Pretest probability of a woman with ischemic heart disease. (B) Specificity of exercise treadmill test, assumed to be .7. (C) Sensitivity of exercise treadmill test, assumed to be .7. (D) Visual representation of posttest probabilities.

how to represent the sensitivity and the specificity of the exercise treadmill test. By convention, one color, usually blue, represents patients who may have ischemic heart disease (the pretest value). Participants are instructed to place one blue mark per square to indicate 1 patient per 100, or percentage, that has ischemic heart disease. For instance, if the pretest probability was determined to be 80%, then 80 squares are marked blue. Next, the sensitivity and the specificity of the diagnostic test are represented. Patients with a positive test are usually represented by the color red. Thus, if a woman with a pretest of 80% undergoes an exercise treadmill test with a sensitivity and specificity of 70%, 56 of the women with ischemic heart disease will have a positive test (marked red), and 24 of the women with ischemic heart disease will have a negative test. Of the 20 women without ischemic heart disease, 14 will have a negative test, and 6 will have a positive test (marked red). Therefore, by using 100 squares and 2 colors, the posttest probability is represented visually for the participants (Fig. 1).

EVALUATION AND FEEDBACK

The sessions allow for direct interaction with the learners while they are online and performing the decision-making tasks. The majority of participants accomplish the 6 tasks by the end of the hour-long session. There is one-on-one instruction by the instructor in accomplishing the computer-related tasks. Peers with more computer-related skills also assist students with fewer skills. At the conclusion of the exercise, an open discussion concerning the merits of the diagnostic strategy employed by the authors of the referenced article, in comparison to the learners' own calculation of posttest risk, occurs. Participants are given a 17-item questionnaire (see Tables 1 and 2) that employs a 5-part (1 = strongly disagree to 5 = strongly agree) Likert scale to assess participants' impression of the need for instruction concerning accessing and interpretation of online sources of information and their evaluation of this alternative approach to instruction in medical decision making. The exercise has a high rate of attendance and is frequently requested by residents and medical students.

From July 1999 to December 1999, 34 participants (12 medical students and 22 internal medicine residents) attended at least 1 of 6 conferences. Two participants attended two conferences; however, only the first questionnaire filled out by the participants was included in the analysis. Two residents were unable to complete the questionnaire. The data from the questionnaire were analyzed using the SPSS version 9 statistical package (SPSS for Windows: Release 9.0. Chicago: SPSS, Inc; 1998). A nonparametric test, a 1-sample Kolmogorov-Smirnov test, was conducted for each of the 17 items to determine whether ratings differed from what would be expected by chance (i.e., from a uniform distribution across ratings). Tests that assume a normal distribution around a rating of "3," i.e., no opinion, yielded similar results. Results for

Table 1. Participants' Ratings of Interactive Evidence-based Experience (Agree or Disagree)

Question	n	Mean (SD)*	Agree†, %	P Value‡
I learned new information concerning evaluating chest pain in women.	29	4.21 (0.77)	79	<.001
I found new tools to retrieve information concerning evaluating chest pain in women.	29	4.31 (0.85)	90	<.001
I learned improved methods to determine posttest probability of disease, e.g., the probability of coronary artery disease after a positive or negative test.	29	4.14 (0.69)	83	.004
My clinical knowledge would improve if I attended regular interactive conferences similar to this exercise.	29	4.55 (0.51)	100	<.001
My ability to retrieve evidence-based medical information would improve if I attended regular interactive conferences similar to this exercise.	29	4.41 (0.57)	97	<.001
My ability to interpret diagnostic tests would improve if I attended regular interactive conferences similar to this exercise.	29	4.48 (0.57)	97	<.001
I need no further training in finding clinical information available online.	30	1.6 (1.00)	7	<.001
I am not likely to use computers in acquiring clinical information.	30	1.53 (1.11)	7	<.001
I am not likely to use computers in retrieving clinical information.	30	1.4 (0.86)	3	<.001
I am not likely to use computers to improve my ability to interpret diagnostic tests.	30	1.57 (0.94)	7	<.001
I am not likely to attend conferences such as this again.	30	1.5 (1.04)	7	<.001
I am likely to learn more with routine, i.e., noninteractive case conference.	29	2.55 (1.12)	21	.003
Interpretation of diagnostic tests is better taught in routine case conference.	29	2.76 (1.09)	24	<.041
The medical literature is more likely to be accessible after routine case conference.	29	2.86 (1.19)	24	<.041

* Mean rating on a 5-point Likert scale (1 = strongly disagree, 2 = disagree, 3 = no opinion, 4 = agree, 5 = strongly agree).

† Percentage responding agree or strongly agree (4 or 5).

‡ One-Sample Kolmogorov-Smirnov Test P value.

each of the 17 items were statistically significant at $P < .05$. The distribution of ratings was not uniform: A significant proportion of participants "agreed" or "strongly agreed" with items that favored the alternative instruction and "disagreed" or "strongly disagreed" with items that did not favor the alternative instruction. The pattern of ratings indicated a high level of endorsement that this exercise improved participants' clinical knowledge, information retrieval skills, and the ability to interpret diagnostic tests (Table 1). Furthermore, participants indicated that they were significantly "more likely" or "very likely" to use online tools for accessing evidence-based research and to apply improved reasoning processes to such evidence (Table 2).

In addition to structured questionnaires, students were able to verbalize concerns with the tasks openly during the sessions. Frequently expressed concerns included difficulty in accessing information using the most intuitive search terms (e.g., "women" and "chest pain") and difficulty interpreting recommendations for posttest probability of disease, both of which were addressed by the instructor during the exercises (see Table 3 for comments).

Since the tasks in this exercise involve the ability to locate specific information with Web-based tools (and databases) and to interpret the literature, learners of all skill levels were appropriately challenged. All participants were instructed regarding postdiagnostic test updating of probabilities using published clinical recommendations.⁴

As predicted by fuzzy-trace theory, when participants used the online clinical calculator, they often missed the question, "What is the probability of disease if the test is negative?" The observed response given most often is the calculated value of the negative predictive value. (If we assume the pretest probability is 80%, then the negative predictive value is 37%). When the visual representation is given, the participants are usually able to see that the probability of ischemic heart disease remains very high even if the test is negative. (Using the same example, of the 38 women with a negative exercise treadmill test, 24 of these women, or 63%, will have ischemic heart disease; see Fig. 1). Although the exact pretest value in the given example is open to debate, the patient is clearly high risk. We illustrated a pretest probability of ischemic heart

Table 2. Participants' Ratings to Benefits from Interactive Evidence-based Exercise (Likely or Not)

Question	n	Mean (SD)*	Likely†, %	P Value‡
As a result of this exercise, how likely are you to augment your knowledge base of clinical problems such as evaluating chest pain in women with tools you used today?	29	4.03 (0.50)	90	<.001
As a result of this exercise, how likely are you to retrieve clinically useful information for clinical decision making?	29	4.13 (0.64)	86	.001
As a result of this exercise, how likely are you to improve the process you use to interpret diagnostic tests like exercise treadmill tests performed on women at risk for unstable angina?	29	3.79 (0.73)	69	.001

* Mean rating on a 5-point Likert scale (1 = not likely, 2 = less likely, 3 = likely as not, 4 = more likely, 5 = very likely).

† Percentage responding more likely or very likely (4 or 5).

‡ One-Sample Kolmogorov-Smirnov Test P value.

Table 3. Sample of Participants' Comments Concerning the Interactive Evidence-based Exercise

MEDLINE search and online journals

- If I limit the search too early, I am less likely to find the articles others found.
- The more specific the search, the less successful I am.
- The Windows-based version of MEDLINE is much more intuitive than the text-only version.
- Online electronic journals will help me manage patients with data at the patients' bedside.
- Downloading Adobe Acrobat articles makes the tables in journal articles easier to read.

Web of Science: Science Citation Index

- Finding relevant articles that impact patient care is easier since the session.
- This will be helpful in preparing for teaching conferences during rounds, morning report, and journal club.
- The authors of the "Evaluating Chest Pain in Women" review article from the *New England Journal of Medicine* have an extensive publication history in cardiology journals.

Clinical recommendations for posttest evaluation and treatment

- For high-risk patients with a negative test, the posttest probability of cardiac ischemia is still high (>50%).
- The published clinical recommendations do not agree with the posttest probability.
- The Website that calculates the posttest prediction is confusing because it uses terms like "prevalence" and abbreviations.
- I am not sure I understand what the calculated numbers mean from the online Bayesian calculator.
- The visual representation of post-test risk is more understandable than the online calculator because I intuitively see what the numbers mean.
- I would like to be able to calculate posttest clinical risk accurately, especially for board tests.
- I rarely calculate posttest values during routine patient care. I just use my clinical judgment to decide posttest risk.

disease of 80% in instruction because that is at the lower margins of pretest probability for a high-risk patient. The actual risk is likely higher; therefore, the probability of ischemic heart disease given a negative exercise treadmill test would be even higher than 63%.

The referenced article on clinical guidelines for evaluation of women with chest pain also illustrates similar problems in human reasoning, predicted by fuzzy-trace theory, concerning anticipating the effects of pretest probability on subsequent interpretation of diagnostic tests. That is, if pretest probability is sufficiently high, catheterization is indicated as a diagnostic test because a negative result of other, less accurate diagnostic tests would, nevertheless, yield a high posttest probability of disease (as calculated using Bayes theorem). Recommendations in the referenced article suggest, contrary to the calculation of posttest probability, that high-risk patients with a negative result need not receive any further testing.⁴ Thus, this highly cited article provides a specific example of difficulty in judging posttest probability. The article illustrates that, despite excellent up-to-date knowledge of the risk factors for coronary artery disease and the variety of presenting symptoms women experience, medical decision-makers often do not perform the optimal diagnostic tests when indicated because of predictable difficulties in anticipating the effects of pretest probability on interpretation of diagnostic tests.³ Learners judge posttest probabilities intuitively, using the visual representation, and realize that their judgments fail to match those of the referenced article.

DISCUSSION

This approach to instruction differs from traditional methods of clinical education in three ways:

1. Technology is not the focus of instruction, but merely a means to acquire information;

2. Instruction is aimed at improving clinical decision making, rather than simply acquiring new knowledge; and
3. Instruction is informed by an empirically supported theory of cognition that predicts points of incompatibility between humans and technology, such as the effect of counterintuitive procedures in popular search engines and specific difficulties in understanding posttest probability.

This medical informatics decision-making exercise requires learners to use search engines, access, and evaluate evidence-based research, and engage in clinical decision making (e.g., make recommendations based on published evidence and posttest probability of disease). The instructor uses fuzzy-trace theory to anticipate tasks that may be particularly difficult for the learners, including the cognitive errors expected in posttest updating of probability estimates, so that these receive appropriate emphasis. Fuzzy-trace theory was also used to break down the sequence of decision making into realistic stages and to select appropriate learning interventions at each stage: At the beginning, learners find information on the appropriate pretest probability from the medical literature, as well as the referenced article's citation history. Later, they use this information to perform calculations of posttest probability using an online calculator. After experiencing predictable difficulties, they are shown an intuitive visual representation that is designed to provide insight into correct posttest updating of probabilities. Finally, recommendations for diagnostic testing in the referenced article (selected to illustrate difficulties with posttest updating, as well as evidence-based research) are compared to the learners' visual representation of risk using the decision-making tool. We emphasize probability estimation, coordination, and updating because of their obvious clinical relevance, but also because our theoretic-

tical perspective predicts that they will be sources of human error.

Fuzzy-trace theory also allows us to anticipate events that occur during the sessions such as "verbatim interference in gist tasks."² A gist task involves meaningful interpretation of data. Verbatim interference occurs when arbitrary, superficial details or procedures make it more difficult to accomplish such meaningful interpretation of data. For example, research has shown that focusing on quantitative details and on performing calculations interferes with accurate estimation of posttest probability. Decision makers perform better at estimating posttest probability when they ignore details about exact probabilities and focus instead on relations among classes (e.g., there are many women with positive test results, but few of them have the disease), as captured in the visual representation. This tendency to perform better at meaningful tasks when representing information as gist, rather than as arbitrary verbatim details, also explains a variety of human-computer mismatches. For example, search engines operate at a strictly verbatim level, which can be frustrating to the human user of the search engine who operates at mainly a gist level.

Although participants' self-reports indicate high levels of endorsement of the teaching intervention, which is grounded in the published literature in cognitive sciences, evidence of effectiveness with more learners would be desirable. In addition, behavioral evidence of improved clinical decision making (and, ultimately, improved patient outcomes) that corroborates learners' ratings is needed. Behavioral outcomes should include measures of efficient accessing of evidence-based research and appropriate clinical application of such research.

Because medical decision making is the basis of most tasks in internal medicine, a structured approach to teaching decision making is imperative. Fuzzy-trace theory provides a framework for such a structured approach, one in which instructional interventions can be developed that address a variety of human-computer mismatches and pitfalls in applications to clinical decision making. According to fuzzy-trace theory, the availability of information from online sources can augment the knowledge base of the decision maker, but information alone is not sufficient to improve decision making. Instead, research indicates that focusing simply on acquiring facts and performing calculations can interfere with important clinical judgments, such

as posttest updating of disease probability. Although demonstrations of these interference effects have been documented in research studies, further work is needed to establish whether fuzzy-trace theory can be used to improve medical education and the quality of clinical decision making.

Preparation of this manuscript was supported in part by grants to both authors from the Academic Medicine and Managed Care Forum (Integrated Therapeutics Group/Schering-Plough, Inc.) and the National Science Foundation (SBR973-0143) and grants to the second author from the National Institutes of Health (NIH 316201 and P50AT00008) and the United States Department of Commerce (SUB005).

We gratefully acknowledge the assistance of Rachael Anderson and David Piper in the Arizona Health Sciences Center Library.

REFERENCES

1. Cochrane Library Database [database online]. The Cochrane Collaboration; 1998. Updated January 1999.
2. Reyna VF, Brainerd CJ. Fuzzy-trace theory: an interim synthesis. *Learn Individ Differ*. 1995;5:1-75.
3. Reyna VF, Lloyd FJ, Brainerd CJ. Memory, development, and rationality: an integrative theory of judgment and decision-making. In: Schneider S, Shanteau J, eds. *Emerging Perspectives in Judgment and Decision Making*. New York, NY: Cambridge University Press. In press.
4. Douglas PS, Ginsburg GS. Evaluating chest pain in women. *N Engl J Med*. 1996;334:1311-5.
5. Rich-Edwards JW, Manson JE, Hennekens CH, Buring JE. The primary prevention of coronary heart disease in women. *N Engl J Med*. 1995;332:1758-66.
6. Willich SN, Lowel H, Lewis M, Arntz R, Schubert F, Schroder R. Unexplained gender differences in clinical symptoms of acute myocardial infarction. *J Am Coll Cardiol*. 1993;21:238A. Abstract.
7. Weiner DA, Ryan TJ, McCabe CH, et al. Exercise stress testing, correlations among history of angina, ST-segment response and prevalence of coronary-artery disease in the Coronary Artery Surgery Study (CASS). *N Engl J Med*. 1979;301:230-5.
8. McGettigan P, Sly K, O'Connell D, Hill S, Henry D. The effects of information framing on the practices of physicians. *J Gen Intern Med*. 1999;14:633-42.
9. Dawson NV, Arkes HR. Systematic errors in medical decision-making: judgment limitations. *J Gen Intern Med*. 1987;2:183-7.
10. Reyna VF, Brainerd CJ. Fuzzy-trace theory and framing effects in choice: gist extraction, truncation, and conversion. *J Behav Decis Making*. 1991;4:249-62.
11. Reyna VF, Brainerd CJ. Fuzzy memory and mathematics in the classroom. In: Davies GM, Logie RH, eds. *Memory in Everyday Life*. Amsterdam: North Holland Press; 1993:91-119.