Patients Derogate Physicians Who Use a Computer-Assisted Diagnostic Aid

Hal R. Arkes, PhD, Victoria A. Shaffer, PhD, Mitchell A. Medow, PhD, MD

Objective. *To ascertain whether a physician who uses a computer-assisted diagnostic support system (DSS) would be rated less capable than a physician who does not.* ***Method***. *Students assumed the role of a patient with a possible ankle fracture (experiment 1) or a possible deep vein thrombosis (experiment 2). They read a scenario that described an interaction with a physician who used no DSS, one who used an unspecified DSS, or one who used a DSS developed at a prestigious medical center. Participants were then asked to rate the interaction on 5 criteria, the most important of which was the diagnostic ability of the physician. In experiment 3, 74 patients in the waiting room of a clinic were randomly assigned to the same 3 types of groups as used in experiment 1. In experiment 4, 131 3rd- and 4th-year medical students read a scenario of a physician-patient interaction and were randomly assigned to 1 of 4 groups: the physician used no DSS, heeded the recommendation of a DSS, defied a recommendation of a DSS by treating in a less aggressive manner, or defied a recommendation of a DSS by treating in a more aggressive manner.* ***Results***. *The participants always deemed the physician who used no decision aid to have the highest diagnostic ability.* ***Conclusion***. *Patients may surmise that a physician who uses a DSS is not as capable as a physician who makes the diagnosis with no assistance from a DSS.* ***Key words:*** *decision support techniques; diagnosis computer assisted; patient satisfaction.* ***(Med Decis Making 2007;27:189–202)***

A large number of computer-based diagnostic support systems (DSSs) have been developed during the past 30 years.[1] DSSs have been heralded as offering a way to decrease various types of errors,[2] foster the implementation of evidence-based medicine,[3] and reduce inappropriate admissions and costs.[4]

There are strong a priori reasons for believing that computer-based DSSs would improve diagnostic accuracy. Dawes and his colleagues[5] reviewed approximately 100 studies, the overwhelming majority of which showed that actuarial predictions were superior to the unaided diagnostician. Actuarial judgments are characterized by combining evidence using a formula such as a regression equation. Of course, combining pieces of evidence in such a manner is easily done with a simple computer program. As a prototypic actuarial method, a computer-based DSS would thus enjoy an advantage over unaided diagnosticians, consistent with the evidence summarized by Dawes and colleagues.[5]

The research reviewed by Dawes and colleagues[5] pertained to a wide variety of domains with a substantial number emanating from the domain of psychology. However, specifically within the domain of medicine, a large number of studies have found that computer-based DSSs perform better than physicians in a wide variety of diagnostic contexts.[6–11] In an early study, de Dombal and colleagues[7] showed that the use of a computer program resulted in significantly more accurate diagnoses of acute appendicitis. Similarly, Corey and Merenstein[6] showed that use of a predictive index for acute cardiac ischemia resulted in far more accurate classification of patients than occurred when

Received 10 February 2005 from the Center for Health Outcomes, Policy, & Evaluation Studies and the Department of Psychology, The Ohio State University, Columbus, Ohio (HRA); the Department of Psychology, Wichita State University, Wichita, Kansas (VAS); and the Division of General Internal Medicine, College of Medicine, The Ohio State University, Columbus, Ohio (MAM). Financial support for this study was provided by a grant from the Program in Decision, Risk, and Management Science at the National Science Foundation. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report. Revision accepted for publication 6 September 2006.

Address correspondence to Hal R. Arkes, PhD, Center for Health Outcomes, Policy, and Evaluation Studies, The Ohio State University, 466 Cunz Hall, 1841 Millikin Road, Columbus, OH 43210-1229; e-mail: arkes.1@osu.edu.

DOI: 10.1177/0272989X06297391

physicians did not use the index. Ridderikhoff and van Herk[8] found that a DSS used in a general practice more than doubled the diagnostic accuracy of unaided physicians. In one of the experiments to be presented in this article, we presented respondents with a scenario describing a sore throat. Several prediction rules have been developed to help physicians diagnose strep throat.[12–14] Another of our scenarios pertains to a possible ankle fracture. A well-known scoring rule has been developed in this domain as well.[15]

Reviews by Kaplan[1] and Hunt and colleagues[16] confirm that although many studies verify the superiority of DSSs in the diagnostic process, some studies do not.[17] However, there is unanimity with regard to 1 characteristic of DSSs: they are grossly underused.[1] To cite 1 example of underutilization, the acute ischemic heart disease predictive instrument put in place by Corey and Merenstein[6] reduced the false-positive diagnosis rate from 71% to 0%. Following the use of the aid in randomized controlled trials, physicians were free to use the aid or not. Utilization during this latter phase was only 2.8%. Other examples of underutilization abound.[18–22]

A number of reasons have been proposed to explain why diagnosticians might not want to use computer-assisted DSSs or any other actuarial system such as a practice guideline.[20,23,24] These reasons include the reluctance to use "cookbook medicine," inertia of previous practice, and lack of agreement with the output of the DSS.[19,20] One additional and potentially important reason why physicians might be reluctant to use computer-based DSSs was suggested in a short report by Cruikshank.[25] Using only 3 British physicians in his study, Cruikshank[25] assessed patients' opinions of their physicians before and after a computer-assisted diagnostic tool was introduced. Even though such a decision aid was designed to reduce diagnostic errors, patients were markedly less positive about the thoroughness, cleverness, decisiveness, and thoughtfulness of their physician after the decision aid was introduced. A physician who values his or her reputation may be understandably reluctant to use a computer-based DSS if its use has this detrimental effect. However, this study was published in 1984, when the use of computers was drastically less prevalent than it is today. Therefore, the wariness patients might have had concerning the use of a novel machine might not be present in today's computer-saturated environment. Note that we are not investigating patients' attitudes toward physicians' use of computerized medical records, which have generally not been negative.[26,27] Instead, we are specifically interested in whether patients would derogate the diagnostic ability of a physician who uses a computer-based DSS compared to the rating given by a patient to a physician who does not use such an aid.

EXPERIMENT 1

Method

Participants

In total, 347 undergraduate students at The Ohio State University participated in the experiment in partial fulfillment of a course requirement. The participants were randomly assigned to 1 of 3 experimental groups: no aid (n = 108), aid (n = 127), and prestigious aid (n = 112).

Materials

The experiment employed 3 short scenarios that described an interaction between a doctor and her patient, a brief questionnaire containing the target-dependent variables, and a numeracy scale that contained 3 questions. The questions were the "General Numeracy Items" from Lipkus and colleagues,[28] which were in turn adapted from Schwartz and colleagues.[29] We hypothesized that facility with numerical concepts might foster greater acceptance of DSSs.

All stimulus materials were presented, and all data were collected via computer.

Procedure

Each participant read a 1-page scenario in which he or she was asked to take the role of the patient. In the scenario, the patient had an ankle injury and went to see a primary care physician, who provided a diagnosis. In the course of the diagnostic process, the doctor used no decision aid, an unspecified decision aid, or a prestigious decision aid. After reading the scenario, the participants were asked to rate the following 5 criteria: thoroughness of examination, length of wait, diagnostic ability of the physician, professionalism of the physician, and overall satisfaction with the examination. Participants were asked to respond to each question separately using 5 different Likert-type scales; each scale ranged from 1 to 7. The more negative evaluations corresponded to the left side of the scale and the more positive evaluations with the right. Each point on the scale was given both a numeric value and a descriptive label. After they responded to the dependent variables, the participants were asked to respond to the numeracy scale (see Appendix A for the scenarios and numeracy scale).

Table 1 Experiment 1 Group Means

Dependent Variable	Group		
	No Aid	Aid	Prestigious Aid
Diagnostic ability	4.69[a]	3.70[b]	3.63[b]
Overall satisfaction	4.71[a]	4.23[b]	4.15[b]
Professionalism	5.20[a]	4.60[b]	4.50[b]
Thoroughness of examination	5.27[a]	4.80[ac]	4.70[bc]
Length of wait	3.17[a]	3.37[a]	3.18[a]

Note: Within a row, means not sharing a common superscript differ significantly by a Kruskal-Wallis test ($P < 0.05$).

Results

A numeracy score was calculated for each participant by tallying the number of correct items from the 3-item scale. The average numeracy score was 1.83.

For all dependent variables except "length of wait," the physician who used no decision aid was given the highest evaluation. See Table 1 for the means of all dependent variables. Significant Pearson correlations existed between the 5 dependent variables, ranging from .302 to .781.

The 5 dependent variables were subjected to a multivariate analysis of covariance (MANCOVA); the covariate was numeracy score, and the single factor was experimental condition: no aid, aid, and prestigious aid. The numeracy score was not a significant covariate for any of the 5 dependent variables. Numeracy data were also collected in the subsequent 3 experiments; however, numeracy was not a significant covariate in any study. Therefore, the numeracy data will not be addressed again in this article.

Significant *F* values were obtained for the factor of experimental condition in 4 of the 5 dependent variables: thoroughness of examination, diagnostic ability, professionalism, and overall satisfaction. The evaluation of a patient's length of wait did not differ significantly between the 3 experimental groups. However, responses in the no-aid condition were significantly less variable for 4 of the dependent variables (thoroughness of examination, diagnostic ability, professionalism, and overall satisfaction). Therefore, the parametric analysis of variance (ANOVA) was rejected in favor of the nonparametric Kruskal-Wallis test, which provides an omnibus test for differences in ranks. Kruskal-Wallis tests yielded significant differences ($P < 0.05$) in thoroughness of examination, $\chi^2(2) = 6.70$; diagnostic ability, $\chi^2(2) = 33.90$; professionalism, $\chi^2(2) = 15.82$; and overall satisfaction, $\chi^2(2) = 8.86$. Subsequent post hoc comparisons indicated that the no-aid group was rated significantly higher than the prestigious-aid group for all 4 dependent variables ($P < 0.05$). For 3 of these 4 variables, the no-aid group was rated significantly higher than the aid group ($P < 0.05$), with "thoroughness of the examination" not reaching significance. In contrast, the aid and prestigious-aid groups did not differ significantly on any of the 4 dependent variables.

Discussion

The data from our 1st experiment suggest that patients' perceptions of their physicians are influenced by the use of decision aids during routine visits. The participants' evaluation of the physician's diagnostic ability, the thoroughness of the examination, the professionalism of the physician, and their overall evaluation of the physician were affected by our experimental conditions. The physician who did not use a decision aid was rated more favorably on all 4 dimensions by our participants than the physicians who used an unspecified decision aid or a highly prestigious decision aid. Numeracy scores, found to be a factor in other studies about medical judgments,[28,29] did not significantly affect the evaluations given in this experiment.

Based on the results of Cruikshank,[25] we expected group differences on the dependent variable pertaining to the diagnostic ability of the physician. We did not anticipate that the use of a computer-assisted DSS would also influence such factors as the thoroughness of the examination because the groups did not differ in the amount of data collected by the physician or offered by the patient. Nevertheless, group differences did occur on 4 of the 5 factors.

Perhaps our conclusions were specific to the type of injury; maybe participants believed that a potential ankle fracture was such a common occurrence that physicians should be able to treat the injury without the use of a diagnostic aid. Therefore, we decided to replicate our findings with a 2nd, but potentially more serious, medical condition—deep vein thrombosis (DVT). Because DVT is potentially more dangerous than an ankle fracture, we thought that it would be advisable to ascertain if the same general pattern of findings would occur in this new situation. In addition, because the use of a "prestigious" aid did not affect diagnostic ability, only 2 groups (no aid and aid) were used with this population.

EXPERIMENT 2

Method

Participants

In total, 128 undergraduate students at The Ohio State University participated in the experiment in partial fulfillment of a course requirement. The participants were randomly assigned to 1 of 2 experimental groups: no aid ($n = 55$) and aid ($n = 73$).

Materials

The experiment employed 2 short scenarios that described an interaction between a doctor and her patient and a brief questionnaire containing the target dependent variables. All stimulus materials were presented, and all data were collected via computer.

Procedure

Each participant read a 1-page scenario in which he or she was asked to take the role of the patient. In the scenario, the patient had a leg injury and was at risk for DVT. The patient went to see a primary care physician, who provided a diagnosis. In the course of the diagnostic process, the doctor used either no decision aid or an unspecified decision aid (see Appendix B for the text of the scenario). After reading the scenario, the participants were asked to rate the following 6 criteria: perceived difficulty of diagnosis, thoroughness of examination, length of wait, diagnostic ability of the physician, professionalism of the physician, and overall satisfaction with the examination. Participants were asked to respond to each question separately using 6 different Likert-type scales constructed in a manner identical to those in experiment 1.

Results

Responses to diagnostic ability and thoroughness of examination were significantly more variable in the aid condition. Therefore, the appropriate statistical correction in the degrees of freedom was made for t tests involving these variables. In addition, significant positive Pearson correlations existed between 5 of the dependent variables (thoroughness of examination, length of wait, diagnostic ability of the physician, professionalism of the physician, and overall satisfaction with the examination); correlations ranged from .342 to .820. Perceived difficulty of diagnosis was negatively correlated with professionalism ($r = -.202$, $P < 0.05$). See Table 2 for the means of all dependent variables.

Table 2 Experiment 2 Group Means

	Group	
Dependent Variable	**No Aid**	**Aid**
Diagnostic ability[a]	4.91	4.05
Perceived difficulty	3.98	3.88
Overall satisfaction	4.64	4.15
Professionalism[a]	5.44	4.78
Thoroughness of examination[a]	5.55	5.11
Length of wait	3.02	3.00

a. According to a t test, the 2 groups differed on this dependent variable ($P < 0.05$).

Again, participants in the aid condition gave significantly lower ratings to the target dependent variable, diagnostic ability, $t(125.93) = 3.53$, $P < 0.01$, Cohen $d = .54$. In addition, participants in the aid condition also deemed the physicians to be less professional, $t(126) = 2.62$, $P < 0.01$, Cohen $d = .43$, and less thorough, $t(125.95) = 1.743$, $P < 0.05$, Cohen $d = .27$. There were no significant differences between the no-aid and aid conditions on perceived difficulty, overall satisfaction, and length of wait.

Discussion

The data from experiment 2 replicated the findings from experiment 1 using a more serious medical condition. The evaluation of a physician was again influenced by the use of a diagnostic aid.

To improve the generality of our conclusions, we decided to replicate our study with a nonstudent population and a scenario that used a different malady. We wanted to ascertain whether the results found in experiment 1 would generalize to patients who were seeking medical care.

EXPERIMENT 3

Method

Participants

In total, 74 patients from the University Health Connections clinic, an urgent and primary care service for faculty and staff at The Ohio State University, participated in this experiment. Each patient was paid $25 for participation. The participants were randomly assigned to 1 of 3 experimental groups: no aid ($n = 24$), aid ($n = 27$), and prestigious aid ($n = 22$).

Materials

The experiment employed 3 short scenarios that described an interaction between a doctor and his patient and a brief questionnaire containing the target dependent variables. All data were collected using paper-and-pencil materials in the lobby of the University Health Connections clinic. Appendix C contains the scenarios.

Procedure

When participants checked in at the receptionist desk at the University Health Connections clinic to receive medical treatment, they were approached by either a receptionist or a researcher and asked to participate in the experiment. Those patients who gave their consent were given a packet that contained the patient-doctor scenario and the questionnaire containing the 5 dependent variables. The dependent variables (thoroughness of examination, evaluation of length of wait, diagnostic ability, professionalism, and overall satisfaction with the examination) were identical to those in experiment 1. However, the scenario in experiment 1, which described an ankle injury, was replaced with a scenario in which the patient had a persistent cough. Again, the participant was asked to take the role of the patient who went to see his or her primary care physician and was subsequently provided with a diagnosis. In the course of the diagnostic process, the doctor employed no decision aid, an unspecified decision aid, or a prestigious decision aid. After reading the scenario, the participants were asked to rate the physician on the 5 criteria. Participants were asked to respond to each question separately using 5 different Likert-type scales constructed in the same manner as those used in the prior 2 studies.

Results

The 5 dependent variables (thoroughness of examination, evaluation of length of wait, diagnostic ability, professionalism, and overall satisfaction) were subjected to a MANOVA with the single independent variable of experimental condition: no aid, aid, and prestigious aid. See Table 3 for the means of all variables. A significant *F* value was obtained for the target variable, diagnostic ability, $F(2, 68) = 4.94$, $P < 0.05$, $\eta^2 = .12$. Tukey post hoc tests revealed significant differences between the aid and no-aid groups, $P < 0.05$. Patients at the University Health Connections clinic awarded the highest rating for diagnostic ability to the physician who used no decision aid. There were no significant differences in experimental condition for the remaining 4 dependent variables: thoroughness of examination, evaluation of length of wait, professionalism, and overall satisfaction. Significant Pearson correlations were obtained between pairs of dependent variables ranging from .261 to .862.

Table 3 Experiment 3 Group Means

	Group		
Dependent Variable	No Aid	Aid	Prestigious Aid
Diagnostic ability[a]	4.71	3.52	4.41
Overall satisfaction	5.00	4.19	4.64
Professionalism	4.79	4.04	4.55
Thoroughness of examination	5.04	5.22	5.50
Length of wait	3.46	3.44	3.58

a. According to a Tukey post hoc test, the aid and no-aid groups differed significantly on this dependent variable ($P < 0.05$).

Discussion

The diagnostic ability of the physician was again rated significantly higher when she used no computer-assisted DSS compared to the rating given when she did use a computer-assisted DSS. This result replicated experiment 1 and experiment 2. Unlike experiment 1, however, the use of a prestigious aid—developed at the Mayo Clinic—mitigated the patients' derogation of the physician; no significant difference was detected between the no-aid and prestigious-aid groups.

The other dependent variables did not differ significantly as a function of DSS use. A likely reason for this variation between the 2 experiments is that there were far fewer participants in experiment 3 than in both experiment 1 and experiment 2, which meant lower power to detect significant differences; in this experiment, the power for thoroughness of examination, professionalism of physician, and overall satisfaction ranged from .181 to .429. Furthermore, the difference between the ratings given to the overall satisfaction and thoroughness of the examination dependent variables by the aid and no-aid groups was actually greater in experiment 3 than in experiments 1 and 2.

To expand the generality of our findings, we decided to use medical students in experiment 4 rather than undergraduate students (experiments 1 and 2) or patients (experiment 3). We thought it was important to ascertain if those who were training to be physicians had the same opinion as laypersons with regard to doctors' use of decision aids.

EXPERIMENT 4

Method

Participants

In total, 131 3rd- and 4th-year medical students (students in their clinical years) at The Ohio State University participated in this experiment.

Materials

The experiment employed both the ankle and DVT scenarios that described an interaction between a doctor and a patient plus a brief questionnaire containing the target-dependent variables.

Procedure

An e-mail was sent to all 3rd- and 4th-year medical students at The Ohio State University soliciting their participation in this study. All participants were promised and paid $25. Of the 403 eligible medical students, 131 (33%) responded to our invitation to participate.

Approximately half of the participants read a modified version of the ankle fracture scenario used in experiment 1 ($n = 68$). The other participants read the DVT scenario ($n = 63$).

Some of the groups in this experiment differed from those used in the prior 3 experiments. In addition to the usual no-aid group ($n = 33$) were 3 new ones. In the 1st new group, the physician used a decision aid and explicitly heeded the recommendation of the aid ($n = 30$). In the 2nd group, the physician used a decision aid, explicitly defied the recommendation of the aid, and treated the patient in a less aggressive manner than the recommendation ($n = 38$). In the final new group, the physician used a decision aid, explicitly defied the recommendation of the aid, and treated the patient in a more aggressive manner than the recommendation ($n = 30$). We wanted to contrast these final 2 groups because of the results of prior research that suggests that physicians prefer overtreatment to undertreatment,[30,31] and medical student raters might therefore be more favorably disposed toward a physician who treats more aggressively—but not less aggressively—than a decision aid might recommend.

The same dependent variables were used as in the prior 3 studies with the addition of 2 questions. First, all participants were asked, "How do you think that other physicians would regard you if you did use a computer-assisted diagnostic aid in making a diagnosis in a case such as this one?" Respondents could answer by marking a 7-point scale whose labels varied from *extremely negative* to *extremely positive*. Second, all participants were asked, "If you did use such a decision aid in making a diagnosis and an adverse medical outcome occurred, do you think using the aid would make you more vulnerable or more protected to a claim of malpractice?" Respondents could answer by marking a 7-point scale whose labels varied from *extremely vulnerable* to *extremely protected*.

Results

The 8 dependent variables (malpractice vulnerability, opinion of colleagues, perceived difficulty of diagnosis, thoroughness of examination, evaluation of length of wait, diagnostic ability, professionalism, and overall satisfaction) were subjected to a MANOVA with 2 factors: scenario (ankle or DVT) and experimental group (no aid, heed aid, defy aid by treating less aggressively, defy aid by treating more aggressively).

As was the case in experiment 2, only the diagnostic ability of the physician was significantly affected by the aid manipulation (no aid, aid heeded, aid defied less aggressively, and aid defied more aggressively), $F(3, 128) = 3.71$, $P < 0.05$, $\eta^2 = .08$. See Table 4 for means. The physician who used no DSS was given the highest rating (4.70), whereas the physician who used the aid but defied the output of the aid by treating in a less aggressive manner was given the lowest rating (3.84). A significant difference (using Tukey post hoc procedure) was found between the no-aid group and the "defy—less aggressive" group ($P < 0.05$).

Although it was not of central interest, there was a significant main effect of scenario for several dependent variables: perceived difficulty, $F(1, 130) = 86.90$, $P < 0.05$; thoroughness of examination, $F(1, 130) = 30.09$, $P < 0.05$; wait time, $F(1, 130) = 10.83$, $P < 0.05$; diagnostic ability, $F(1, 130) = 14.91$, $P < 0.05$; and professionalism, $F(1, 130) = 5.00$, $P < 0.05$. The ankle scenario received significantly greater ratings of wait time and perceived difficulty of diagnosis. Physicians in the DVT scenario were perceived as more thorough and professional, and they were given higher ratings of diagnostic ability.

Although no group differences were found on the question pertaining to vulnerability to malpractice, it is interesting to note that the medical students provided an average rating of 4.21, which was slightly toward the "protected" end of the scale.

Table 4 Experiment 4 Group Means

Dependent Variable	Group			
	No Aid	Aid Heeded	Aid Defy—Less Aggressive	Aid Defy—More Aggressive
Physician perception	3.30	3.57	3.45	3.87
Malpractice vulnerability	4.15	4.13	4.08	4.53
Perceived difficulty	4.24	4.40	4.50	4.30
Diagnostic ability[a]	4.70	4.07	3.84	4.10
Overall satisfaction	4.55	4.47	3.95	3.93
Professionalism	4.85	4.73	4.66	4.43
Thoroughness of examination	4.30	4.33	4.42	4.37
Length of wait	3.27	3.50	3.37	3.10

a. According to a Tukey post hoc test, the no-aid and aid defy—less aggressive groups differed significantly on this dependent variable ($P < 0.05$).

Discussion

We again found that the physician who used no decision aid was rated the highest in diagnostic ability. The physician who used an aid but defied it by treating in a less aggressive manner was deemed to have significantly less diagnostic ability. This mirrors the results of Pezzo and Pezzo,[32] who found in their experiment 2 that a physician who defied the recommendation of a decision aid was considered to be less competent than one who heeded its recommendation. Pezzo and Pezzo[32] found no significant difference in rated competence between those physicians who defied an aid and those who used no aid, a difference we did find in our study. However, a major difference between the procedure in their study and ours was that in their study, an adverse outcome occurred. As we suspected, the physician in our study who defied the aid and treated in a more aggressive manner was rated higher than the physician who defied the aid and treated in a less aggressive manner, but this difference was not significant.

GENERAL DISCUSSION

In both experiment 1 and experiment 2, we found that respondents derogated the diagnostic ability of physicians who used a computer-based DSS. In experiment 1, we also found evidence that other characteristics of the physician, such as professionalism, were also rated lower if a DSS was used. Unfortunately, patients' derogation of physicians who use a decision aid might be a defensible response, even if the aid does promote accurate judgment performance. According to social psychology's discounting principle,[33] whenever an effect has 2 possible causes, a perceiver tends to emphasize one cause and discount the other. Thus, if a correct diagnosis is made, and its accuracy can be attributed to either the skill of the physician or the output of the computer-assisted decision aid, it may be understandable why the perceiver may attribute some component of the accuracy to the aid rather than attributing all of it to the person. Consistent with this analysis, Pezzo and Pezzo[32] found that compared to a situation in which a physician did not use a decision aid, if a physician did use the decision aid, he or she was rated less positive following a good outcome and less negative following a bad one. Raters apparently attributed to the aid some of the credit or blame for the outcome, thus reducing the magnitude of the impact of the outcome on their evaluation of the physician. An unfortunate consequence of this attribution is that it may make decision makers less willing to use the decision aid for fear of appearing less capable should the outcome be a typical positive one. Because patients want their physician to be extremely knowledgeable (if not omniscient), the physician who appears to rely on a DSS may be risking the patient's otherwise high regard. The physician prudent enough to employ a helpful aid should be thanked—not derogated. We hasten to point out that this attributional complexity is germane not only to the medical arena but to virtually all other domains in which professional judgment is required. The discounting principle may be involved whenever a decision maker's ability and the decision aid's precision are each possible causes of a professional judgment.

We were surprised in experiment 1 that the independent variable significantly influenced not only the rated diagnostic ability of the physician—the dependent variable on which we anticipated the greatest impact—but 3 of the 4 additional dependent variables too. Some of these additional factors, such as the thoroughness of the examination, would not seem to be directly related to the use of a diagnostic aid. However, another social psychology principle, the "halo effect," might apply here.[34] This effect pertains to the phenomenon in which an evaluation of 1 particular salient trait of a person influences the evaluation of other traits of that same person. Thus, if one thinks that a physician is a good diagnostician, one is likely to believe in addition that the physician has other positive characteristics—very thorough, for example—even if there exists no relevant evidence bearing on these other factors.

Possible Theoretical Frameworks

Loss of the Aura of Omniscience

Kaplan[35] suggests that doctors' negative attitudes toward and underutilization of decision aids might both be fueled by the fear that the use of such aids might diminish their professional status. Our data suggest that this fear might be well founded in that patients rate the physician who uses an aid as a less competent diagnostician than a physician who acts without such assistance. Abundant research suggests that people think they can do better than a decision aid. For example, Arkes and colleagues[36] demonstrated that baseball experts were reluctant to use a decision rule despite the fact that it would have improved their performance significantly on a baseball-related judgment task. The experts were confident that they could do well without the helpful rule. Whitecotton[37] showed that there was a significant negative correlation between the confidence that professional financial analysts manifested toward the forecasting task that confronted them and their willingness to use a decision aid to perform the task. If we assume that those who hire the experts are as confident in the experts' performance as the experts themselves are, this would explain why patients have a relatively negative view toward expert diagnosticians who do not seem self-assured enough to eschew a decision aid.

Intuition v. Analysis

Insight and creativity are often associated with intuitive thinking. We admire such attributes. Analysis, on the other hand, seems tedious and opaque to a layperson. Latham and Whyte[38] and Whyte and Latham[39] investigated the relative persuasiveness of advice provided to managers either by an unaided expert or by an expert who had performed a utility analysis to support his or her recommendation. Although the advice was identical, the recommended course of action was less likely to be supported by the managers if it were predicated on the utility analysis. The fact that the merits of the utility analysis were explained by an expert who was an "internationally recognized authority" did not seem to impress the managers in a positive way.

We hypothesize that patient-raters might have an opinion similar to that of the managers in the aforementioned research. The unaided physician is analogous to the unaided expert who uses only intuition. The physician who uses a DSS is analogous to the expert who uses a decision analysis. In both situations, the person who uses his or her own intuition benefits, either by having the advice heeded or by being given higher evaluations. One difference in the 2 domains is that in the utility analysis study by Whyte and Latham,[39] the support of a prestigious expert did not seem to help. In our experiment 3, we did find that when patients learned that the aid had been developed at the prestigious Mayo Clinic, the derogation usually visited upon the physician who used a decision aid was largely mitigated.

Several authors have suggested reasons why people trust intuition more than analysis.[40–42] Among the most prominent is the suspicion that some important intuitions are performed without awareness and therefore cannot be expressed in analytic terms; this would therefore render them unavailable for inclusion in a decision aid. Whether this suspicion is warranted is a matter of debate.[43,44]

Case-Specific Data v. Base Rates

DSS tools gain their power from being based on a large amount of prior information—namely, the base rates of various diseases and the likelihood ratios that describe the probability of having a particular symptom given the presence of each disease. An abundance of psychological research compels the conclusions that most people do not 1) think that base rates are relevant,[45] 2) use them appropriately when they are presented,[46] or 3) consider them to be as probative as individuating information.[47] We want our physician to see us as individuals. Each of us is unique, and each of us wants to be treated accordingly. If a DSS lumps each of us with all the other people whose data go into the base rates or likelihood ratios, we feel minimized or even disregarded.[40] A physician who uses a DSS that offends us in this manner is going to be derogated compared to a physician who appears to appreciate our individuality, even though the latter physician may not be as accurate a diagnostician as the former.

The Issue of Generality

A number of factors that varied between and within the 4 experiments did not qualify the general finding that physicians who used no decision aid were deemed to be more capable than at least some physicians who did.

In experiments 1, 2, and 4, the pronouns used in the scenario would lead one to conclude that the

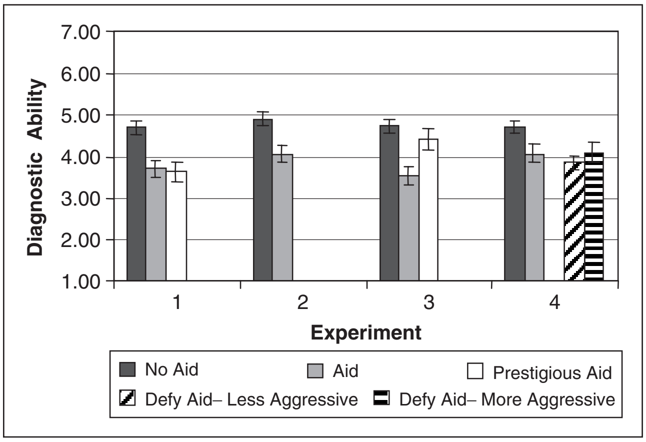

Figure 1 Ratings of diagnostic ability across experiments.

physician was a female. In experiment 3, the physician was a male. The principal result was the same in all studies.

In experiment 2, the medical condition was more serious than that in experiments 1 and 3. The principal results were the same in all 3 studies.

In experiments 1 and 2, the participants were younger than those in experiments 3 and 4. The principal results were the same.

In experiment 4, the participants were in training to be physicians, whereas the participants in the other studies were laypersons. Again, the principal result was the same.

We conclude that our principal finding has substantial generality. The common pattern of results in all 4 experiments is depicted in Figure 1.

External Validity

Reviewers of a prior draft of this article have questioned whether various aspects of our scenarios faithfully depict a typical doctor-patient interaction. For example, would a patient normally be aware of the fact that the physician has used a decision aid? In order for us to manipulate this independent variable, the participant in our research who was taking the role of the patient had to be made aware of whether the physician was using a decision aid. Thus, to investigate the influence of this factor, we had to make the presence of the aid quite explicit. From personal experience, we know that patients sometimes are aware of the fact that the physician has consulted a reference book, a colleague, or a journal article. We wanted to investigate what would be the patient's attitude if a computer-assisted aid were consulted instead.

Another closely related issue pertains to the context in which the aid was used. Perhaps a patient's attitude toward the physician who used a DSS would vary depending on whether the aid was used to make the initial diagnosis, corroborate a physician's own diagnosis, amalgamate a large number of complex laboratory tests, make the final diagnosis, or even override a physician's intuition. Of course, all of these are potentially important qualifications of our main finding. In our experiment 1, our scenario included the following: "she is using a computer program to decide whether to order an X-ray of your ankle." In experiment 3, we used the following: "he explains that to help make the diagnosis, he is using a computer program." Compared to its helping or consultative role in experiment 3, the DSS seems to be playing a more direct role in making the diagnosis in experiment 1. In both cases, the physician who used no decision aid was rated as more competent than the physician who did use an aid. Future research might wish to diminish the role of the decision aid even further to ascertain if the effect remains when the aid's role becomes truly nominal.

Another feature of our scenarios is that the use of a DSS may have delayed or frustrated a person who expected the physician to respond immediately with the recommended course of action. In experiment 1, the patient in the scenario had already waited 45 minutes, and the additional wait while the DSS was being used may have been particularly aversive. This extra delay may have been a source of some of the displeasure directed toward the physician who used a diagnostic aid. To the extent the use of an aid does not result in any delay or does not follow a wait for an already tardy and harried health care provider, patients may be less negative toward the physician who employs a DSS.

The results of these studies pose a problem for a physician who believes that computer-assisted DSSs improve diagnostic accuracy, as several studies suggest.[6–11] Using such decision aids may indeed increase diagnostic accuracy, but they might also simultaneously decrease the patient's opinion of the physician's diagnostic ability.

APPENDIX A
Experiment 1 Ankle Scenarios

NO-AID SCENARIO

On Saturday afternoon, during an informal game of softball at the local park, you hurt your left ankle. You jumped up to catch a line drive, and when you landed, your ankle turned in. You fell to the ground and were unable to get up or walk because of the pain. Your teammates, one of whom is a physical therapist, helped you to the side and got some ice for your ankle from the concession stand. Your friend gave you a ride home and helped you to your couch. She recommended that you see your physician tomorrow. Until then, you kept your ankle elevated, used ice, and took some ibuprofen for the pain.

Early the next morning, at 7:30 AM, you called your physician's office. The recorded message said they do not take calls for appointments for another hour and a half. You called back promptly 90 minutes later and got an appointment later that afternoon. You managed to go to work, using an ornamental walking stick you brought back from a trip.

In the afternoon, you left work early to get to your appointment. After a 30-minute wait, a nurse takes you to an examination room. The nurse asks you what the problem is, and you respond that you injured your left ankle yesterday playing softball. The nurse takes your temperature (98.7°), measures your blood pressure (122/78), takes your pulse (78 beats per minute), measures your respiratory rate (16 breaths per minute), and asks you to step on the scale. The nurse writes this information in the chart and then leaves. You want to ask her if you should get an X-ray to help move things along, but she left before you had a chance.

About 15 minutes later, the doctor comes into the room and asks you a number of questions:

1. When did this happen?
   Last evening.
2. What were you doing?
   Playing softball.
3. Describe the accident.
   When I landed, my ankle turned in. I could not walk on it. My friends helped me to the side and got some ice for it. The ankle swelled up, and there seems to be a bruise over the outer part of it.
4. Can you walk on it now?
   No, I cannot put my full weight on it. I must use this walking stick.
5. Have you ever injured your ankle before?
   Yes, but only minor twists.

In addition, the doctor asks you some more questions:

Do you have any drug allergies? *No.*
Do you have any other major health problems? *No.*

The doctor uses her stethoscope to listen to your lungs as you breathe deeply and to your heart while you lie back quietly. She then examines your ankles: your left ankle is puffy. There is a visible black and blue mark on the outside, over the bone. When she pushes over the bone on the outside of the ankle, it is tender. She moves it side to side and back and forth. It is beginning to throb. She asks you to walk on it without the walking stick. You are unable to. The doctor asks you to sit down again.

[Additional text for aid scenario: The doctor then turns to a computer in the room, explaining that she is using a computer program to decide whether to order an X-ray of your ankle. According to this decision aid, you should have an X-ray of your ankle to see if it is fractured. So she orders an X-ray of your ankle.] [Additional text for prestigious-aid scenario: The doctor then turns to a computer in the room, explaining that she is using a computer program developed at the prestigious Mayo Clinic, one of the nation's premier medical facilities. She explains that she is using their computer program to decide whether to order an X-ray of your ankle. According to this decision aid, you should have an X-ray of your ankle to see if it is fractured. So she orders an X-ray of your ankle.] [Concluding sentence in the no-aid scenario: Your doctor explains to you that she is concerned that you might have fractured your ankle during the injury. So she orders an X-ray of your ankle.]

NUMERACY SCALE

1. Imagine that we rolled a fair, 6-sided die 1000 times. Out of 1000 rolls, how many times do you think the die would come up even (2, 4, or 6)?
2. In the BIG BUCKS LOTTERY, the chance of winning a $10.00 prize is 1%. What is your best guess about how many people would win a $10.00 prize if 1000 people each buy a single ticket to BIG BUCKS?
3. In the ACME PUBLISHING SWEEPSTAKES, the chance of winning a car is 1 in 1000. What percent of tickets to ACME PUBLISHING SWEEPSTAKES wins a car?

APPENDIX B
Experiment 2 Deep Venous Thrombosis Scenarios

In this study, we are interested in what factors lead people to like or dislike their interactions with medical personnel. We are going to present you with a scenario describing a patient-physician interaction. Please take the role of a patient, read the story, and then, using your own medical knowledge and experience, give us your candid opinion of this physician.

After 5 days at home with a painful, swollen right leg, you decide to go to the doctor's office. You wait in the waiting room for about 30 minutes with approximately a half-dozen other people. Some of them are coughing, sneezing, and blowing their noses; you find the atmosphere very noisy.

After the 30-minute wait, a nurse takes you to an examination room. The nurse asks you what the problem is, and you respond that you've had a very swollen and sore leg for 5 days, after spraining your ankle skiing. The nurse takes your temperature (98.7°), measures your blood pressure (122/78), takes your pulse (78 beats per minute), measures your respiratory rate (16 breaths per minute), and asks you to step on the scale so you can be weighed. The nurse writes this information in the chart and then leaves you alone in the room. About 15 minutes later, the doctor comes into the room.

The doctor asks you a large number of questions:

- Why did you come to the doctor's office today? *You repeat that your leg has been painful and swollen for 5 days after you sprained your ankle skiing.*
- Can you walk on the leg? *You could after the accident, but for the last 3 days, you have stayed off of your leg. In fact, the ankle feels much better, but the leg itself hurts.*
- Has anyone examined the leg since the injury? *After the accident, the local emergency department examined and X-rayed the leg. The emergency department physician told you that it was just a sprain—you did not tear any ligaments and did not fracture any bones.*
- Do you have paralysis, paresis (weakness), or recent cast immobilization of the lower extremities? *No, you used an Ace wrap on the ankle for a couple of days.*
- Have you recently been bedridden for more than 3 days? *No, you have not stayed in bed but have been lying on the couch with your leg elevated most of the time since the accident. You have been getting up to go to the bathroom, make yourself meals, and stretch out occasionally.*
- Have you recently traveled anywhere, sitting still for 4 or more hours? *Yes, the trip home from the ski resort took 6 hours.*
- Do you have active cancer, are you being treated for cancer, have you been treated for cancer within the previous 6 months, or are you now having palliative care? *No.*
- Have you had major surgery within 4 weeks? *No.*
- Have you ever had a blood clot? *No.*
- Has anyone in your family ever had a blood clot? *No.*
- Have you been running any fevers or had shaking chills? *No.*
- Have you had any recent bug bites, cuts, or scrapes on the leg? *No.*
- Do you have a cough? *No.*
- Have you coughed up any blood? *No.*
- Can you feel your heart pounding? *No.*
- Are you short of breath? *No.*

In addition, the doctor asks you some more questions:

- Do you have any drug allergies? *No.*
- Do you have any other major health problems? *No.*
- Do you bleed easily? *No.*
- Have you ever had a stroke or an ulcer? *No.*

The doctor uses her stethoscope to listen to your heart and your lungs as you breathe deeply. She examines your legs: your right leg is swollen to just above the knee and is tender to gentle squeezing throughout. It is warm to the touch but not red. The right calf is 1.5 inches in diameter larger than your left leg. If you push on the leg, you can see the indentation after you remove your finger. The veins on your leg are not evident.

Your doctor explains to you that she is very concerned that you have a blood clot in your leg (a "deep venous thrombosis," she calls it). This can be dangerous, even fatal. So, she wants to check some blood test and send you immediately for a test (a "duplex ultrasound") to see if there is a clot in your right leg. She says she will be right back, after she arranges for the test.

Twenty minutes later, the nurse returns to the room. First, the nurse draws some blood. Then the nurse tells you that the test has been ordered. The nurse then takes you over to the test in a wheelchair and tells you that someone will come and pick you up when the test is over.

The test takes 20 minutes, but then you wait an hour before someone brings you back to the clinic. In the clinic, you wait for another 15 minutes before the doctor comes in to see you. She sits down and examines the test results. After a brief pause, she tells you that these tests do not show a blood clot in the leg.

The doctor says that you should take ibuprofen and use hot/cold packs for the pain.

[NO DECISION AID] However, the doctor says that you are at intermediate risk for a deep venous thrombosis, and so the duplex ultrasound should be done again in 1 week. You make an appointment 1 week from now.

[COMPUTER DECISION AID] The doctor then turns to a computer in the room, explaining to you that she is using a computer program to decide what to do next. According to this decision aid, you are at intermediate risk for a deep venous thrombosis, and the "advice" from the decision aid is that the duplex ultrasound should be done again in 1 week. So you make an appointment 1 week from now.

Before you leave, the doctor proceeds to explain to you how a blood clot is treated, just in case you do eventually develop a clot. You would be admitted to the hospital for 4 to 7 days. First an intravenous (IV) blood thinner will be started, and then a pill form of a blood thinner will be started. When your blood is thin enough, the IV thinner will be stopped and you can go home on the pill. You will probably remain on the pill for 6 months. After hearing this explanation, you leave the doctor's office.

APPENDIX C
Experiment 3 Cough Scenarios

NO-AID SCENARIO

After 5 days with a particularly bad sore throat, you decide to go to the doctor. You wait in the waiting room for about 30 minutes with approximately a half-dozen other people. Some of them are coughing, sneezing, and blowing their nose.

After the 30-minute wait, a nurse takes you to an examination room. The nurse asks you what the problem is, and you respond that you've had a very sore throat for 5 days. It's really hard for you to speak to the nurse because your throat is so sore. Your voice is extremely hoarse. The nurse takes your temperature (99.5°), measures your blood pressure (120/80), and asks you to step on the scale. The nurse writes this information in the chart and then leaves. About 15 minutes later, the doctor comes into the room.

The doctor asks you a large number of questions concerning the following:

1. Whether you have been coughing up blood (*no*)
2. What color the material is that you do cough up (*yellow*)
3. When the sore throat started (*5 days ago*)
4. Whether you have any chest pain (*no*)
5. Whether you are on any medications at the current time (*no*)
6. Whether the sore throat is more painful when you swallow (*yes*)
7. Whether you've had a lot of nasal congestion (*yes*)

The doctor also uses his stethoscope to listen to your heart and then to listen to your lungs as you breathe deeply.

He then asks more questions:

1. Do you have allergies? (*yes, ragweed*)
2. Do you have asthma? (*no*)
3. Have you ever had pneumonia? (*no*)

Although the doctor did not ask you about the topic, you volunteer the information that during the last several years, you've been treated for several urinary and kidney infections.

He then feels the lymph nodes under your jaw in order to determine if they are swollen. (They are.) He looks in your ears and asks if you have had any earaches in the last few days. You say that you have had a couple of earaches.

The doctor then says that he will order some laboratory tests. First, he takes a swab and brushes it against the back of your throat. He says that he wants to get a culture in case there is a strep infection. He says that he wants to get a chest X-ray because he wants to get a better idea of the condition of your lungs. He says that he detected some fluid when he listened to your lungs with the stethoscope. He says that the nurse will return to take a blood sample. He says that the laboratory tests will be back tomorrow and that you should stop by tomorrow early in the afternoon. He writes a prescription, which he tells you to get filled right away. The nurse comes in and draws some blood from your left arm. Then you go get the chest X-ray. When you finish with the X-rays, you go to the pharmacy to get the prescription filled. You go home, take the medicine, and spend the rest of the day in bed.

The following day, you stop back in the early afternoon for your 1 PM appointment. You wait about 30 minutes, and then you are called into the examination room. The nurse again takes your temperature (99.5°) and blood pressure (120/80). The doctor comes in 10 minutes later. He sits down and examines the laboratory test results, which have just come in. He has not examined the results before your appointment. He says that you don't have a strep infection. He then looks at the chest X-ray and looks again at the various laboratory results, including the results from the blood tests. [The following sentence is only in the no-aid scenario: He spends a minute or two mulling over this information.] [Additional text for aid scenario: He then opens a laptop computer and types in the results of the various laboratory tests. He also types in the information you provided yesterday concerning your coughing, earaches, etc. He explains that to help make the diagnosis, he is using a computer program {The following clause is inserted only in the prestigious-aid scenario: developed at the prestigious Mayo Clinic, one of the nation's premier medical facilities}. To run the program, he says, he merely types in the information you've provided. The computer program then assigns a likelihood to each of several possible diagnoses. He says that based on his examination of you yesterday, he thought at that time that you might have pneumonia, but based on the output of the computer program, he now is pretty sure that you have acute bronchitis.] He says that based on his examination of you yesterday, he thought at that time that you might have pneumonia, but based on the laboratory tests and X-ray, he now is pretty sure that you have acute bronchitis. He tells you to continue to take the medication he prescribed yesterday, and in addition, he writes you a 2nd prescription for a cough suppressant. He tells you to come back in 5 days if there is no improvement. He says that you will not be entirely well for a couple of weeks; the cough will last quite a long time.

REFERENCES

1. Kaplan B. Evaluating informatics applications: clinical decision support systems literature review. Int J Med Inform. 2001;64:15–37.

2. Balas EA. Information systems can prevent errors and improve quality. J Am Med Inform Assoc. 2001;8:398–9.

3. Sim I, Gorman P, Greenes RA, et al. Clinical decision support systems for the practice of evidence-based medicine. J Am Med Inform Assoc. 2001;8:527–34.

4. 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;92:238–42.

5. Dawes RM, Faust D, Meehl PE. Clinical versus actuarial judgment. Science. 1989;243:1668–74.

6. Corey GA, Merenstein JH. Applying the acute ischemic heart disease predictive instrument. J Fam Pract. 1987;25:127–33.

7. de Dombal FT, Dallos V, McAdam WAF. Can computer aided teaching packages improve clinical care in patients with acute abdominal pain? BMJ. 1991;310:1495–7.

8. Ridderikhoff J, van Herk E. A diagnostic support system in general practice: is it feasible? Int J Med Inform. 1997;45:133–43.

9. Friedman DP, Elstein AS, Wolf FM. Enhancement of clinicians' diagnostic reasoning by computer-based consultation: a multisite study of 2 systems. JAMA. 1999;282:1851–5.

10. Getty DJ, Pickett RM, D'Orsi CJ, Swets JA. Enhanced interpretation of diagnostic images. Invest Radiol. 1988;23:240–52.

11. Chase CR, Vacek PM, Shinozaki T, Giard AM, Ashikaga T. Medical information management: improving the transfer of research results to presurgical evaluation. Med Care. 1983;21:410–24.

12. Ebell MH, Smith MA, Barry HC, Ives K, Carey M. Does this patient have strep throat? JAMA. 2000;284:2912–8.

13. Centor RM, Witherspoon JM, Dalton HP, Brody CE, Link K. The diagnosis of strep throat in adults in the emergency room. Med Decis Making. 1981;1:239–46.

14. Walsh BT, Bookheim WW, Johnson RC, Tompkins RK. Recognition of *Streptococcal pharyngitis* in adults. Arch Intern Med. 1975;135:1493–7.

15. Stiell I. Ottawa ankle rules. Can Fam Physician. 1996;42:478–80.

16. Hunt DL, Haynes B, Hanna SE, Smith K. Effects of computer-based clinical decision support systems on physician performance and patient outcomes. JAMA. 1998;280:1339–46.

17. Kassirer JP. A report card on computer-assisted diagnosis—the grade C. New Engl J Med. 1994;330:1824–5.

18. Rocha BHSC, Christenson JC, Evans RS, Gardner RM. Clinicians' response to computerized detection of infections. J Am Med Inform Assoc. 2001;8:117–25.

19. Overhage JM, Tierney WM, McDonald CJ. Computer reminders to implement preventive care guidelines for hospitalized patients. Arch Intern Med. 1996;156:1551–6.

20. Keefe B, Subramanian U, Tierney WM, et al. Provider response to computer-based care suggestions for chronic heart failure. Med Care. 2005;43:461–5.

21. Graham ID, Stiell IG, Laupacis A, et al. Awareness and use of the Ottawa Ankle and Knee Rules in 5 countries: can publication alone be enough to change practice? Ann Emerg Med. 2001;37:259–66.

22. Hilton NZ, Simmons JL. The influence of actuarial risk assessment in clinical judgments and tribunal decisions about mentally disordered offenders in maximum security. Law Hum Beh. 2001;25:393–408.

23. Cabana MD, Rand CS, Powe NR. Why don't physicians follow clinical practice guidelines? A framework for improvement. JAMA. 1999;282:1458–65.

24. Meehl PE. Causes and effects of my disturbing little book. J Pers Assess. 1986;50:370–5.

25. Cruikshank PJ. Patient ratings of doctors using computers. Soc Sci Med. 1985;21:615–22.

26. Garrison GM, Bernard ME. 21st century health care: the effect of computer use by physicians on patient satisfaction at a family medicine clinic. Fam Med. 2002;34:362–8.

27. Solomon GL, Dechter M. Are patients pleased with computer use in the examination room? J Fam Pract. 1995;41:242–4.

28. Lipkus IM, Samsa G, Rimer BK. General performance on a numeracy scale among highly educated samples. Med Decis Making. 2001;21:37–44.

29. Schwartz LM, Woloshin S, Black WC, Welch GH. The role of numeracy in understanding the benefit of screening mammography. Ann Intern Med. 1997;127:966–71.

30. Ayanian JZ, Berwick DM. Do physicians have a bias toward action? A classic study revisited. Med Decis Making. 1991;11:154–8.

31. Bakwin H. Pseudodoxia pediatrica. New Engl J Med. 1945;232:691–7.

32. Pezzo MV, Pezzo SP. Physician evaluation after medical errors: does having a computer decision aid help or hurt in hindsight? Med Decis Making. 2006;26:48–56.

33. Kelley HH. The process of causal attribution. Am Psychol. 1973;28:107–28.

34. Thorndike EL. A constant error on psychological rating. J Appl Psychol. 1920;4:25–9.

35. Kaplan B. Culture counts: how institutional values affect computer use. MD Comput. 2000;17:23–6.

36. Arkes HR, Dawes RM, Christensen C. Factors influencing the use of a decision rule in a probabilistic task. Organ Behav Hum Dec. 1986;37:93–110.

37. Whitecotton SM. The effects of experience and confidence on decision aid reliance: a causal model. Behav Res Account. 1996; 8:194–216.

38. Latham GP, Whyte G. The futility of utility analysis. Pers Psychol. 1994;47:31–46.

39. Whyte G, Latham GP. The futility of utility analysis revisited: when even an expert fails. Pers Psychol. 1997;50:601–10.

40. Dawes RM. The robust beauty of improper linear models. Am Psychol. 1979;34:571–82.

41. Meehl PE. Causes and effects of my disturbing little book. J Pers Assess. 1986;50:370–5.

42. Yates JF, Veinott ES, Patalano AL. Hard decisions, bad decisions: on decision quality and decision aiding. In: Schneider SL, Shanteau J, eds. Emerging Perspectives on Judgment and Decision Research. New York: Cambridge University Press; 2003. p 13–63.

43. Hammond KR. Human Judgment and Social Policy: Irreducible Uncertainty, Inevitable Error, Unavoidable Injustice. New York: Oxford University Press; 1996.

44. Hogarth RM. Educating Intuition. Chicago: University of Chicago Press; 2001.

45. Arkes HR. Impediments to accurate clinical judgment and possible ways to minimize their impact. J Consult Clin Psych. 1981;49: 323–30.

46. Arkes HR, Rothbart M. Memory organization, retrieval and contingency judgments. J Pers Soc Psychol. 1985;49:598–606.

47. Kahneman D, Tversky A. On the psychology of prediction. Psychol Rev. 1973;80:237–51.