

# The Influence of a Physician's Use of a Diagnostic Decision Aid on the Malpractice Verdicts of Mock Jurors

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

**Background.** One reason why physicians may be reluctant to use diagnostic decision aids is that such usage might increase the likelihood of an unfavorable malpractice verdict. The authors tested this hypothesis by sending a DVD of a malpractice trial to a national sample of jury-eligible adults. **Methods.** There were 3 independent variables: 1) the physician did or did not use a diagnostic aid, 2) the patient's symptoms either were or were not consistent with a diagnosis of probable appendicitis, and 3) the physician's decision to operate or not operate was either concordant or discordant with the severity of the patient's symptoms. Jurors rendered a verdict, and if they deemed the physician not to have met the standard of care, they indicated how punitive they felt toward the physician. **Results.** Mock jurors were more likely to side with the

physician-defendant if he recommended an operation when there were many symptoms and refrained when there were few symptoms compared with a physician who did the converse. The use of a decision aid had no influence on this binary standard-of-care decision. Among those physicians deemed liable by the jurors, defying the aid resulted in heightened punishment compared with heeding it. **Conclusion.** Contrary to many physicians' fears, use of a diagnostic decision aid did not influence the likelihood of an adverse malpractice verdict. Complying with the aid's recommendation provided a measure of protection against jurors' punitiveness for those physicians deemed liable for malpractice. **Key words:** malpractice; decision aids; jury decision making. (*Med Decis Making* 2008;28:201–208)

**D**iagnostic support systems are to be distinguished from the larger category of decision support systems, which can assist the health care provider in tasks other than diagnosis, such as drug

dosing or other therapeutic decisions. Although some reviews have questioned the usefulness of diagnostic aids,<sup>1,2</sup> many successful examples have been reported in the literature.<sup>3–16</sup> However, there is unanimity among researchers with regard to 1 characteristic of diagnostic aids: They are grossly underutilized.<sup>17</sup> One reason why diagnostic aids might be underutilized is that such usage might increase the likelihood of an unfavorable malpractice verdict. Should this hypothesis be true, then physicians might understandably eschew the use of such aids.

The goal of the present research was to investigate the influence of a physician's use of a diagnostic aid on jurors' malpractice verdicts. Unlike the few prior studies in this area, we did not use written scenarios but instead used a highly realistic DVD of a malpractice trial. A national sample of jury-eligible adults was shown the DVD and asked to render a verdict as to whether the physician met the standard of care. Two competing hypotheses were pitted against each other. Perhaps the use of the diagnostic aid might be seen by the mock jurors as an abrogation of the physician's responsibility,

Received from the Department of Psychology; Division of Health Services, Management, and Policy; and the Center for Health Outcomes, Policy, and Evaluation Studies, The Ohio State University, Columbus (HRA); the Department of Psychology, Wichita State University, Kansas (VAS); and the Department of Internal Medicine, Division of General Internal Medicine, The Ohio State University, Columbus (MAM). Presented at the annual meetings of the Society for Medical Decision Making, Boston, 2006, and the Society for General Internal Medicine, Toronto, Canada, 2007, and the Society for Judgment and Decision Making, Houston, 2006. This research was supported by a grant from the Program in Decision, Risk, and Management Science at the National Science Foundation (NSF). The NSF had no role in the design, conduct, data analysis, or reporting of this research other than to evaluate the original proposal on which the research was based. Revision accepted for publication 28 July 2007.

Address correspondence to Hal R. Arkes, PhD, Department of Psychology, The Ohio State University, 1827 Neil Avenue, Lazenby 240N, Columbus, OH 43210-1222; e-mail: arkes.1@osu.edu.

DOI: 10.1177/0272989X07313280

thus rendering him more blameworthy.<sup>18</sup> On the other hand, a physician who uses a diagnostic aid might be seen as an up-to-date practitioner who is relying on a helpful and modern aid, thus rendering him less blameworthy.

## METHODS

### Participants

Participants were recruited by Knowledge Networks (KN), a company that conducts surveys in the United States. KN offers free Internet service to a large national sample of Americans who are randomly recruited for the purpose of filling out surveys online. The randomly selected persons are not limited to those who did not have Internet service before being recruited by KN. The 1000 persons sampled for this study were drawn from the entire US residential population with telephone service.

### Materials

All survey respondents received 1 of 8 DVDs in the mail. Each DVD depicted a mock malpractice trial filmed in an actual courtroom. Professional actors portrayed the judge, 2 attorneys, and 2 physicians, who served as expert witnesses. Each DVD lasted approximately 11 min and contained testimony by each expert witness, cross-examination of each expert witness, and instructions by the judge to the jury. Before each respondent watched the DVD, he or she was instructed to first read a 1.5-page introduction to the dispute. The introduction and the DVD contained the following basic facts. Ms Marcia Rodman went to the emergency department of a hospital following the onset of abdominal pain. She was seen by Dr Stephens, who suspected that she might have appendicitis.

One experimental (independent) variable pertained to whether Dr Stephens did or did not use the Alvarado Score<sup>19</sup> in deciding whether Ms Rodman did indeed have appendicitis. This decision aid awards points for each symptom consistent with a diagnosis of appendicitis, with some symptoms counting for more points than others.

The 2nd experimental variable was symptom severity. For half of the participants, Ms Rodman's symptoms were less severe, corresponding to an Alvarado Score of 4 (not indicative of appendicitis). For the other half of the participants, Ms Rodman's symptoms were more severe, corresponding to an Alvarado

Score of 7 (indicative of probable appendicitis). Of course, if subjects were among those whose physician did not use the Alvarado Score, Ms Rodman's symptoms were described without any reference to how they would contribute toward the calculation of the Alvarado Score.

The 3rd experimental variable was whether Dr Stephens' behavior was symptom concordant or symptom discordant. If Ms Rodman's symptoms were less severe, then symptom-concordant behavior would be not to operate, and that is what the Alvarado Score would recommend if it were to be used. If Ms Rodman's symptoms were more severe, then symptom-concordant behavior would be to operate, and that is what the Alvarado Score would recommend if it were to be used. To operate when the symptoms were less severe or to refrain from operating when they were more severe would comprise symptom-discordant behavior.

In all 8 combinations of these 3 experimental variables, precisely the same adverse outcome occurred: Ms Rodman experienced a serious infection, and she also lost wages because of her inability to work for several weeks. A complete transcript of the introductory written material plus a transcription of 1 of the 8 DVDs can be accessed at <http://faculty.psy.ohio-state.edu/arkes> by clicking on the "Rodman versus Stephens" link.

Before respondents were allowed to respond to the 2 queries that solicited the outcome (dependent) variables, they had to answer correctly 3 very simple multiple-choice questions to ensure that they had watched the DVD. The 3 questions were the following: 1) What medical problem did Ms Rodman's children have? (asthma, flu, diabetes, hearing loss, or hepatitis); 2) How many expert witnesses did each side have? (1, 2, 3, 4, or 5); and 3) About how many times was the defendant, Dr Stephens, shown during the trial? (4, 3, 2, 1, or 0).

After answering these 3 screening questions, all respondents were asked the following question: "Now we'd like you to imagine yourself as one of the jurors. Do you think that Dr Stephens met the standard of care in his treatment of Ms Rodman?" The 2 options were, "Yes, Dr Stephens DID meet the standard of care, and he is therefore NOT GUILTY of medical malpractice" and "No, Dr Stephens DID NOT meet the standard of care, and he is therefore GUILTY of medical malpractice." Those who decided that Dr Stephens did not meet the standard of care (i.e., was liable for medical malpractice) were asked to respond to a 2nd question: "How deserving of punishment do you think Dr Stephens

is?" A 9-point Likert-type scale followed, which was anchored at 1 = *not at all deserving of punishment* and 9 = *highly deserving of punishment*. We opted to use this scale rather than a scale soliciting opinions of appropriate monetary damages because of the extreme skewness of responses to the latter type of scale.<sup>20</sup>

**Procedure**

All potential respondents first received a solicitation letter explaining the study and seeking their consent to participate. Those who agreed to participate were randomly assigned to 1 of the 8 different scenarios. They subsequently received in the mail the 1.5-page introduction to the trial plus 1 of the 8 DVDs of the trial. Participants then answered the 3 screening questions and the 2 outcome variable questions via the Internet connection provided by KN. However, before they did so, participants had to indicate how much time had elapsed since they viewed the DVD. If more than 30 min had elapsed, they were instructed to view it again before answering any questions about it.

**Statistical Methods**

To identify the statistically significant independent predictors of finding Dr Stephens liable for malpractice among the demographic and experimental variables, we used forward selection in a logistic regression model. The demographic variables were sex, age, educational level, race/ethnicity, income, employment, US region, and urbanization. The experimental variables were the severity of the symptoms, use of the aid (yes/no), and the doctor's symptom-concordant/-discordant behavior. A bivariate analysis, using contingency tables, was run with each variable and the malpractice verdict. Variables with a  $P \leq 0.20$  were used in the logistic regression analysis. We then used forward selection, retaining variables with a  $P \leq 0.05$  on the likelihood ratio test for the logistic regression model. Appropriate interaction terms were also checked. In the logistic regression, variables with multiple categories were entered as treatment contrasts.

We performed a 2 (use of aid: yes/no)  $\times$  2 (symptom concordant/discordant)  $\times$  2 (symptom severity: symptoms consistent with an Alvarado Score of 4/7) analysis of variance on the punitiveness measure. Tukey a posteriori tests were done using the Cicchetti<sup>21</sup> correction. Calculations were done using SPSS version 14.0 and R version 2.5.1. This research

**Table 1** Demographic Characteristics of Survey Respondents

	n	%
Sex		
Male	329	50.1
Female	328	49.9
Age <sup>a</sup>	47.0	34.0–59.0
Race/ethnicity		
White, non-Hispanic	489	74.4
Black, non-Hispanic	59	9.0
Hispanic	72	11.0
Other	37	5.6
Highest educational level obtained		
Less than high school	65	9.9
High school	223	33.9
Some college	188	28.6
Bachelor's degree or higher	181	27.5
Income		
\$0–\$34,999	235	35.8
\$35,000–\$74,999	270	41.1
$\geq$ \$75,000	152	23.1
Employment		
Employed	397	60.4
Not employed	260	39.6
US region		
Northeast	123	18.7
Midwest	162	24.6
South	220	33.5
West	152	23.1
Urbanization		
Metropolitan	541	82.3
Nonmetropolitan	116	17.6

a. Data are presented as the median and interquartile range.

protocol was approved both by The Ohio State University Institutional Review Board and the Western Institutional Review Board.

**RESULTS**

One thousand persons were solicited, and 660 agreed to participate. All but 3 provided a usable verdict, resulting in a sample of 657. Demographic characteristics of the 660 persons who agreed to participate are shown in Table 1.

Of the 657 participants, 398 (60.6%, 95% confidence interval [CI] = 56.7%–64.3%) thought that Dr Stephens met the standard of care (i.e., was not liable for medical malpractice) and 259 (39.4%, 95% CI = 35.7%–43.3%) thought that he did not (i.e., was liable for medical malpractice). Table 2 depicts the proportion of respondents in each of the 8 scenarios who thought that Dr Stephens was liable for malpractice.

**Table 2** Percentage of Respondents in Each Group Who Thought That Dr Stephens Was Liable

Decision Aid Used?	Symptoms	Physician Behavior (%)	
		Concordant (95% CI)	Discordant (95% CI)
No	Not probable appendicitis	34.1 (24.3–45.0)	43.4 (32.1–55.3)
	Probable appendicitis	35.6 (24.7–47.7)	43.4 (32.1–55.3)
Yes	Not probable appendicitis	31.4 (21.8–42.3)	42.8 (32.1–54.1)
	Probable appendicitis	34.1 (24.2–45.2)	50.6 (39.8–61.3)

Note: CI = confidence interval.

In the bivariate analysis, the following variables were associated ( $P < 0.20$ ) with finding Dr Stephens liable: sex, age, educational level, race/ethnicity, income, employment, and symptom-discordant behavior (Table 3). Aid use and symptom severity were not associated with malpractice verdict.

The logistic regression analysis selected age, educational level, race/ethnicity, and the doctor's symptom-concordant/-discordant behavior as statistically significant independent predictors of the malpractice verdict. When Dr Stephens behaved in a symptom-concordant manner, a potential juror was about two-thirds (odds ratio = 0.606, 95% CI = 0.437–0.840) as likely to find Dr Stephens liable for medical malpractice compared with those instances in which Dr Stephens behaved in a symptom-discordant manner (Table 4). Mock jurors who were older than 60 y, had at least a bachelor's degree, or were not members of a racial or ethnic minority were also less likely to find Dr Stephens liable for medical malpractice.  $R^2$  was 0.097, and the C index was 0.65 for the model. The 2-way interaction terms were not significant.

The 259 respondents who deemed Dr Stephens liable for medical malpractice were asked how deserving of punishment they felt the physician was. One subject did not respond to this question, leaving 258 subjects with complete data. We performed an analysis of variance on this punitiveness measure. The only significant result was the aid  $\times$  symptom-concordant/-discordant behavior interaction,  $F(1, 250) = 6.13$ ,  $P = 0.013$ ,  $\eta^2 = 0.024$ . When the physician did not use a decision aid, there was no significant difference in punitiveness ratings for symptom-concordant versus symptom-discordant behavior (5.36 [95% CI = 4.80–5.932] v. 4.95 [95%

**Table 3** Relation between Variables and a Verdict That the Physician Was Liable for Malpractice

Demographic Variable	N	Liable Verdict		P
		n	%	
Sex				
Male	329	119	36.2	0.088
Female	328	140	42.9	
Age (y)				
18–29	130	57	43.4	0.002
30–44	164	66	40.2	
45–59	200	92	46.0	
$\geq 60$	163	44	27.0	
Education level				
Less than high school	65	34	52.3	0.010
High school	223	93	41.7	
Some college	188	77	41.0	
Bachelor's degree or higher	181	55	30.4	
Race/ethnicity				
White, non-Hispanic	489	171	35.0	0.000
Other	168	88	52.4	
Income				
< \$35,000/y	235	105	44.7	0.119
\$35,000–\$74,999/y	270	98	36.3	
$\geq$ \$75,000/y	152	56	36.8	
Employment				
Employed	397	165	41.6	0.165
Unemployed	260	94	36.2	
US region				
Northeast	123	52	42.3	0.616
Midwest	162	66	40.7	
South	220	79	35.9	
West	152	62	40.8	
Urbanization				
Metropolitan	541	209	38.6	0.371
Nonmetropolitan	116	50	43.1	
<b>Experimental variables</b>				
Symptoms				
Not probable appendicitis	334	126	37.7	0.365
Probable appendicitis	323	133	41.2	
Aid use				
Yes	344	137	39.8	0.824
No	313	122	39.0	
Physician behavior				
Concordant	332	112	33.7	0.003
Discordant	325	147	45.2	

CI = 4.41–5.49]). When an aid was used, those who heeded it—that is, those who behaved concordantly with its recommendation—were rated significantly less punitively (4.74 [95% CI = 4.10–5.30]) according to a Tukey a posteriori test ( $q = 3.63$ ,  $P < 0.05$ ) than those who defied the advice of the aid by doing the

**Table 4** Logistic Regression on a Verdict That the Physician Was Liable

Variable	Coefficient (β)	Standard Error	Z	P	Odds Ratio	95% Confidence Interval
Intercept	0.885	0.325	1.184	0.236	—	—
Education level						
Less than high school	—	—	—	—	1	Referent category
High school	−0.419	0.296	−1.414	0.157	0.658	0.368 – 1.176
Some college	−0.494	0.302	−1.632	0.103	0.610	0.337 – 1.104
Bachelor’s degree or higher	−0.949	0.311	−3.050	0.00229	0.387	0.210 – 0.712
Age (y)						
18 – 29	—	—	—	—	1	Referent category
30 – 44	−0.037	0.246	−0.150	0.880	0.964	0.595 – 1.563
45 – 59	0.212	0.236	0.900	0.368	1.236	0.779 – 1.967
≥60	−0.665	0.260	−2.559	0.0105	0.514	0.309 – 0.856
Physician behavior						
Symptom discordant	—	—	—	—	1	Referent category
Symptom concordant	−0.501	0.166	−3.009	0.00262	0.606	0.437 – 0.840
Race/ethnicity						
Other	—	—	—	—	1	Referent category
White, non-Hispanic	−0.594	0.188	−3.162	0.00157	0.552	0.382 – 0.798

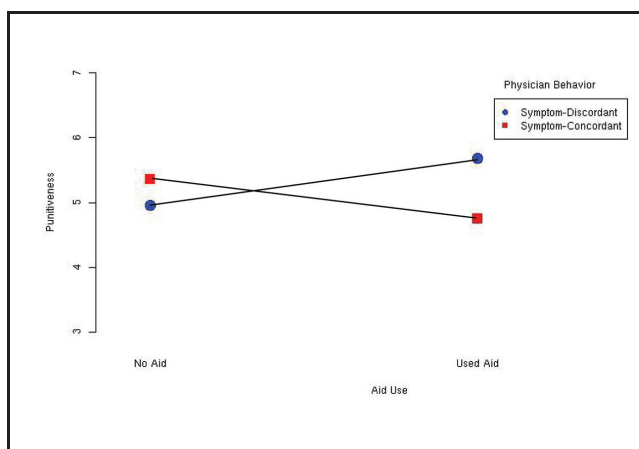


Figure 1 Interaction of the punitiveness ratings as a function of the concordance of the physician behavior with the patient’s symptoms and a physician’s use of the diagnostic aid.

opposite (5.68 [95% CI = 5.22–6.14]; Figure 1). Physicians were not significantly worse off if they defied the aid than if they had performed the same symptom-discordant behavior without using an aid. However, using and heeding an aid provided an advantage to Dr Stephens if he were to be sued following the adverse outcome.

In light of this interaction between the use of the aid and symptom-concordant behavior, and at the suggestion of 1 of the reviewers, we performed a post hoc analysis of the same interaction on verdict. The interaction between aid use and symptom-concordant behavior was significant in the binary

analysis,  $\chi^2(3) = 9.624, P = 0.022$ . When we repeated the forward stepwise logistic regression analysis with the interaction term included as an additional candidate variable, the interaction term was not selected and the final model was identical. We conclude that the main effect of behaving in concordance with the symptoms on the liability judgment was not qualified by any interaction.

**DISCUSSION**

There were 2 principal results of this investigation. First, mock jurors were more likely to side with the physician-defendant if he recommended an operation when there were many symptoms and refrained when there were few symptoms compared with a physician who did the converse. The use of a decision aid had no influence on this binary standard-of-care decision.

The 2nd main result pertained to the punitiveness measure, which was collected only for those mock jurors who thought that the physician did not meet the standard of care. If no diagnostic aid was used, mock jurors gave punitiveness ratings that did not differ significantly irrespective of what the physician did. However, physicians who used the aid and behaved concordantly with its recommendation were punished significantly less severely than those who defied it. Consistent with this finding, Pezzo and Pezzo<sup>22</sup> reported in their 1st experiment that compared with a physician who used no diagnostic

aid, a physician who used one was rated less negatively following an adverse outcome. However, these researchers also found that a physician who used a computer-assisted diagnostic aid was rated less positively after a favorable outcome compared with the unaided physician. Apparently, the participants gave the diagnostic aid some of the credit for a good outcome and some of the blame for a bad one, distributing the responsibility for the decision between the physician and the diagnostic aid. However, the groups did not differ in their likelihood of suing the doctor. In a 2nd experiment, Pezzo and Pezzo<sup>22</sup> presented a scenario in which a radiologist was accused of negligence leading to a patient's death. The researchers found that a physician who defied the recommendation of a computerized diagnostic aid was rated as less competent and more at fault than a physician who agreed with the aid or who initially disagreed with the aid but opted to obey its recommendation anyway. The results of this 2nd experiment are consistent with our finding that complying with the recommendation of a decision aid provides a measure of protection.

Secondary results included the findings that age and educational level were significant predictors of finding Dr Stephens liable. A large number of studies have found that older patients are more satisfied with their care than are younger patients,<sup>23-28</sup> although satisfaction levels may tend to drop among the very oldest patients who are in poor health.<sup>29</sup> This generally positive relation between age and satisfaction with care is consistent with our finding that the mock jurors older than 60 y were less likely to find Dr Stephens liable. Our finding that less educated mock jurors are more likely to deem the physician liable is also consistent with research indicating that such persons are more likely to be dissatisfied with their care.<sup>30,31</sup>

There is a substantial amount of literature on the relation between malpractice and the use of practice guidelines.<sup>32-38</sup> Although practice guidelines generally pertain to therapy and treatment rather than diagnosis, such guidelines bear some similarity to diagnostic aids in that they prescribe a recommended course of action to a medical professional. (In fact, readers of this article disagreed whether the Alvarado Score is more of a diagnostic decision aid or a practice guideline. Its calculation helps the physician decide whether appendicitis is very likely or very unlikely to be the correct diagnosis; however, it also provides a guideline for when to check an imaging study for further evaluation.<sup>39</sup>) Therefore, some of the research pertaining to the relation between malpractice and

the use of practice guidelines might be relevant to our discussion relating malpractice and the use of decision aids. Some observers have pointed out a very significant benefit of adhering to a practice guideline: Such adherence might serve as an effective shield against an accusation of malpractice.<sup>32-35</sup> Because practice guidelines have generally been endorsed by a reputable professional organization or published in a respected journal, a successful lawsuit against a physician who followed the guideline would seem unlikely. On the other hand, a physician's deviation from a practice guideline might provide a plaintiff with an effective "sword."<sup>36-38</sup> Generalizing from the domain of practice guidelines, one might speculate that 1 reason why diagnostic aids might not be used is that if the physician uses one, its recommendation must be followed to shield that physician from a successful malpractice suit should an adverse outcome occur. However, if the diagnostic aid is not even consulted, then the physician is less constrained with regard to the diagnosis and treatment he or she might favor. We found that if a physician wanted to act in a manner that was discordant with what an aid would recommend, the use of an aid did not result in a statistically significant increase in either liability verdicts or in juror punitiveness compared with the case in which the aid was eschewed. However, our results suggest that adherence to a decision aid may result in a significant protective benefit to the physician compared with the case in which the aid is defied.

Some researchers<sup>18,22,40</sup> have found that the use of a computer in rendering a diagnosis has an influence on a patient's opinion of the physician who uses such an aid. Arkes and others<sup>18</sup> have reported that the use of a computer-assisted diagnostic aid lowers the opinion a patient has of a physician compared with a physician who makes an unaided diagnosis. Apparently, some patients think that a physician who needs a diagnostic aid is not as capable as one who can render diagnoses without it. However, this potential cost of using a decision aid has to be weighed against the benefit we have demonstrated: Adherence to the recommendation of an aid mitigates punitiveness directed toward a physician found liable for malpractice. Of course, another potential benefit of using a decision aid is that, as several studies suggest, the use of a decision aid results in improved care.<sup>3-16</sup>

Among the reasons why physicians do not follow practice guidelines is that they may disagree with them.<sup>41</sup> We suggest that a similar situation may exist with regard to diagnostic aids. Once a physician uses a diagnostic aid, because of the fear of a

punitive jury should an adverse verdict occur, he or she might feel compelled to heed the diagnostic aid's suggested diagnosis even if the physician does not agree with it. For this reason, a physician might be reluctant to use a diagnostic aid. This situation is unfortunate and somewhat ironic, in that much research<sup>3-16</sup> has suggested that diagnostic aids promote diagnostic accuracy. If physicians are fearful of accepting beneficial assistance, then the diagnostic accuracy that might have been fostered by the use of diagnostic aids may go unrealized.

### Limitations

We acknowledge several limitations in the current research.

First, we did not have the individual jurors interact with each other, as would be the case in a normal jury trial. This would have been very difficult given our desire to obtain a nationally representative sample. To assemble 12-person juries in all parts of the United States would have been prohibitively expensive. Also, the group dynamics present in jury deliberations would have made it extremely problematic to ascertain the influence of our independent variables on each juror's decision. Given that this study was breaking new ground, we opted to start with a design that eliminated the influence of group dynamics.

Second, we acknowledge that a short DVD is not a faithful replication of an entire malpractice trial. However, our materials were of far higher fidelity than the written materials used in much jury research pertaining to malpractice.<sup>42-45</sup> All persons depicted in the DVD were professional actors, 1 of whom was also an attorney. We are confident that our respondents found the DVD to be highly engaging.

Third, because we used a national sample, our results may not generalize to jurisdictions in which the demographics of the jury pool differ substantially from that of the nation. For example, Helland and Tabarrok<sup>46</sup> have found that the average personal injury tort award increases as minority county population proportions increase and as minority poverty rates increase. Our demographic analyses confirmed these earlier findings that such factors are important. Thus, we suggest that our results generalize most readily to those jurisdictions whose demographics most closely approximate national averages.

Finally, we used a single case with 1 particular diagnostic aid. How our results might generalize to other decision aids or other types of diagnoses has yet to be determined.

### CONCLUSION

We tested the effects of using a diagnostic decision aid on potential jurors' verdicts in a medical malpractice case. Using or not using a decision aid did not affect the finding of liability for medical malpractice, but using the aid and heeding its advice was protective in the punishment phase of the trial. These data suggest that physicians should not be afraid of using a diagnostic decision aid for fear of increasing their vulnerability to findings of medical malpractice. Using and heeding an aid might be protective in cases in which a finding of medical malpractice occurs. This study thus debunks a potential reason for not using a diagnostic decision aid. In light of the fact that many studies show that aids improve diagnostic decision making, we hope our results may remove a hurdle to more widespread acceptance of such aids.

### REFERENCES

1. Hunt DL, Haynes RB, Hanna SE, Smith K. Effects of computer-based clinical decision support systems on physician performance and patient outcomes. *JAMA*. 1998;280:1339-46.
2. Delaney BC, Fitzmaurice DA, Riaz A, Hobbs FDR. Can computerized decision support systems deliver improved quality in primary care? *BMJ*. 1999;319:1-3.
3. Corey GA, Merenstein JH. Applying the acute ischemic heart disease predictive instrument. *J Fam Pract*. 1987;25:127-33.
4. 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.
5. Ridderikhoff J, van Herk E. A diagnostic support system in general practice: is it feasible? *Int J Med Inform*. 1997;45:133-43.
6. 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.
7. Getty DJ, Pickett RM, D'Orsi CJ, Swets JA. Enhanced interpretation of diagnostic images. *Invest Radiol*. 1988;23:240-52.
8. 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.
9. Fraser HSF, Long WJ, Naimi S. Evaluation of a cardiac diagnostic program in a typical clinical setting. *J Am Med Inform Assoc*. 2003;10:373-81.
10. Friedman CP, Elstein AS, Wolf, FM, et al. Enhancement of clinicians' diagnostic reasoning by computer-based consultation. *JAMA*. 1999;282:1851-6.
11. Reilly BM, Evans AT, Schaidler JJ, Wang Y. Triage of patients with chest pain in the emergency department: a comparative study of physicians' decisions. *Am J Med*. 2002;112:154-6.

12. Wang SJ, Ohno-Machado L, Fraser HSF, Kennedy RL. Using patient-reportable clinical history factors to predict myocardial infarction. *Comput Biol Med.* 2001;31:1–13.
13. Selker HP, Beshansky JR, Griffith JL, et al. Use of the acute cardiac ischemia time-insensitive predictive instrument (ACI-TIPI) to assist with triage of patients with chest pain or other symptoms suggestive of acute cardiac ischemia: a multicenter, controlled clinical trial. *Ann Intern Med.* 1998;129:845–55.
14. Garg AX, Adhikari NK, McDonald H, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. *JAMA.* 2005;293:1223–38.
15. Marsan RJ, Shaver KJ, Sease KL, et al. Evaluation of a clinical decision rule for young adult patients with chest pain. *Acad Emerg Med.* 2005;12:26–31.
16. Bachmann LM, Kolb E, Koller MT, Steurer J, ter Riet G. Accuracy of Ottawa ankle rules to exclude fractures of the ankle and mid-foot: systematic review. *BMJ.* 2003;326:417.
17. Kaplan B. Evaluating informatics applications—clinical decision support systems literature review. *Int J Med Inform.* 2001;64:15–37.
18. Arkes HR, Shaffer VA, Medow MA. Patients derogate physicians who use a computer-assisted diagnostic aid. *Med Decis Making.* 2007;27:187–202.
19. Alvarado A. A practical score for the early diagnosis of acute appendicitis. *Ann Emerg Med.* 1986;15:557–64.
20. Sunstein CR, Kahneman D, Schkade D. Assessing punitive damages. *Yale Law J.* 1998;107:2071–153.
21. Cicchetti DV. Extension of multiple-range tests to interaction tables in the analysis of variance: A rapid approximation. *Psychol Bull.* 1972;77:405–8.
22. 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.
23. Svensson I, Sjostrom B, Haljamae H. Influence of expectations and actual pain experiences on satisfaction with postoperative pain management. *Eur J Pain.* 2001;5:125–33.
24. Myles PS, Williams DL, Hendrate M, et al. Patient satisfaction after anaesthesia and surgery: results of a prospective survey of 10,811 patients. *Br J Anaesth.* 2000;84:6–10.
25. Thomas T, Robinson C, Champion D, et al. Prediction and assessment of the severity of post-operative pain and of satisfaction with management. *Pain.* 1998;75:177–85.
26. Rahmqvist M. Patient satisfaction in relation to age, health status and other background factors: a model for comparisons of care units. *Int J Qual Health Care.* 2001;13:385–90.
27. Finkelstein BS, Singh J, Silvers JB, Neuhauser D, Rosenthal GE. Patient and hospital characteristics associated with patient assessments of hospital obstetrical care. *Med Care.* 1998;36(8 Suppl):AS68–78.
28. Nguyen Thi PL, Briancon S, Empereur F, Guillemin F. Factors determining inpatient satisfaction with care. *Soc Sci Med.* 2002;54:493–504.
29. Jaipal CK, Rosenthal GE. Are older patients more satisfied with hospital care than younger patients? *J Gen Intern Med.* 2003;18:23–30.
30. Hendriks AAJ, Smets EMA, Vrieling MR, Van Es SQ, De Haes JC. Is personality a determinant of patient satisfaction with hospital care? *Int J Qual Health Care.* 2006;18:152–8.
31. Quintana JM, Gonzalez N, Bilbao A, et al. Predictors of patient satisfaction with hospital health care. *BMC Health Serv Res.* 2006;6:102.
32. Murphy RN. Legal and practical impact of clinical practice guidelines on nursing and medical practice. *Nurse Pract.* 1997;22:138, 147–8.
33. Rosoff AJ. And the law: the courts confront clinical practice guidelines. *J Health Polit Polic.* 2001;26:327–68.
34. Pennachio DL. Clinical guidelines sword or shield: you're more likely to be sued for not following guidelines than for following them. But you can never count on what a jury will do. *Health Econ.* 2004;81:22–4.
35. Eagan MC. Bariatric surgery: malpractice risks and risk management guidelines. *Am Surg.* 2005;71:369–75.
36. Hyams AL, Brandenburg JA, Lipsitz SR, et al. Practice guidelines and malpractice legislation: a two-way street. *Ann Intern Med.* 1995;122:450–5.
37. Hyams AL, Shapiro DW, Brennan TA. Medical practice guidelines in malpractice litigation: an early retrospective. *J Health Polit Polic.* 1996;21:289–313.
38. Woolf SH. Practice guidelines: a new reality in medicine. *Arch Intern Med.* 1993;153:2646–55.
39. Douglas CD, Macpherson NE, Davidson PM, Gani JS. Randomised controlled trial of ultrasonography in diagnosis of acute appendicitis, incorporating the Alvarado Score. *BMJ.* 2000;321(7266):919–22.
40. Cruikshank PJ. Patient ratings of doctors using computers. *Soc Sci Med.* 1985;21:615–22.
41. Cabana MD, Rand CS, Powe NR, et al. Why don't physicians follow clinical practice guidelines. *JAMA.* 1999;282:1458–65.
42. Vidmar N, Rice JJ. Assessments of noneconomic damage awards in medical negligence: a comparison of jurors with legal professions. *Iowa L Rev.* 1993;78:883–911.
43. Vidmar N. Empirical evidence on the “deep pockets” hypothesis: jury awards for pain and suffering in medical malpractice cases. *Duke Law J.* 1993;43:217–66.
44. Vidmar N, Lee J, Cohen E, Stewart A. Damage awards and jurors' responsibility ascriptions in medical versus automobile negligence cases. *Behav Sci Law.* 1994;12:149–60.
45. Feigenson N, Park J, Salovey P. Effect of blameworthiness and outcome severity on attributions of responsibility and damage awards in comparative negligence cases. *Law Hum Behav.* 1997;21:597–617.
46. Helland E, Tabarrok A. Race, poverty, and American tort awards: evidence from three data sets. *J Legal Stud.* 2003;32:27–58.