Pediatricians’ Clinical Decision Making

Results of 2 Randomized Controlled Trials of Test Performance Characteristics

Colin M. Sox, MD, MS; Thomas D. Koepsell, MD, MPH; Jason N. Doctor, PhD; Dimitri A. Christakis, MD, MPH

Objective: To determine whether presenting test performance characteristics influences clinical management.

Design: Two questionnaire-based, randomized controlled trials.

Setting: Mailed surveys with 2 clinical vignettes.

Participants: Randomly selected US pediatricians (N=1502).

Intervention: Vignette-specific, randomly assigned test information: no additional information (control), test characteristics (TC), or TC defined. In the pertussis vignette, the TC group received the direct fluorescent antibody test’s sensitivity and specificity, and the TC defined group received the same information with definitions. In the urinalysis vignette, the TC group received the false-positive rate of persistent microhematuria in predicting renal disease, and the TC defined group received a definition of this information.

Main Outcome Measures: In the pertussis vignette, diagnostic test choice and management of erythromycin therapy and hospital discharge plans. In the urinalysis vignette, serum laboratory testing and nephrology referral plans.

Results: Six hundred fifty-three participants (49.5% of those eligible) returned completed surveys. In the pertussis vignette, significantly more of the TC (73%) and TC defined (71%) groups ordered the best-performing test than did controls (21%) (P<.001 for both comparisons). Receiving test characteristics did not significantly affect erythromycin therapy or hospital discharge plans (P=.40). In the urinalysis vignette, the TC defined group referred to nephrology (30%) and checked laboratory tests (88%) significantly more often than did controls (19%, P=.01; 78%, P=.01, respectively), but the TC and control groups’ testing and referral plans did not differ significantly (22% vs 19%, P=.36; 75% vs 78%, P=.48, respectively).

Conclusion: Providing test performance characteristics influenced certain clinical decisions, sometimes in unexpected ways.

Arch Pediatr Adolesc Med. 2006;160:487-492

The Institute of Medicine’s report To Err Is Human prioritized improvement in the quality of health care and increased awareness of its importance. Although improving care can be difficult, advances in information technology, such as electronic medical records and computerized provider order entry, have provided useful tools for this task. Decision support linked to computer provider order entry has been shown to improve many types of care, including medication prescribing, guideline compliance, and diagnostic test ordering. The health information technology revolution under way in the United States has been influenced by requirements to report quality, incentives to improve billing efficiencies, and financial support from the Department of Health and Human Services. It is only a matter of time before health information technology tools become more widely available.

As outlined in Medical Decision Making, physicians approach clinical problems in a stepwise fashion. A list of hypotheses (differential diagnosis) is generated, the hypotheses are tested, the test results evaluated and interpreted, and a course of action is chosen. During the hypothesis-testing phase of medical decision making, appropriate diagnostic test ordering is important, as excessive use of diagnostic tests is an expensive problem in medicine. Physicians’ discomfort with uncertainty sometimes motivates them to order tests for defensive reasons. Interventions targeting physician behavioral factors and using quality improvement strategies have had some success at improving test ordering.
The format used to present test information can independently influence the hypothesis-testing phase. In a randomized controlled trial, general practitioners were given a clinical vignette with test results presented in 1 of 3 ways (test result only, test result and test sensitivity and specificity, or test result and the positive likelihood ratio presented in plain language). Physicians who received the test result and the positive likelihood ratio in plain language more accurately estimated the patient's posttest probability of disease. Although this study evaluated the hypothesis-testing phase of medical decision making, the influence of the format of presentation of test characteristics on physicians' choice of tests or patient management has not been rigorously tested, to our knowledge. While decision support linked to computerized provider order entry has reduced unnecessary serologic and radiologic tests ordered among adult patients, the influence of providing test performance with test results on pediatricians' subsequent clinical management has not been studied.

Our study examined 2 sets of research questions to address our hypothesis that presenting general pediatricians with information about test accuracy would influence patient management. Our first research question sought to determine whether presenting pediatricians with the sensitivity and specificity of direct fluorescent antibody (DFA) testing for pertussis would influence their initial choice of a diagnostic test and subsequent patient management. We hypothesized that pediatricians receiving test performance characteristics would choose the test with the highest sensitivity and specificity and would choose an appropriate management plan. Our second research question sought to determine whether presenting pediatricians with the high false-positive rate of persistent microhematuria detected on screening urinalysis (96%) would influence their subsequent patient evaluation. We hypothesized that subjects receiving test performance information would be less likely to evaluate the patient further.

**METHODS**

**STUDY DESIGN**

We implemented 2 questionnaire-based, randomized controlled trials in a mailed survey that asked pediatricians to manage 2 clinical vignettes. The questionnaire required approximately 7 minutes to complete. For each of 2 vignettes, we randomized subjects in a factorial design into 1 of 3 vignette-specific test characteristics groups (Figure 1). Therefore, each subject received 1 of 9 versions of the questionnaire. We randomized subjects using a computerized random-number generator. Subjects were blinded to randomization, and allocation was concealed.

In February 2002, we mailed potential subjects a 4-page questionnaire, a prepaid return envelope, a dollar bill as a token of appreciation, and a letter stating that participation was voluntary and that responses would be kept confidential. Subjects were given the opportunity to decline to participate. We sent nonresponders a replacement survey at 4-week intervals, for a maximum of 4 mailings per subject, and sent a reminder postcard after the second mailing. We contacted the institutions of nonresponders to confirm that the mailing addresses were correct. The University of Washington Institutional Review Board approved this study.

**SUBJECTS**

Our target study population was a random selection of pediatricians practicing general pediatrics in the United States. We recruited our sample from the 2002 American Medical Association Masterfile of licensed physicians in the United States. Compiled for the US Department of Defense, this list is considered a complete collection of physicians' names, addresses, specialty, age, and sex and is not limited to members of the American Medical Association. From among 44 561 pediatricians without listed secondary specialties, 1502 were randomly selected for participation. To power the study to detect a 15% difference in management choices between test performance information groups and assuming a 50% response rate, we chose our sample size to ensure at least 175 participants in each group.

**QUESTIONNAIRE VIGNETTES**

We developed the questionnaire and pilot tested it among 22 nonstudy pediatricians. We will refer to the questionnaire's 2 clinical vignettes as "the pertussis vignette" and "the urinalysis vignette." For each vignette, subjects were presented with a diagnostic test result followed by their randomly assigned, vignette-specific information about test performance. Both vignettes had a control group, a TC group, and a TC defined group. The TC groups were presented information that included the scientific terms sensitivity, specificity, or false-positive rate, with the corresponding numerical values as frequencies (because medical students understand probabilities easier when they are presented as frequencies). The TC defined groups received the same information defined using nonscientific terminology.

After being provided with test information, subjects were asked to make management decisions about the case. We assumed that subjects would estimate the posttest probability of disease during the hypothesis-testing phase of medical decision-making, before making management decisions. We hypoth-

---

The format used to present test information can independently influence the hypothesis-testing phase. In a randomized controlled trial, general practitioners were given a clinical vignette with test results presented in 1 of 3 ways (test result only, test result and test sensitivity and specificity, or test result and the positive likelihood ratio presented in plain language). Physicians who received the test result and the positive likelihood ratio in plain language more accurately estimated the patient's posttest probability of disease. Although this study evaluated the hypothesis-testing phase of medical decision making, the influence of the format of presentation of test characteristics on physicians' choice of tests or patient management has not been rigorously tested, to our knowledge. While decision support linked to computerized provider order entry has reduced unnecessary serologic and radiologic tests ordered among adult patients, the influence of providing test performance with test results on pediatricians' subsequent clinical management has not been studied.

Our study examined 2 sets of research questions to address our hypothesis that presenting general pediatricians with information about test accuracy would influence patient management. Our first research question sought to determine whether presenting pediatricians with the sensitivity and specificity of direct fluorescent antibody (DFA) testing for pertussis would influence their initial choice of a diagnostic test and subsequent patient management. We hypothesized that pediatricians receiving test performance characteristics would choose the test with the highest sensitivity and specificity and would choose an appropriate management plan. Our second research question sought to determine whether presenting pediatricians with the high false-positive rate of persistent microhematuria detected on screening urinalysis (96%) would influence their subsequent patient evaluation. We hypothesized that subjects receiving test performance information would be less likely to evaluate the patient further.

**METHODS**

**STUDY DESIGN**

We implemented 2 questionnaire-based, randomized controlled trials in a mailed survey that asked pediatricians to manage 2 clinical vignettes. The questionnaire required approximately 7 minutes to complete. For each of 2 vignettes, we randomized subjects in a factorial design into 1 of 3 vignette-specific test characteristics groups (Figure 1). Therefore, each subject received 1 of 9 versions of the questionnaire. We randomized subjects using a computerized random-number generator. Subjects were blinded to randomization, and allocation was concealed.

In February 2002, we mailed potential subjects a 4-page questionnaire, a prepaid return envelope, a dollar bill as a token of appreciation, and a letter stating that participation was voluntary and that responses would be kept confidential. Subjects were given the opportunity to decline to participate. We sent nonresponders a replacement survey at 4-week intervals, for a maximum of 4 mailings per subject, and sent a reminder postcard after the second mailing. We contacted the institutions of nonresponders to confirm that the mailing addresses were correct. The University of Washington Institutional Review Board approved this study.

**SUBJECTS**

Our target study population was a random selection of pediatricians practicing general pediatrics in the United States. We recruited our sample from the 2002 American Medical Association Masterfile of licensed physicians in the United States. Compiled for the US Department of Defense, this list is considered a complete collection of physicians' names, addresses, specialty, age, and sex and is not limited to members of the American Medical Association. From among 44 561 pediatricians without listed secondary specialties, 1502 were randomly selected for participation. To power the study to detect a 15% difference in management choices between test performance information groups and assuming a 50% response rate, we chose our sample size to ensure at least 175 participants in each group.

**QUESTIONNAIRE VIGNETTES**

We developed the questionnaire and pilot tested it among 22 nonstudy pediatricians. We will refer to the questionnaire's 2 clinical vignettes as "the pertussis vignette" and "the urinalysis vignette." For each vignette, subjects were presented with a diagnostic test result followed by their randomly assigned, vignette-specific information about test performance. Both vignettes had a control group, a TC group, and a TC defined group. The TC groups were presented information that included the scientific terms sensitivity, specificity, or false-positive rate, with the corresponding numerical values as frequencies (because medical students understand probabilities easier when they are presented as frequencies). The TC defined groups received the same information defined using nonscientific terminology.

After being provided with test information, subjects were asked to make management decisions about the case. We assumed that subjects would estimate the posttest probability of disease during the hypothesis-testing phase of medical decision-making, before making management decisions. We hypoth-
esized that providing test performance information would aid subjects in their disease probability estimations, because they are needed to make Bayesian calculations of the posttest probability of disease.17

The pertussis vignette presented a 5-month-old girl with perioral cyanosis and a hacking cough; subjects were instructed to assume that the likelihood she had pertussis was 30% (vignette details are available in an appendix from the corresponding author). Subjects were asked to choose one of several diagnostic tests to evaluate her nasopharyngeal secretions for pertussis, including culture, DFA testing, or polymerase chain reaction. Subjects in the control group were presented with no further information and were asked if they “happened to know the sensitivity and specificity” of the 3 tests. The other 2 groups were given the sensitivity and specificity of culture (35% and 100%, respectively), DFA testing (50% and 95%, respectively), and polymerase chain reaction (95% and 99%, respectively) for Bordetella pertussis.24,25 The vignette continued with information that the patient was hospitalized, DFA testing of her sputum was ordered, and treatment with erythromycin was started. Three days later, her symptoms were somewhat improved, and results of the DFA testing were negative for pertussis. Controls received no additional information about this test result, the TC group was presented the sensitivity and specificity of DFA testing (30% and 93%, respectively), while the TC defined group was given the following information: “The pertussis DFA has a sensitivity of 50%, meaning that if 100 patients infected with pertussis were tested with the pertussis DFA, 50 would test positive and 50 would test negative by DFA. Similarly, the pertussis DFA has a specificity of 95%, meaning that if 100 patients who are not infected with pertussis were tested with the DFA for pertussis, 5 would test positive and 95 would test negative.” We then asked subjects to decide whether to continue the patient’s treatment with erythromycin and whether to discharge the patient from the hospital.

The urinalysis vignette presented a healthy 5-year-old boy with persistent microhematuria detected by routine screening urinalysis (vignette details are available in an appendix from the corresponding author). Using data obtained from the literature,26-31 we calculated the false-positive rate of persistent microhematuria detected by screening urinalysis (vignette details are available in an appendix from the corresponding author). Subjects were asked to choose one of several diagnostic tests to evaluate her nasopharyngeal secretions for pertussis, including culture, DFA testing, or polymerase chain reaction. Subjects in the control group were presented with no further information and were asked if they “happened to know the sensitivity and specificity” of the 3 tests. The other 2 groups were given the sensitivity and specificity of culture (35% and 100%, respectively), DFA testing (50% and 95%, respectively), and polymerase chain reaction (95% and 99%, respectively) for Bordetella pertussis.24,25 The vignette continued with information that the patient was hospitalized, DFA testing of her sputum was ordered, and treatment with erythromycin was started. Three days later, her symptoms were somewhat improved, and results of the DFA testing were negative for pertussis. Controls received no additional information about this test result, the TC group was presented the sensitivity and specificity of DFA testing (30% and 93%, respectively), while the TC defined group was given the following information: “The pertussis DFA has a sensitivity of 50%, meaning that if 100 patients infected with pertussis were tested with the pertussis DFA, 50 would test positive and 50 would test negative by DFA. Similarly, the pertussis DFA has a specificity of 95%, meaning that if 100 patients who are not infected with pertussis were tested with the DFA for pertussis, 5 would test positive and 95 would test negative.” We then asked subjects to decide whether to continue the patient’s treatment with erythromycin and whether to discharge the patient from the hospital.

The urinalysis vignette presented a healthy 5-year-old boy with persistent microhematuria detected by routine screening urinalysis (vignette details are available in an appendix from the corresponding author). Using data obtained from the literature,26-31 we calculated the false-positive rate of persistent microscopic hematuria detected by screening urinalysis to be 96%. The control group received the test result only, with no additional information. The TC group was given the following information: “When screening asymptomatic patients with urinary tract infection for diseases whose clinical outcome might be improved with early treatment, persistent microscopic hematuria has an estimated false-positive rate of 96%.” The TC defined group was given the following nontechnical explanation of this information: “The literature suggests that, of 100 asymptomatic children who have persistent microscopic hematuria detected by routine urinalysis, 96 will not have significant nephrologic renal disease for which early treatment might improve clinical outcome.” We then asked subjects if they would refer the patient to a pediatric nephrologist or check his serum electrolyte, urea nitrogen, and creatinine levels.

ADDITIONAL DEMOGRAPHIC DATA

To collect information about subjects’ training, we asked if they were board certified in pediatrics, a US medical school graduate, and currently in a pediatric residency program. We defined foreign medical school graduate as a graduate of a medical school located outside of the United States.

To collect information about clinical practice, we asked subjects to report the percentage of clinical time spent in general pediatrics and whether they worked in a practice that had 1 or more. We defined general pediatrician as one who reported spending more than 80% of clinical time in general pediatrics. We defined small primary practice as a primary practice that included 1 or 2 pediatricians.

MAIN OUTCOME MEASURES

In the pertussis vignette, we assessed subjects’ choice of an initial pertussis diagnostic test and whether they continued the patient’s treatment with erythromycin and discharged the patient from the hospital. In the urinalysis vignette, we assessed whether subjects referred the patient to a pediatric nephrologist and/or checked the patient’s serum electrolyte, urea nitrogen, and creatinine levels.

STATISTICAL ANALYSIS

For each outcome, we performed comparisons between the vignette-specific control group and its TC or TC defined group. We conducted χ² tests to examine the effect of receiving test characteristics alone; TC defined, pediatrics who received test characteristics defined.

RESULTS

STUDY PARTICIPATION

We mailed questionnaires to 1502 physicians identified in the American Medical Association Masterfile as pediatrics without subspecialization. Of these, 106 questionnaires (7%) were returned by the post office without forwarding addresses, and 43 physicians (3%) did not meet inclusion criteria. Of the 1333 potentially eligible subjects, 59 (4%) declined to participate and 653 returned completed surveys, leaving 641 potentially eligible nonresponders (Figure 2), for an estimated response rate32 of 49.75% (653/1313). There were no statistically significant differences by age, sex, or test characteristics group between participants and nonresponders.
For the pertussis vignette, 202 participants were in the control group, 231 in the TC group, and 220 in the TC defined group. For the urinalysis vignette, 208 participants were in the control group, 231 in the TC group, and 214 in the TC defined group (Table 1). The only significant difference in participant characteristics by randomization groups was by sex for the urinalysis vignette ($P = .02$) (Table 1).

### PERTUSSIS VIGNETTE

Providing participants with the sensitivity and specificity of the 3 tests influenced which test they chose, as significantly more of the TC (73%) and TC defined (71%) groups chose the polymerase chain reaction test (the only test for which sensitivity and specificity were ≥95%) than did controls (21%, $P < .001$ in both comparisons) (Table 2). Among the controls, 45% reported that they knew the sensitivity and specificity of pertussis culture, 34% knew the values for DFA testing, and 18% knew the sensitivity and specificity of pertussis culture, 34% knew the values for DFA testing, and 18% knew the values for DFA testing.

The vast majority of participants chose to continue the patient’s erythromycin therapy, and receiving test characteristics had no significant influence on these decisions (90% of the TC group vs 92% of controls, $P = .55$; and 94% of the TC defined group vs 92% of controls, $P = .40$). Similarly, most participants chose to discharge the patient from the hospital, and receiving test characteristics had no significant influence on these decisions (86% of the TC group vs 87% of controls, $P = .48$) (Table 3).

### URINALYSIS VIGNETTE

Significantly more participants in the TC defined group than controls chose to refer the patient to a nephrologist (30% vs 19%, $P = .01$), while there was no significant difference in referral between the TC group and the controls (22% vs 19%, $P = .36$). Significantly more subjects in the TC defined group than controls decided to check the patient’s serum electrolyte, urea nitrogen, and creatinine levels (88% vs 78%, $P = .01$), while there was no significant difference in laboratory test ordering between the TC group and the controls (75% vs 78%, $P = .48$) (Table 3). Controlling for sex in logistic regression analyses had little effect on any of these results.

---

**Table 1. Subject Characteristics by Randomization Groups***

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Overall (n = 653)</th>
<th>Control (n = 202)</th>
<th>TC Group (n = 231)</th>
<th>TC Defined Group (n = 220)</th>
<th>Control (n = 208)</th>
<th>TC Group (n = 231)</th>
<th>TC Defined Group (n = 214)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean ± SD, y</td>
<td>43 ± 11</td>
<td>43 ± 11</td>
<td>43 ± 11</td>
<td>44 ± 11</td>
<td>43 ± 11</td>
<td>43 ± 11</td>
<td>44 ± 11</td>
</tr>
<tr>
<td>Board certified in pediatrics</td>
<td>81</td>
<td>81</td>
<td>82</td>
<td>79</td>
<td>80</td>
<td>79</td>
<td>84</td>
</tr>
<tr>
<td>Foreign medical school graduate</td>
<td>20</td>
<td>19</td>
<td>22</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>19</td>
</tr>
<tr>
<td>General pediatric</td>
<td>80</td>
<td>81</td>
<td>82</td>
<td>76</td>
<td>80</td>
<td>77</td>
<td>84</td>
</tr>
<tr>
<td>Male sex</td>
<td>44</td>
<td>40</td>
<td>43</td>
<td>50</td>
<td>31</td>
<td>45</td>
<td>37</td>
</tr>
<tr>
<td>Small primary practice</td>
<td>22</td>
<td>13</td>
<td>17</td>
<td>17</td>
<td>15</td>
<td>17</td>
<td>14</td>
</tr>
<tr>
<td>Resident</td>
<td>16</td>
<td>22</td>
<td>21</td>
<td>23</td>
<td>22</td>
<td>23</td>
<td>22</td>
</tr>
</tbody>
</table>

Abbreviations: TC, test characteristics.

*Data are given as percentages unless otherwise indicated. The only significant difference between randomization groups was by sex for the urinalysis vignette ($P = .02$).

**Table 2. Effect of Providing Sensitivity and Specificity of the Bordetella pertussis DFA on Pediatricians’ Patient Management**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Control (n = 202)</th>
<th>TC Group (n = 231)</th>
<th>TC Defined Group (n = 220)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continued erythromycin therapy</td>
<td>92</td>
<td>90</td>
<td>94</td>
<td>.55</td>
</tr>
<tr>
<td>Discharged patient from the hospital</td>
<td>87</td>
<td>86</td>
<td>87</td>
<td>.79</td>
</tr>
</tbody>
</table>

Abbreviations: DFA, direct fluorescent antibody (DFA); TC, test characteristics.

*The TC group received the DFA’s sensitivity and specificity; the TC defined group received the same information with nontechnical definitions of these terms.

**Table 3. Effect of Providing False-Positive Rate of Persistent Microhematuria Found on Screening Urinalysis on Pediatricians’ Subsequent Patient Evaluation**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Control (n = 208)</th>
<th>TC Group (n = 231)</th>
<th>TC Defined Group (n = 214)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Referred to pediatric nephrology</td>
<td>19</td>
<td>22</td>
<td>20</td>
<td>.36</td>
</tr>
<tr>
<td>Checked serum electrolyte, urea nitrogen, and creatinine levels</td>
<td>78</td>
<td>75</td>
<td>88</td>
<td>.01</td>
</tr>
</tbody>
</table>

Abbreviations: DFA, direct fluorescent antibody (DFA); TC, test characteristics.

*The TC group received the false-positive rate of persistent microhematuria in predicting renal disease; the TC defined group received a nontechnical definition of this information.
Providing laboratory test performance characteristics influenced pediatricians’ clinical decision making, but to different degrees. In the pertussis vignette, few participants reported familiarity with the performance characteristics of commonly used tests. Providing the sensitivity and specificity of diagnostic tests for pertussis dramatically increased the proportion of participants who chose the best-performing test, but it did not influence their clinical management decisions. The subjects’ post-test probability of pertussis was perhaps higher than their treatment threshold. On the other hand, it is possible that their decisions to continue erythromycin therapy were primarily influenced by its low cost or favorable adverse effect profile. In the urinalysis vignette, receiving test characteristics alone did not significantly influence participants’ decisions to refer the patient or to conduct serum laboratory tests, but receiving test characteristics defined increased the frequency of both evaluations. The direction of this effect contradicted our hypothesis, as we expected that presenting pediatricians with the high false-positive rate of persistent microhematuria would decrease their subsequent workup of the patient’s hematuria. The increased evaluation of microhematuria may have been in part due to the “certainty effect,” as presenting the nonzero probability of renal disease may have instilled concern. This result supports others’ findings that false-positive laboratory results lead to increased resource use.

The format of the test performance information influenced practice to different degrees, as patient management after receiving test characteristics defined differed significantly from that controls for 3 of the 5 measures, while receiving the test characteristics alone significantly affected only 1 measure. Perhaps participants in the TC group did not understand the scientific terminology that was used; the TC defined group was given the same information described without using scientific language. Presenting the test characteristics defined might be difficult to implement using computerized provider order entry, because the time required to read lengthy decision support might limit the influence of the intervention. The test characteristics alone information was more concise, but it did not significantly influence most management decisions.

Several limitations to this study warrant comment. First, physicians who returned completed surveys may be different from those who did not. However, there were no significant differences in age or sex between responders and nonresponders, and our response rate was similar to the mean response rates in published studies. Second, decisions made by participants in our vignettes may not represent their actual clinical practice. However, vignettes have been shown to be a valid tool for measuring the quality of clinical practice, and our primary comparison was between randomized groups. Third, in retrospect it might have been preferable to avoid the use of the term false-positive rate presented to the TC group in the urinalysis vignette. Although some have used this term to refer to 1 minus predictive value positive (as our survey did), it also has been used to refer to 1 minus specificity. A false-positive rate would be termed a false-alarm rate in other disciplines. An occasional participant may have been confused by the potential ambiguity of this term. However, we believe that an interpretation of a false-positive rate of 96% as 1 minus specificity would be implausible because it would imply a specificity of 4%, a value that would essentially render any such test useless for clinical diagnostic purposes.

Providing diagnostic test performance characteristics sometimes influenced pediatricians’ patient management in unexpected ways. Presenting a representative sample of US pediatricians with the same information in different formats influenced their clinical decisions to different degrees. Future research should determine the format for presenting test characteristics that maximally influences physicians’ clinical decision making. In addition, interviews should be performed among pediatricians who do not respond to decision support as expected in an effort to understand why sometimes this information is not influential. In light of the substantial investment that the Department of Health and Human Services recently made to promote the use of health information technology in the United States, our findings reinforce the importance of verifying that point-of-care decision support has the desired effect.

Accepted for Publication: October 27, 2005.

Correspondence: Colin M. Sox, MD, MS, Center for Child Health Care Studies, Department of Ambulatory Care and Prevention, Harvard Medical School, and Harvard Pilgrim Health Care, Inc, 133 Brookline Ave, Sixth Floor, Boston, MA 02215 (colin_sox@hms.harvard.edu).

Funding/Support: This study was supported in part by the Nesholm Family Foundation, Seattle, and by the Clinical Scholars Program (Dr Sox) and the Generalist Physician Faculty Scholars Program (Dr Christakis) of The Robert Wood Johnson Foundation, Princeton, NJ.

disclaimer: The opinions are those of the authors and do not necessarily represent the views of The Robert Wood Johnson Foundation.

Acknowledgment: We acknowledge the helpful feedback received on this project from Alison Galbraith, MD, MPH, and from Hal Sox, MD. We also acknowledge the assistance of the Clinical Scholars Program work-in-progress seminar at the University of Washington.

REFERENCES

8. Christakis DA, Zimmerman FJ, Wright JA, Garrison MM, Rivara FP, Davis RL.


