

*Research Paper* ■

## Automated Evidence-based Critiquing of Orders for Abdominal Radiographs: Impact on Utilization and Appropriateness

LINDA H. HARPOLE, MD, MPH, RAMIN KHORASANI, MD, FRCPC, JULIE FISKIO, GILAD J. KUPERMAN, MD, PhD, DAVID W. BATES, MD, MSc

**Abstract** **Objective:** Inappropriate utilization of diagnostic testing has been well documented. The purpose of this study was to measure the impact of presenting real time, evidence-based critiques about the appropriateness of abdominal radiograph (KUB) orders on physician decision making.

**Design:** Prospective trial where evidence-based critiques were presented to ordering clinicians in two kinds of situations: (1) a KUB was likely to have a low probability of providing useful information, or (2) an alternative view(s) was more appropriate given the clinical circumstance. There were two phases of the trial: Phase 1 was a 9-week period where evidence-based critiques were presented at the time of ordering a KUB, followed by Phase 2, a 19-week period in which orderers were randomized to receive critiques either amended to include both institutional data regarding the utility of the critiques and stronger messages about the lack of utility of the study, or the same critiques as presented in Phase 1, depending upon indication. Based upon the radiologist's report of their interpretation of the exams, the results of the examinations were scored as positive, equivocal, or negative using structured criteria.

**Results:** 299 KUBs in Phase 1 and 385 KUBs in Phase 2 received at least one critique. Cancellation rates of low yield films were low, and were similar in Phase 1 and 2, 8/258 (3%) vs. 10/283 (4%). Compliance with the recommendation for alternative view(s) was higher: 19/104 (38%) in Phase 1 vs. 96/176 (55%) in Phase 2 ( $p = 0.006$ ). The rules differentiated low-yield from non-low-yield films: 5% of low-yield films vs. 20% of non-low-yield films were positive in Phase 2 ( $p < 0.0001$ ). Surgical physicians were less likely to cancel ( $p = 0.07$ ) or to change to the suggested view(s) ( $p < 0.0001$ ) than medical physicians or nurses.

**Conclusions:** The intervention identified clinical situations in which KUBs appeared to have a low clinical yield. In response to evidence-based critiques, providers were reluctant to cancel their order, but were more willing to change to different views. To reduce the number of inappropriate radiographic films, stronger incentives or interventions may be required.

■ *J Am Med Inform Assoc.* 1997;4:511–521.

Inappropriate utilization of diagnostic testing has been well documented.<sup>1–4</sup> In attempts to improve the

appropriateness of testing, multiple methodologies, including audit and feedback, education, rationing,

Affiliations of the authors: Section for Clinical Epidemiology, Division of General Internal Medicine, Department of Medicine (LHH, JF, DWB); Department of Radiology (RK); and the Center for Applied Medical Information Systems Research (GJK, DWB), Brigham and Women's Hospital, Harvard Medical School, Boston, MA.

Supported in part by R01 HS08927 from the Agency for Health Care Policy and Research, Rockville, MD.

Reprint requests: David W. Bates, MD, MSc, Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115. Correspondence: Linda H. Harpole, MD, MPH, Duke University Medical Center, P.O. Box 3228, Durham, NC 27710. E-mail: harpo003@mc.duke.edu

Presented in part at the Annual Meeting of the Society of General Internal Medicine, May 1996, Washington, DC.

Received for publication: 4/7/97; accepted for publication: 7/14/97.

and financial incentives, have been attempted.<sup>1-8</sup> These strategies have enjoyed some success<sup>6,9</sup> but require considerable time and effort to implement. Also, even when these approaches have been successful, their effects have often diminished soon after the interventions were discontinued.<sup>10</sup>

Interventions that improve physician decision making may be most useful when applied at the time of ordering. The advent of physician computer order entry (POE) allows for this possibility.<sup>11-13</sup> In POE, orders are entered directly into an automated information system. The system can require structured ordering, and reminders can be presented at the time orders are written.

To date, studies using computer order entry have evaluated the impact of real-time reminders on the cost of test ordering,<sup>10,14,15</sup> redundant laboratory testing,<sup>16</sup> drug-drug interactions, and drug-lab interventions.<sup>11</sup> However, the effect of real-time, automated critiques on the ordering of radiologic tests has not been reported.

In this study, we presented real-time comments about the appropriateness of abdominal radiographs (KUB) to physicians using a POE system, and we evaluated the impact of these critiques on their decision-making behavior.

Abdominal radiographs were chosen because they are performed frequently yet often provide little information; additionally there is substantial evidence regarding their utility for specific indications.<sup>17-32</sup>

We hypothesized that real-time critiquing during the use of POE could decrease inappropriate KUB ordering, thereby eliminating unnecessary films, as well as improve the usefulness of the test by suggesting alternative view(s) that could provide better clinical information.

## Methods

### Setting

The study was conducted at the Brigham and Women's Hospital (BWH), a 720-bed tertiary care teaching hospital. Its integrated, computerized hospital information system runs on a 3,000-node local area network (LAN) with 486-based personal computers serving as workstations.<sup>11</sup> Within this computerized information system, a POE application system was developed and was implemented in 1993.<sup>11</sup> At BWH, the primary users of POE are housestaff and nurses. At the time of this study, all in-patient orders were entered through the POE system. Most orders are written using menus, and more than 90% are captured in coded form.<sup>11</sup>

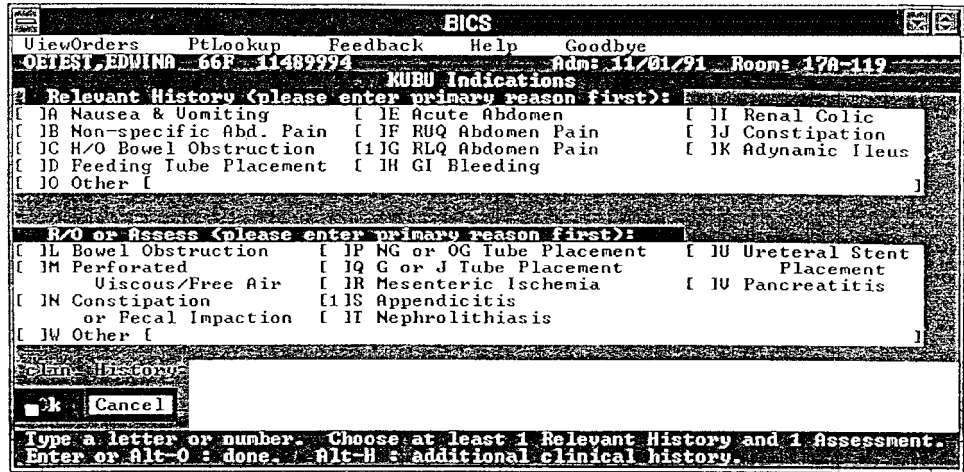
### Capturing the Reason for the Radiograph in Coded Form

To order a radiologic study at our institution, orderers are required to provide the "relevant history items" and items to "rule out or assess." Prior to the study, physicians ordering abdominal radiographs were required to enter the "relevant history items" and items to "rule out or assess" into a free text box. For this study, to permit critiquing, the patient's relevant history as well as items the physician wanted ruled out or assessed when ordering an abdominal radiograph were captured in coded form. To develop lists for these two fields, the free text reasons for a 2-month period were reviewed, put into categories, and ranked. The most common categorical reasons then were used for the structured KUB ordering form, which contained the 11 most commonly entered "relevant history" items and 10 commonly entered "rule out or assess" items. This structured ordering form was then tested and re-ranked before determining the final version (Fig. 1). Physicians were required to indicate at least one item from each section when ordering a KUB. A box for "other," in which free text could be entered if none of the items applied, was included. For all indications physicians could add free text to describe the patient's situation in more detail.

### Development of Critiquing Messages

The literature regarding KUBs was reviewed,<sup>17-32</sup> and lists of appropriate and inappropriate reasons for ordering a KUB were developed. From the common history and assessment items listed on the structured ordering form, eight combinations of history and assessment items that reflected inappropriate reasons for ordering a KUB were determined, based upon the literature and expert opinion in our medicine, surgery, obstetrics/gynecology, and radiology departments. These situations reflected instances in which, given the identified history and assessment items, a KUB was unlikely to add diagnostic information (see Appendix). The POE system was modified so that, when a physician entered certain history and assessment items (Appendix), a screen would be displayed stating that the test was unlikely to yield worthwhile information. For example, if G (right lower quadrant pain) and S (appendicitis) were chosen from the structured KUB ordering form (Fig. 1), a message would be presented to the effect that a KUB would be low yield and that an ultrasound test might be more useful (Fig. 2). When receiving a message that indicated that a film was likely to be of little value, the orderer had the option to cancel the order or to continue despite the critique. We refer to these kinds of messages as "low yield" messages.

**Figure 1** Structured KUB ordering screen. Capturing history/assessment items; user must enter at least one of each.



Two situations in which an alternate view(s) might be superior to a KUB were also identified (Appendix). For example, if a KUB with both upright and supine views is ordered when perforation was suspected, a message would be displayed suggesting that a chest PA view and a KUB (supine alone) film would be preferable, unless the patient was unable to stand, in which case a lateral decubitus film of the abdomen and a KUB (supine alone) film would be recommended (Fig. 3). If both upright and supine views were ordered but neither perforation nor obstruction was suspected, a message would be displayed suggesting that the supine view alone would be sufficient. When a message suggested an alternate view(s), the orderer had the option to change to the suggested view(s), continue as ordered, or to cancel altogether.

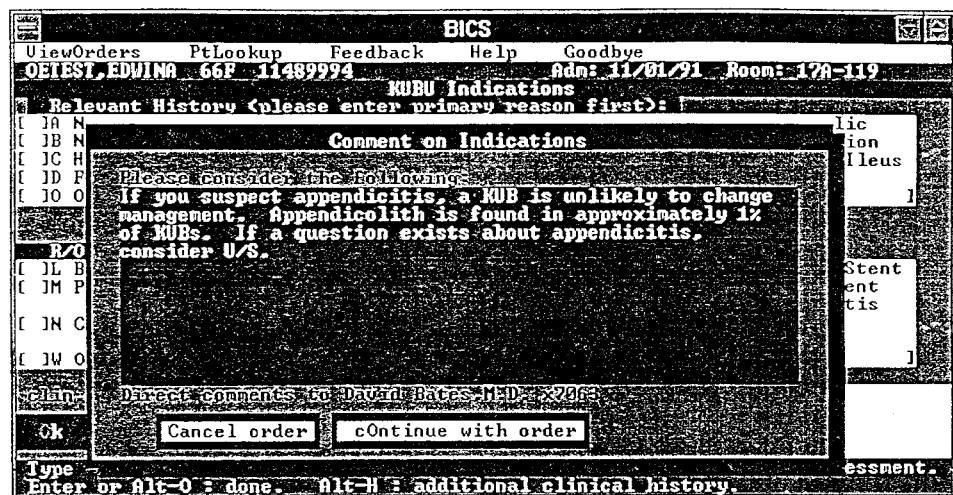
For all exams, instructions directed the orderer to

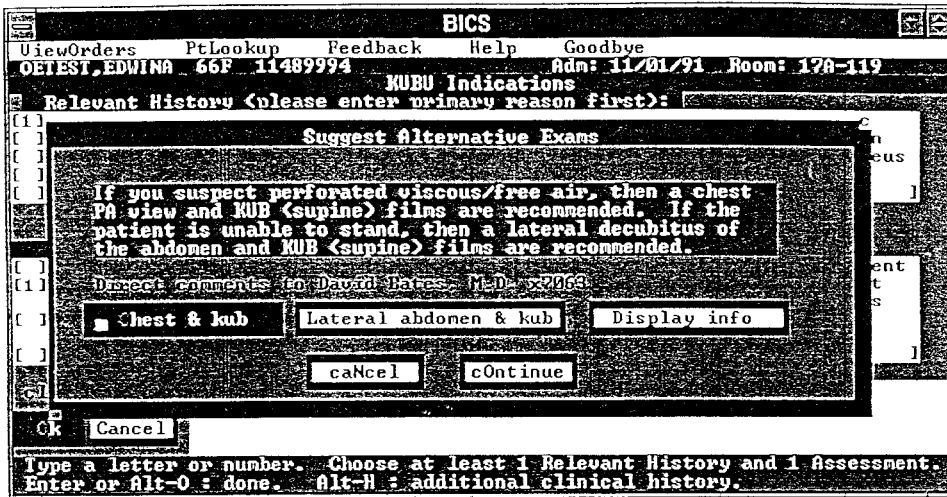
enter the most likely indication for the study first, the second next, and so forth. If more than one selection were chosen by the orderer for the "relevant history" or "item to rule out or assess," the evidence-based message would be triggered from the first item entered only. When available, full abstract(s) from scientific papers supporting the messages were accessible on-line at the time of receiving a message.<sup>18,19,21,22,26-31</sup>

**Intervention**

Prior to implementation of the evidence-based messages, a letter of introduction and an explanation of the intervention were mailed via computer to all order-entry users. The letter was endorsed by the chairmen of the departments of Medicine, Surgery, Obstetrics/Gynecology, and Radiology.

**Figure 2** Example of low-yield critique. The user is notified that KUB is not likely to yield worthwhile information and that ultrasound might be more useful.





**Figure 3** Example of alternate-view critique. In some situations, an alternate view(s) might be superior to KUB.

### Study Period and Outcomes

In Phase 1, from August 1, 1995, to September 30, 1995, all KUB in-patient orders were subject to the evidence-based messages. The main outcome measures were cancellation rates and the frequency with which orders were changed to suggested view(s). A secondary outcome was the yield of important diagnostic information in films that were predicted to be of low utility but were nevertheless obtained. (Films were graded by guided implicit review, described in the following section.)

In Phase 2, based on the findings from Phase 1, the evidence-based messages were amended, and orders were randomized to receive either the same ("control group") or amended ("intervention group") messages. One message (Appendix, low yield rules, no. 5) was supposed to be amended for all future orders, as too many positive films were found when suggestions to cancel were overridden; however, the original rule was inadvertently displayed to the control group of orderers. Other "low yield" messages were amended to include data from the Phase 1 experience, further emphasizing that, for the given indications, KUB was likely to have little value. In addition, one of the alternate exam messages was re-written more emphatically. In Phase 2 of the evaluation of evidence-based messages (November 10, 1995, to March 21, 1996), all users of order-entry were randomized to receive either the initial or amended evidence-based messages to evaluate whether feedback of institutional data and stronger messages might have a larger impact on ordering behavior than delivery of messages without local data alone.

### KUB Film Radiographic Evaluation

The radiologists' transcribed reports of all KUBs were reviewed using a guided implicit approach by one reviewer (LH) and were scored as positive, equivocal, or negative. Specifically, the reviewer graded films as positive if any of the following were present: bowel obstruction, bowel ischemia perforated viscous, free air not following surgery, volvulus, misplaced feeding tube, evaluation of ureteral stent placement, evaluation of barium prior to CT scan, foreign body; as equivocal if one of the following were present: possible ileus, possible ileus vs. small bowel obstruction, moderately dilated loops of bowel, gastric distention, questionable ischemic changes, possible urinary tract stone, possible pneumatosis; or as negative in the presence of none of the above (normal film). To evaluate the consistency of interpretation, a random sample of 50 films were reviewed by a second reviewer using the same criteria. Percent agreement was 94%, and the kappa between the reviewers was 0.80, suggesting good agreement.

### Analysis

Rates of cancellation and change to suggested view(s) were calculated for both Phase 1 and 2 and compared by the chi square test. Rates of cancellation and change to suggested views were also compared within Phase 2 between the "control" and "intervention" groups. Rates of cancellation and change to suggested view(s) for all films were compared by type of orderer (medicine physician, surgery physician, nurse). Characteristics of patients (age, gender, race, clinical service) who received the KUBs were obtained from the computerized information system and compared across the two phases using the chi-square test

Table 1A ■

## Patient Characteristics

Variable	Phase 1 (N = 190)	Phase 2 (N = 491)	P*
Age (mean, range)	58.8 (18–90)	58.5 (16–93)	0.84
Gender			
% Female	91 (48)	267 (54)	0.13
% Male	99 (52)	224 (46)	
Race			0.49
White	149 (78)	366 (74)	
Black	23 (12)	72 (15)	
Hispanic	11 (6)	24 (5)	
Other	7 (4)	29 (6)	
Clinical service			0.03
Medicine	81 (42)	187 (38)	
Surgery	100 (53)	268 (55)	
Obstetrics/gynecology	9 (5)	36 (7)	

\*P—phase 1 vs. phase 2.

for categorical variables, and Student's *t* test for continuous variables. The KUB results for low-yield and non-low-yield films were compared within the two phases by the chi-square test. The potential impact of KUB results upon clinical care for all positive films was evaluated through medical record chart review. The potential impact of canceled KUBs upon clinical care was evaluated through medical record chart review of patients in whom a KUB was canceled. Interrater reliability for KUB film results was determined using the kappa statistic. Time trend analyses of the number of KUBs during a 26-month period was performed using piece-wise regression.<sup>33</sup>

## Results

Overall, 681 patients had 1,244 KUB films performed during the 6-month trial: 190 patients during Phase 1 (380 films) and 491 patients during Phase 2 (864 films). Neither the patient characteristics nor the orderer type differed between the two phases (Tables 1A,1B).

In Phase 1, 79% of KUBs ordered resulted in at least one critique, either low yield or alternate exam, as compared with Phase 2, in which 45% of KUBs ordered resulted in at least one message (Table 2). Between Phase 1 and the start of Phase 2, a rule that had accounted for 32% of low-yield film reminders in Phase 1 was removed for one group of orderers in Phase 2, accounting for the majority of this difference. In Phase 2, there were no differences in the rate of cancellation of low-yield films, change to suggested view(s), or results of low-yield films between the two randomized groups. For this reason, all Phase 2 films are considered together.

Cancel rates for orders receiving a low yield message were low in both Phase 1—8/258 (3%)—and in Phase 2—10/283 (4%)—despite strengthening the messages and improving specificity (Table 2). The results were better for orders receiving an alternate exam message: 38% changed as suggested in Phase 1, versus 55% in Phase 2 ( $p = 0.006$ , Table 2). More specifically, for messages suggesting a KUB supine alone (Appendix, alternate exam rule no. 1), 33% (13/39) of the messages were followed as suggested in Phase 1, versus 40% (29/71) in Phase 2. For messages suggesting upright chest x-ray in addition to supine KUB (Appendix, alternate film rule no. 2), 40% (26/65) were followed in Phase 1 and 64% (67/105) in Phase 2.

The review of the KUB reports (Table 3) showed that the percentage of positive films was significantly lower and the percentage of negative films significantly higher for films that received a low-yield critique as opposed to those that did not in both Phase 1 and Phase 2. After the removal of one rule that included 80% of the positive films in Phase 1, only 3% of the remaining films within the low-yield category of Phase 1 were positive. With removal of this rule from one group of orderers prior to Phase 2 of the trial, 5% of films were positive. Of note, this rule was to have been removed from all orderers in Phase 2, but it inadvertently was displayed to the "control group" of orderers. If this had not occurred, the positive film rate in Phase 2 would have been even lower, approximately 3.5%.

After excluding the rule that was problematic from Phase 1, review of the positive films that resulted from indications prompting a low-yield message revealed a few results of clinical importance (Table 4). For Phase 1, of 3 films which were positive, none had any apparent effect on clinical outcome. In Phase 2, 12 positive films resulted from indications receiving a low-yield message. Of the 12 films, 6 were determined to have a significant effect on clinical outcome, and 6 were determined to have an indeterminate effect. Of these 12 films, 5 resulted from the orderer entering an inaccurate indication for the study, thus resulting in

Table 1B ■

## Provider Characteristics

Variable	Phase 1 (N = 67)	Phase 2 (N = 236)	P*
Provider ordering KUB			0.27
Medicine physician	21 (31)	85 (36)	
Surgical physician	8 (12)	42 (18)	
Nurses/other	38 (57)	109 (46)	

\*P—phase 1 vs. phase 2.

Table 2 ■

## Actions: Phase 1 vs. Phase 2

Variable	Phase 1 (N, % 95% CI)	Phase 2 (N, % 95% CI)
KUB receiving $\geq 1$ critique*	299/380 (79% $\pm$ 4%)	385/864 (45% $\pm$ 3%)
Low-yield KUB canceled	8/258 (3% $\pm$ 2%)	10/283 (4% $\pm$ 2%)
KUB orders changed to suggested* views	39/104 (38% $\pm$ 9%)	96/176 (55% $\pm$ 7%)

\*p &lt; 0.01 Phase 1 vs. Phase 2.

the presentation of a low-yield message that would not have occurred if the orderer had indicated the correct reason for the study. For example, one low-yield critique resulted from the orderer entering “nephrolithiasis” as the reason for ordering a film in a patient who was receiving follow-up KUBs for a known small bowel obstruction. Of the four positive films remain-

Table 3 ■

## Findings of Films: Phase 1 vs. Phase 2

Findings	Phase 1*: N (%)		Phase 2†: N (%)	
	Low-yield Films	Non-Low-yield Films	Low-yield Films	Non-Low-yield Films
Positive	20/225 (9)	22/100(22)	12/255 (5)	101/514(20)
Equivocal	54/225(24)	18/100(18)	55/255(22)	165/514(32)
Negative	151/225(67)	60/100(60)	188/255(73)	248/514(48)

\*p &lt; 0.02 low-yield vs. non-low-yield films, Phase 1.

†p &lt; 0.0001 low-yield vs. non-low-yield films, Phase 2.

ing (excluding the films resulting from mistakenly delivered reminders n = 3), two had a significant effect on clinical management (small bowel obstruction resulting in enema and lactulose, partial colonic obstruction resulting in flexible sigmoidoscopy), and two had indeterminate effects (feeding tube looped in mid-esophagus but no further radiologic evaluation;

Table 4 ■

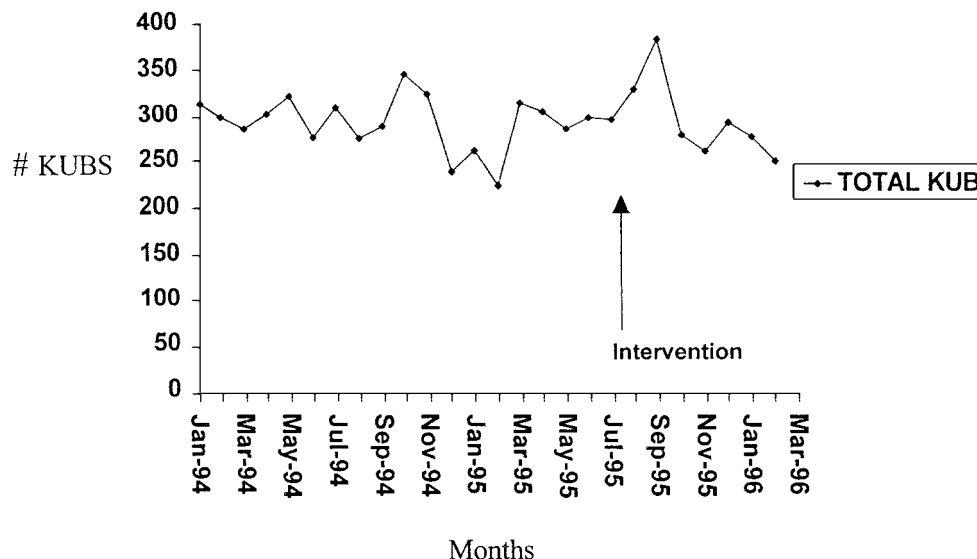
## Positive Findings in Low-yield Films

Phase	Finding	Clinical Outcome	Effect	Comment
1	Improving small bowel obstruction	No further abdominal films or problems during hospitalization	None	—
1	Thumbprinting	No change in management; follow-up KUB 1 week later without evidence of thumbprinting	None	—
1	Thickened sigmoid colon	Abdominal CT performed following day without evidence of pneumatosis; notable for colitis	None	—
2	Small bowel obstruction	Improvement in distention; treatment with lactulose	Positive	Inaccurate reason entered
2	Feeding tube looped in mid-esophagus	No further films for feeding tube placement	Indeterminate	
2	Small bowel obstruction	Enema and lactulose, improvement in film findings	Positive	
2	Colonic obstruction, unchanged	Clinical improvement, patient discharged without further films	Indeterminate	Inaccurate reason entered
2	Barium throughout colon (pre-CT)	Abdominal CT postponed for 10 days due to retained barium	Positive	Inaccurate reason entered
2	Partial small bowel obstruction	Patient improved, then worsened, requiring laparoscopy	Positive	Rule malfunction*
2	Small bowel obstruction, unchanged	Resolution with conservative therapy	Indeterminate	Rule malfunction
2	Worsening of small bowel obstruction	Clinical improvement with conservative therapy	Indeterminate	Inaccurate reason entered
2	Incomplete small bowel obstruction	Nasogastric tube already in place, improvement with conservative therapy	Indeterminate	Rule malfunction
2	Partial colonic obstruction	Flexible sigmoidoscopy performed next day	Positive	—
2	Feeding tube at gastro-esophageal junction	No further films	Indeterminate	—
2	Partial small bowel obstruction	Colonoscopy and decompression from below	Positive	Inaccurate reason entered

Includes all positive low yield films Phase 2, remaining positive films Phase 1 after removal of offending rule.

\*Description of rule malfunction: A rule identified in Phase 1 that was associated with too many positive film findings when overridden was to have been removed for all orderers in Phase 2; however, the “control group” continued to receive this rule.

**Figure 4** KUB frequencies over time (January 1994–February 1996) to evaluate whether there was a time series effect associated with the intervention. This was evaluated using piecewise regression. No significant effect was found.



feeding tube at gastroesophageal junction but no further radiologic evaluation). For films scored as equivocal, 87% were read as possible ileus, 9% as possible stone in the urinary system, and 4% with other questionable findings (possible ischemic changes, possible splenic enlargement).

Low-yield film critiques were more likely to be received by medicine physicians and nurses than by surgical physicians for KUBs ordered (Table 5). In addition, alternative view(s) critiques were also more often received by medicine physicians and nurses than by surgical physicians. However, for orders receiving low-yield critiques, medicine physicians and nurses more often canceled the KUB than did surgical physicians, although this trend did not reach statistical significance ( $p = 0.07$ ). Additionally, for orders receiving alternative view(s) critiques, suggestions to switch to an alternative view were more often adhered to by medicine physicians and nurses than by surgical physicians. The results of films performed, in spite of a low-yield critique, did not differ significantly among nurses, medicine, and surgical physicians ( $p = 0.65$ ).

Because physicians might learn over time not to order

examinations that are low yield, we evaluated the frequency with which KUBs were performed during the study period (Fig. 4) as compared with an 18-month period prior to the intervention period. There was no significant difference in the number of KUBs performed per month in the pre- as compared with the post-intervention period ( $p = 0.19$ ). Similar results were found comparing the number of KUBs per admission per month between the two periods.

With the current results, annual charge savings of only \$6,000 of a potential \$98,500 were realized. This estimate is based upon a 4% cancellation rate of low-yield film orders (50% KUB one-view orders only—\$83 charge—and 50% KUB two-view orders—\$138 charge), and a 40% adherence to the critique to change from two KUB views to one (\$55 charge saving per order).

### Discussion

In this study, we found that orderers were reluctant to cancel radiographs that were likely to be of low yield when presented with evidence-based messages

Table 5 ■

#### Response to Critique by Type of Provider

Type of Provider	No. KUBs Ordered Receiving* Low-yield Critique (%)	No. KUBs Ordered Receiving* Alternate-view Critique (%)	No. Low-yield KUBs Canceled (%)	No. KUB Orders Changed to Suggested* View(s) (%)
Medicine	189/337 (56)	120/337 (36)	9/189 (5)	75/120 (63)
Surgery	205/466 (44)	85/466 (18)	3/205 (1)	26/85 (31)
Nursing	131/231 (57)	69/231 (30)	8/131 (6)	33/69 (48)

\* $p < 0.001$  medicine physicians versus surgery physicians versus nursing.

in real time. This was the case even though review of the results of the films demonstrated that our rules for identifying low-yield indications were valid. Orderers were, however, more amenable to changing to the suggested view(s) when presented with suggestions to change the ordered test to an alternate examination. Thus, providers were willing to substitute but not forgo imaging once the decision to order a KUB had been made.

Changing physician behavior has been difficult,<sup>34</sup> and efforts to improve decision-making behavior have met with a variety of results. A number of studies attempting to decrease inappropriate resource utilization have been unsuccessful, despite using a variety of methodologies, including education and feedback,<sup>3</sup> education and audit,<sup>1</sup> risk estimates, and triage recommendations.<sup>35</sup> Interventions developed to improve the quality of care by suggesting additional and/or alternative testing have been more effective,<sup>15,36-41</sup> compatible with the findings of this study that providers were more willing to substitute for or change the number of view(s) of KUBs than to cancel a planned examination.

There are several possible reasons why the orderers were unwilling to cancel the requested exams. First, the orderers in this study were residents, who were not directly responsible for the financial consequences of their actions. Second, the orderers may not have been the decision makers (e.g., 22% of KUB orders entered by nurses) but merely individuals carrying out the commands of more senior housestaff or attending physicians. They may not have felt that they were in a position to act independently upon the critiques. Third, failure may have been due to the mechanism by which the intervention was introduced. Except for a brief letter of introduction prior to turning on the intervention, no formal, direct education with the housestaff occurred, and critiques regarding the utility of KUBs for selected indications were presented by computer only. Perhaps if the housestaff had received educational lectures on the utility of KUBs in addition to the information provided on-line, a greater impact would have been realized. Alternatively, providing individualized retrospective data to the housestaff about their KUB-ordering behavior and its yield might have been helpful, a technique that has been variably successful.<sup>1,3,9,14</sup> Fourth, the limited impact of the intervention could have been due in part to users' distrust regarding the validity of the critiques because of the display of a rule that, although later removed, initially had poor predictive value. However, it is unlikely that this played a large role, as there were few encounters with the KUB critiques per user, and thus it is unlikely that they "learned" to

distrust the intervention. Involving the housestaff in the development of the criteria, another potentially successful technique,<sup>6</sup> could also have been helpful in getting the housestaff to "buy in" to the validity of the evidence-based critiques.

Although the rules from which the critiques were developed successfully differentiated low-yield from non-low-yield KUBs, the intervention had only a small effect on utilization. The implementation of rules, or "guidelines," can be difficult; guidelines can be perceived to be of varying quality, ranging from clearly defined and proven rules to unproved suggestions.<sup>42</sup> In addition, physicians may be hesitant to make a decision based upon guideline recommendations, particularly if they are concerned about the medical-legal implications of following a guideline that could result in an unfavorable clinical result. This concern has some validity, since guidelines have been used to incriminate as well as exonerate physicians, although generally physicians are on firmer ground when they follow guidelines.<sup>43</sup> Additionally, although the rules utilized here were tested for their ability to identify low-yield utilization during the study and appeared to be effective, including data about their local performance in Phase 2 was no more successful at changing ordering behavior than the presentation of the literature-based critiques in Phase 1. Stronger interventions, such as requiring a radiology consultation of examinations likely to be of low yield or giving individual house officers' data on their respective ordering behavior to their superiors could have had a larger impact, but they would be more intrusive.

If orderers had canceled all low-yield films as suggested, two positive films that significantly affected clinical management would have resulted (0.8% of all low-yield films, excluding those that resulted from the orderer entering an incorrect reason for the KUB). Although it was our impression that films that were scored as equivocal would not result in a change in clinical management, it is possible that not performing these films could have affected patient outcomes. However, none of these films appeared to provide a definitive answer to a clinical question. Further diagnostic testing or reliance on clinical history and exam would be required, suggesting that the KUB was of little clinical benefit to the patient in these circumstances. In addition, there were 18 KUBs canceled in total; upon review of the medical record, none of these cancellations appeared to affect clinical management.

Interestingly, there were significant differences in ordering behavior between medicine and surgical physicians. Although surgical physicians were more resis-



tant to suggestions to cancel and to change their orders, they were also less likely to receive a low-yield or alternate exam critique, suggesting that either their clinical judgment as to when a KUB study was necessary was better than the other orderers, or that the prevalence of disease that a KUB would detect was higher in their population. Different strategies may be necessary to affect behavior for different provider and patient groups. Future studies could focus on this as well as on other important areas, including user satisfaction with the decision-support systems. A satisfaction study performed at our institution prior to this study found high levels of satisfaction overall with the POE system,<sup>44</sup> and a follow-up questionnaire found an average satisfaction score of 3.5 on a scale of 1–7 for suggestions regarding low-yield studies (unpublished data).

This study has several limitations. For one, it is not clear how the results will generalize to other settings. For example, if the KUB decision-support system were instituted in a setting where incentives were better aligned with its purpose—in a more mature managed-care environment, for instance, or one in which the orderers felt the financial impact of their actions—the results might have been more favorable. Also, while the financial savings were small, they could have been larger had the intervention been implemented in emergency and outpatient units, where the literature suggests the yield of KUBs may be lower than in the in-patient setting.<sup>26</sup>

We conclude that an intervention to identify low-yield KUBs was effective for identifying low-yield examinations, although simply presenting a computerized critique was rarely sufficient to convince providers to cancel orders. Providers were, however, more willing to “change direction”—that is, to order more appropriate views. To substantially reduce the number of inappropriate radiographic exams, stronger incentives or interventions may be required. Computers offer the opportunity to give patient-specific, real-time feedback to providers. This will undoubtedly prove an important tool for changing behavior; we are still learning how best to use it.

#### References ■

- Schroeder SA, Kenders K, Cooper JK, Piemme TE. Use of laboratory tests and pharmaceuticals: variation among physicians and effect of cost audit on subsequent use. *JAMA*. 1973;225:969–73.
- Eisenberg JM, Williams SV, Garner L, Viale R, Smits H. Computer-based audit to detect and correct overutilization of laboratory tests. *Med Care*. 1977;15:915–21.
- Williams SV, Eisenberg JM. A controlled trial to decrease the unnecessary use of diagnostic tests. *J Gen Intern Med*. 1986; 1:8–13.
- Dixon RH, Laszle J. Utilization of clinical chemistry services by medical house staff. An analysis. *Arch Intern Med*. 1974; 134:1064–7.
- Tierney WM, McDonald CJ, Hui SL, Martin DK. Computer predictions of abnormal results: effects on outpatient testing. *JAMA*. 1988;259:1194–8.
- Spiegel JS, Shapiro MF, Berman B, Greenfield S. Changing physician test ordering in a university hospital: an intervention of physician participation, explicit criteria, and feedback. *Arch Intern Med*. 1989;149:549–53.
- Martin AR, Wolf MA, Thibodeau LA, Dzau V, Braunwald E. A trial of two strategies to modify the test-ordering behavior of medical residents. *N Engl J Med*. 1980;303:1330–6.
- Eisenberg JM, Williams SV. Cost containment and changing physicians' practice behavior. Can the fox learn to guard the chicken coop? *JAMA*. 1981;246:2195–201.
- Marton KE, Tul V, Soc HC Jr. Modifying test-ordering behavior in the outpatient medical clinic: a controlled trial of two educational interventions. *Arch Intern Med*. 1985;145: 816–21.
- Tierney WM, Miller ME, McDonald C. The effect on test ordering of informing physicians of the charges for outpatient diagnostic tests. *N Engl J Med*. 1990;322:1499–504.
- Bates DW, Kuperman G, Teich J. Computerized physician order enter and quality of care. *Qual Man Health Care*. 1994;2:18–27.
- Sittig DF, Stead WW. Computer-based physician order entry: the state of the art. *J Am Med Inform Assoc*. 1994;1:108–23.
- Tierney WM, Miller ME, Overhage JM, McDonald CJ. Physician inpatient order writing on microcomputer workstations: effects on resource utilization. *JAMA*. 1993;269:379–83.
- Pugh JA, Frazier LM, DeLong E, Wallace AG, Ellenbogen P, Linfors E. Effect of daily charge feedback on inpatient charges and physician knowledge and behavior. *Arch Intern Med*. 1989;149:426–9.
- Bates DW, Kuperman G, Jha A, et al. Effect of displaying charges for clinical laboratory tests on inpatient ordering. *J Gen Intern Med*. 1995;10(S2):61.
- Bates DW, Kuperman GJ, Rittenberg E, et al. Reminders for redundant tests: results of a randomized controlled trial. *Symp Comp Appl Med Care*. 1995;19:935.
- Stower MJ, Mikulin T, Hardcastle JD, Amar SS, Kean DM. Evaluation of the plain abdominal x-ray in the acute abdomen. *J Royal Soc Med*. 1985;78:630–3.
- Greene CS. Indications for plain abdominal radiography in the emergency department. *Ann Emerg Med*. 1986;15:257–60.
- Simpson A, Sandeman D, Nixon SJ, Goulbourne IA, Grieve DC, MacIntyre IMC. The value of an erect abdominal radiograph in the diagnosis of intestinal obstruction. *Clin Radiol*. 1985;36:41–2.
- Brewer RJ, Golden GT, Hitch DC, Rudolf LE, Wangenstein SL. Abdominal pain: an analysis of 1,000 consecutive cases in a university hospital emergency room. *Am J Surg*. 1976; 131:219–23.
- Field S, Guy PJ, Upsdell SM, Scourfield AE. The erect abdominal radiograph in the acute abdomen: should its routine use be abandoned? *BMJ*. 1985;290:1934–6.
- Mutgi A, Williams JW, Nettleman M. Renal colic: utility of the plain abdominal roentgenogram. *Arch Intern Med*. 1991; 151:1589–92.

23. Plewa MC. Emergency abdominal radiography. *Emerg Med Clin NA* 1991;9:827-53.
24. Levine MS. Plain film diagnosis of the acute abdomen. *Emerg Med Clinic NA*. 1985;3:541-62.
25. Goldberg B. Department of inappropriate investigations. *BMJ*. 1977;2:1274-5.
26. Eisenberg RL, Heineken P, Hedgcock MW, Federle M, Goldberg HI. Evaluation of plain abdominal radiographs in the diagnosis of abdominal pain. *Ann Intern Med*. 1982;97:257-61.
27. McCook TA, Ravin CE, Rice RP. Abdominal radiography in the emergency department: a prospective analysis. *Ann Emerg Med*. 1982;11:7-8.
28. Campbell JPM, Gunn AA. Plain abdominal radiographs and acute abdominal pain. *Br J Surg*. 1988;75:554-6.
29. Hayward MWJ, Hayward C, Ennis WP, Roberts CJ. A pilot evaluation of radiography of the acute abdomen. *Clin Radiol*. 1984;35:289-91.
30. Lacey GJ, Wignall BK, Bradbooke S, Reidy J, Hussain S, Cramer B. Rationalizing abdominal radiography in the accident and emergency department. *Clin Radiol*. 1980;31:453-5.
31. Hatfield E, MacMahon H, Hatfield M. Efficacy of plain abdominal radiography in patients with gastrointestinal bleeding. *Ann Emerg Med*. 1988;17:903-5.
32. Boyes RJ, Kruse JA. Nasogastric and nasoenteric intubation. *Crit Care Clinics*. 1992;4:865-77.
33. Gillings E, Makuc D, Siegel A. Analysis of interrupted time series mortality trends: an example to evaluate regionalized perinatal care. *Am J Public Health*. 1981;71:38-46.
34. Eisenberg JM. Doctors' Decisions and the Cost of Medical Care. Ann Arbor: Health Administration Press, 1986;99-124.
35. Lee TH, Pearson SD, Johnson PA, et al. Failure of information as an intervention to modify clinical management. *Ann Intern Med*. 1995;122:434-7.
36. Soumerai SB, Avorn J, Taylor WC, et al. Improving choice of prescribed antibiotics through concurrent reminders in an educational order form. *Med Care*. 1993;31:552-8.
37. McDonald CJ, Wilson GA, McCabe GP. Physician response to computer reminders. *JAMA*. 1980;244:1579-81.
38. McDonald CJ. Use of a computer to detect and respond to clinical events: its effect on clinician behavior. *Ann Intern Med*. 1976;84:162-7.
39. McPhee SJ, Bird JA, Fordham D, Rodnick JE, Osborn EH. Promoting cancer prevention activities by primary care physicians. *JAMA*. 1991;266:538-44.
40. McDonald CJ, Hui SL, Smith DM, et al. Reminders to physicians from an introspective computer medical record: a two-year randomized trial. *Ann Intern Med*. 1984;100:130-8.
41. Litzelman DK, Dittus RS, Miller ME, Tierney WM. Requiring physicians to respond to computerized reminders improves their compliance with preventive care protocols. *J Gen Intern Med*. 1993;8:311-7.
42. McDonald CJ, Overhage JM. Guidelines you can follow and can trust: an ideal and an example. *JAMA*. 1994;271:872-3.
43. Hyams AL, Brandenburg JA, Lipsitz SR, Shapiro DW, Brennan TA. Practice guidelines and malpractice litigation: a two-way street. *Ann Intern Med*. 1995;122:450-5.
44. Lee F, Teich JM, Spurr CD, Bates DW. Implementation of physician order entry: user satisfaction and self-reported usage patterns. *J Am Med Inform Assoc*. 1996;3:42-55.

## APPENDIX

Indication Entered:	Initial Message	Amended Message
Low-yield Rules		
(1) Appendicitis:	If you suspect appendicitis, a KUB is of low yield and is unlikely to change management. Appendicolith is found in approximately 1% of KUBs. If a question exists about appendicitis, consider U/S.	No change
(2) Renal Colic	If a patient has active renal colic, then IVP is generally a superior test.	No change
(3) Pancreatitis	If a patient has pancreatitis or is suspected of having pancreatitis, a KUB is unlikely to change management. If considering etiology of pancreatitis, consider abdominal U/S. If evaluating complications from pancreatitis, consider abdominal CT.	If the patient has pancreatitis or is suspected of having pancreatitis, a KUB is very unlikely to affect your management of the patient. Of all films done at this hospital for pancreatitis in the past 2 months, NONE changed management, and all were negative. If considering etiology of pancreatitis, consider abdominal U/S. If evaluating complications from pancreatitis, consider abdominal CT.

## APPENDIX (continued)

Indication Entered:	Initial Message	Amended Message
(4) Constipation	A KUB is unlikely to change management in constipation unless you're evaluating for severe impaction or bowel obstruction.	A KUB is very unlikely to affect your management of the patient with constipation unless you are truly concerned about severe impaction or bowel obstruction. Of the past 24 films obtained at this hospital to assess constipation or fecal impaction, only 1 was suggestive of impaction. No films demonstrated bowel obstruction.
(5) Nausea and vomiting and/or non-specific abdominal pain	If patient with nausea and vomiting and/or non-specific abdominal pain, KUB is very low yield. Studies show at most a 2% incidence of helpful findings for this indication. Also, KUB is not helpful in gastroenteritis.	If a patient has non-specific abdominal pain, a KUB is of very low yield. Studies show at most a 2% incidence of helpful findings for this indication. Of all studies done at this hospital in the past 2 months for non-specific abdominal pain NONE demonstrated a positive finding.
(6) Gastrointestinal bleeding	If a patient has GI bleeding, a KUB is not indicated to look for a cause of the bleeding. Studies suggest that there is less than a 1% chance of demonstrating a relevant abnormality causing the bleeding.	If a patient has GI bleeding, a KUB is not indicated to look for a cause of the bleeding. Studies suggest that there is less than a 1% chance of demonstrating a relevant abnormality causing the bleeding. Of all KUBs done at this hospital in the past 2 months for GI bleeding, NO relevant findings were found.
(7) Mesenteric ischemia	If you suspect mesenteric ischemia, a KUB is unlikely to change management. You should consider CT or angiography. If questions, consider discussing with radiology.	If you suspect mesenteric ischemia, a KUB is unlikely to change management. You should consider CT or angiography. If questions, consider discussing with radiology. Of all films done at this hospital in the past 2 months for mesenteric ischemia, NONE changed management or had a significant finding.
(8) Nasogastric tube placement	If obtaining this test for placement of an NG tube (NOT small-bore feeding tube), then if able to auscultate and aspirate bowel contents from NG tube, little additional information is obtained from KUB.	A film for the placement of an NG tube (NOT small-bore feeding tube) is only indicated if the tube's position cannot be verified by auscultation and/or aspiration of bowel contents. Did you auscultate? Did you aspirate? If the answer is yes to both of these questions and you still cannot verify position, only then is a KUB (supine) indicated. Of note, of all KUBs obtained for NG tube placement in the past 2 months at this hospital, NONE changed management.
<b>Alternate Film Rules</b>		
(1) KUB vs. KUB and upright	If you do not strongly suspect bowel obstruction or perforated viscus, KUB (supine) alone is recommended.	If you do not strongly suspect bowel obstruction or perforated viscus, KUB (supine) alone is recommended. In general, two views are rarely necessary, except for perforated viscus.
(2) Upright CXR	If suspected perforated viscus/free air, recommend Upright chest film OR If patient unable to stand, L lateral decubitus film	No change