

## Appropriate Use of Abdominal Ultrasonography in Two Hospitals

Isaac W. Hammond, MD, PhD, MPH; David M. Grossman, MD; Julie P. Jones, BS; Helen Hull-Sanders, BS; Philip W. Lundy, DO; and William C. Steinmann, MD, MSc

### Abstract

Criteria for judging the appropriate use of a diagnostic test were applied to all abdominal ultrasounds ordered between January 3, 1994, and February 2, 1994, at the Tulane University Hospital (TMC) and the Veterans Administration Medical Center of New Orleans (VA). The criteria were based on answers to the following three-step sequence of questions. Step A: (1) Are the history and physical examination adequately documented? (2) Is the differential diagnosis supported? Step B: (3) Does the diagnostic study provide discriminatory or confirmatory information regarding the differential diagnosis? (4) Is it required prior to initial therapeutic intervention? Step C: (5) Are the results of the diagnostic study reviewed? (6) Do these results direct the therapeutic course? Two hundred seventy-five abdominal ultrasounds were performed in the two hospitals. Only 160 patients had complete medical information for analyses, and only 73 (45.6%) of the 160 abdominal ultrasounds were judged to be appropriate. Using the chi square test of independence with an expected appropriateness of 80%, we found a statistically significant deviation from the expected  $\chi^2=51.05$ ,  $df=2$ ,  $P<0.0001$ . At TMC, 134 abdominal ultrasounds were performed during the study period. Eighty-one patients were included in the analysis with complete information, and 26 of the 81 abdominal ultrasounds were judged to be appropriate. The chi square test was statistically significant ( $\chi^2=23.23$ ,  $df=1$ ,  $P<0.0001$ ). At the VA, 141 abdominal ultrasounds were performed during the study period. Seventy-nine patients who had abdominal ultrasound and whose charts were available were reviewed. Of the 79 abdominal ultrasounds, 47 (59.5%) were judged to be appropriate. Chi square test results were  $\chi^2=4.13$ ,  $df=1$ ,  $0.025<P<0.05$ . A large percentage of abdominal ultrasounds did not meet criteria for appropriateness. We estimated that \$1.102 million is spent on abdominal ultrasounds per year in the two hospitals, of which more than \$0.5 million (\$599,597) was spent on inappropriate ab-

dominal ultrasound. The adoption of such criteria, while not tested, would likely reduce unnecessary testing, reduce costs, ensure that test performance is optimized, and improve physician documentation.

(*Am J Man Care* 1995;1:271-275)

Abdominal ultrasonography is used in the evaluation of a wide variety of abdominal disorders. As a diagnostic modality, abdominal ultrasonography has the advantages of being noninvasive, inexpensive, and easily obtainable. Among its disadvantages are that it is highly observer dependent and therefore may not be reproducible, and that interpretation of results requires skill. Although abdominal ultrasound is an excellent test for such conditions as biliary obstruction and ureteral obstruction, the utility of the tests can be compromised by inappropriate test ordering and utilization.

Abdominal ultrasonography has known sensitivities and specificities based on the disorder under investigation. The positive and negative predictive values of the results also depend on patient selection, initial history, and/or complaint. Thus, if abdominal ultrasound is ordered inappropriately, the interpretation of the results as well as the positive and/or negative predictive value of the findings are problematic.

Third-party payers increasingly are defining criteria that address the issue of appropriateness of tests ordered by physicians. Their determination of appropriateness has been subjective, however, and has been made in the absence of objective criteria. Recently, the Health Care Financing Administration (HCFA) became involved in the issue of appropriateness but, rather than developing criteria for its determination, placed the burden on physicians. HCFA proposed that physicians prove the medical necessity of all laboratory tests billed to Medicare.<sup>1</sup> Although this policy may be difficult to implement, the rationale may be sound and may represent a new attempt to define appropriateness of tests ordered by physicians. It also makes implicit the need for more generic or broad-based criteria that can be applied across a variety of conditions and tests.

From the Section of General Internal Medicine, Department of Medicine, Tulane University Medical Center, New Orleans, LA

Address correspondence to: Dr Isaac W. Hammond, Section of General Internal Medicine, 1430 Tulane Avenue, SL-16, New Orleans, LA 70112.

**Table 1.** Indications for Abdominal Ultrasound

Indication	TMC	VA	Total
Abscess	0	1	1
Aneurism	3	3	6
Ascites	2	8	10
Biliary Obstruction	0	3	3
Cyst	4	1	5
Gallstones	12	5	17
Hydronephrosis	2	4	6
Liver Disease	5	2	7
Metastasis	0	2	2
Obstruction	1	8	9
Pancreatic Disease	1	1	2
Renal Disease	11	5	16
Renal Stones	2	3	5
Missing Data	38	33	71
<b>Total</b>	<b>81</b>	<b>79</b>	<b>160</b>

Recently, Vosburgh and Kopta<sup>2</sup> published a set of general criteria for determining the appropriateness of medical tests. They applied these criteria to the ordering of magnetic resonance imaging of the spine and extremities and found that, based on these criteria, 21% were not appropriate. The test involves the following three-step sequence of six questions. Step A: (1) Are the history and physical examination adequately documented? (2) Is the differential diagnosis supported? Step B: (3) Does the diagnostic study provide discriminatory or confirmatory information

regarding the differential diagnosis? (4) Is it required prior to initial therapeutic intervention? Step C: (5) Have the results of the diagnostic study been reviewed? (6) Do the results direct the therapeutic course? We used these criteria to evaluate the appropriateness of abdominal ultrasound in a private and a government hospital, both of which are Tulane University Medical Center teaching hospitals.

... METHODS ...

**Patient Selection**

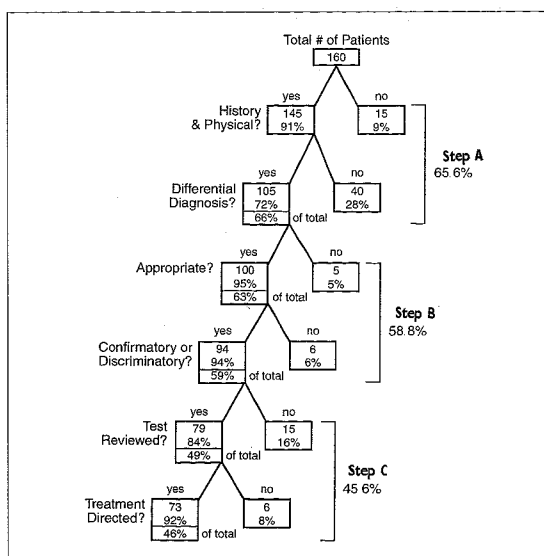
A list of all abdominal ultrasounds performed between January 3, 1994, and February 2, 1994, was obtained from Tulane University Hospital (TMC) and the Veterans Administration Medical Center of New Orleans (VA). All patients who underwent an abdominal ultrasound examination were eligible for participation in the study.

**Data Collection**

Each chart, including inpatient and outpatient medical records, was reviewed by three of the authors of this study and was evaluated using the six questions for the appropriateness criteria of Vosburgh and Kopta.<sup>2</sup> If all the criteria were met, the use of the diagnostic study was judged to be appropriate. The use of the study was judged to be inappropriate if any obvious deviation from the criteria occurred.

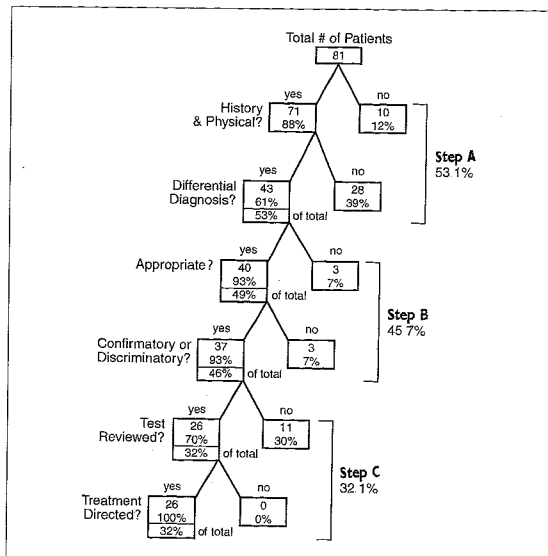
The three investigators who reviewed the charts were not masked to the physicians who ordered the test, but they were unaware of the study objective. The list of patients who had abdominal ultrasound was obtained from the respective radiology departments' listing of procedures. The study period was determined on the basis of the number of abdominal ultrasounds performed per month. The researchers were given a list of accepted indications for abdominal ultrasound in order to determine whether abdominal ultrasounds would provide confirmatory or discriminatory information to address the third question. One of the authors determined the list of accepted indications for abdominal ultrasounds by conducting a comprehensive review of the literature on the utility of abdominal ultrasound for a variety of conditions. The following conditions were deemed appropriate diagnostic considerations for the use of ultrasound of the abdomen: peritonitis, intra-abdominal abscesses, urologic malignancies, fever of undetermined origin, gastrointestinal malignancies, gynecologic malignancies, diseases of the aorta and large arteries, acute diarrheal disease, pancreatic disease, chronic hepatitis, cirrhosis of the liver, nephrolithiasis, pulmonary infections, and tubulointerstitial diseases.

**Figure 1.** Appropriateness Assessment for TMC and VA



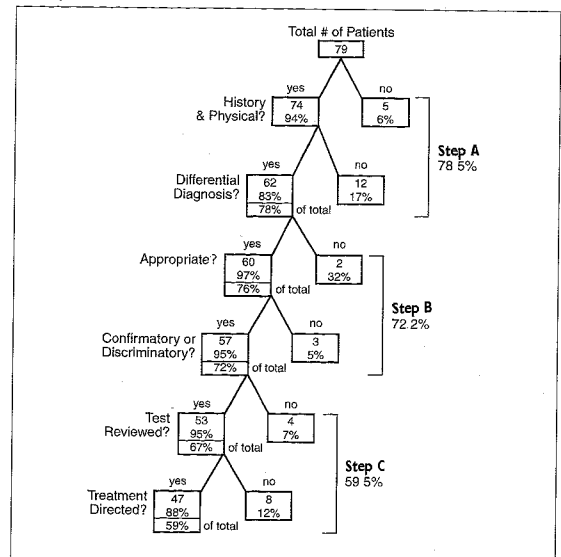
Appropriate use of abdominal ultrasound in study group. Note the percentage by each step on the right.

**Figure 1a.** Appropriateness Assessment for TMC



Appropriate use of abdominal ultrasound for Tulane University Hospital.

**Figure 1b.** Appropriateness Assessment for VA Hospital



Appropriate use of abdominal ultrasound for Veterans Administration Hospital of New Orleans.

In question 1, if there was any form of history and physical examination (complete or brief), the response was recorded as "yes." If the ordering physician stated at least one disease or condition in the visit note at the time that the ultrasound was ordered, then the response to question 2 was recorded as "yes." In Step B, if the disease or condition was one that was included in the list of appropriate indications for abdominal ultrasound, then the response to question 3 was recorded as "yes." If the report of the abdominal ultrasound provided a confirmation of the disease or condition in question, or provided information to be used to support another diagnosis considered appropriate for abdominal ultrasound, then the response to question 4 was recorded as "yes." Three subsequent visits were reviewed to determine whether there was a mention of the results of the abdominal ultrasound. If any reference to the abdominal ultrasound was made, then the response to question 5 was recorded as "yes." If the next step in the work-up of the patient was a logical follow up from the abdominal ultrasound based on the decision of the reviewer, then the response to question 6 was recorded as "yes." Twenty-five percent of all the charts from each reviewer were selected at random and were checked in a blinded fashion for quality control by concordance. If there was discordance between the two reviewers, then a conference was set up to resolve the issue.

**Data Analysis**

The data were abstracted into a computerized database for analysis. The chi square test was used to

compare two ratios; the significance level was set at  $\alpha = 0.05$ , two-tailed. Group comparison was performed as follows: TMC was compared with the VA, and then the overall appropriateness and each hospital's appropriateness percent were compared with 80%, which was selected as an acceptable "expected" for statistically significant difference.

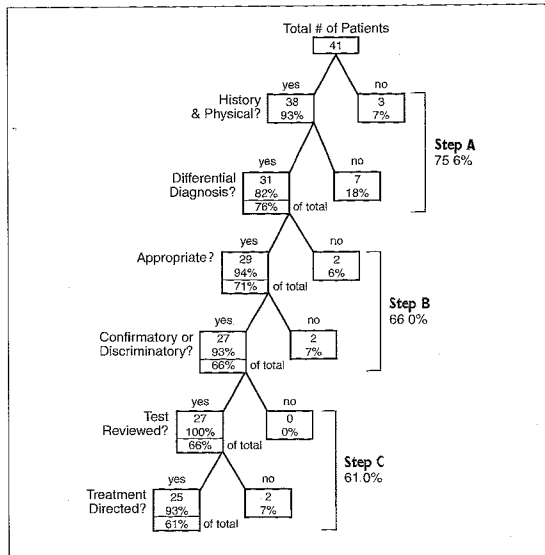
... **RESULTS** ...

Two hundred seventy-five abdominal ultrasounds were performed at the two hospitals during the study period (134 at TMC, and 141 at VA). We were able to obtain complete information on only 160 (58.2%) of the 275 patients; 160 patients therefore were included in the analyses. Of the 160, 81 (50.6%) were from TMC and 79 (49.4%) were from the VA.

The mean age for the study group was  $54.3 \pm 21.1$  years (range, 2 to 91 years). The mean age for the VA patients was  $62.6 \pm 14.3$  years (range, 23 to 87 years); the mean age for TMC patients was  $45.1 \pm 23.6$  years (range, 2 to 91 years). The study sample comprised 61% blacks and 75% men. The indications for ordering the abdominal ultrasounds based on appropriateness criteria are summarized in Table 1.

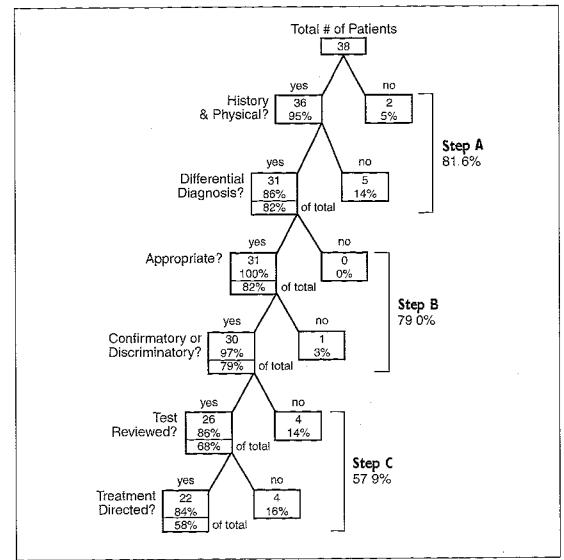
Of the 160 abdominal ultrasounds, 73 (45.6%) were found to have been ordered and used appropriately (Figure 1). At step A, we found that 105 (65.6%) of 160 physicians ordering tests documented a history and physical examination and at least one differential diagnosis. This number dropped to 58.8% when we

**Figure 2a.** Inpatient Abdominal Ultrasounds for VA Hospital



Appropriate use of abdominal ultrasound among inpatients at the Veterans Administration Hospital of New Orleans.

**Figure 2b.** Outpatient Abdominal Ultrasounds for VA Hospital



Appropriate use of abdominal ultrasound among outpatients at the Veterans Administration Hospital of New Orleans.

reviewed the differential diagnoses to evaluate whether abdominal ultrasound was appropriate, or whether it would provide discriminatory or confirmatory information to differentiate between the possibilities listed in the differential diagnoses (step B). At step C, the number decreased to 45.6%, because we found no evidence that the test results were reviewed by the physicians or that it directed the treatment of the patient.

In Figure 1a, we present the data from TMC, which had a 32.1% appropriate use of abdominal ultrasound. Figure 1b presents the data for the VA hospital, which had a 59.5% appropriate use of abdominal ultrasound.

The records from the VA hospital were more detailed than the information obtained from TMC; thus, we were able to perform subgroup comparison on the VA data. These subgroup analyses are presented in Figures 2a and 2b. In Figure 2a, we present data on inpatient use of abdominal ultrasound; we found 61%

appropriate use of this test. In Figure 2b, we present data on outpatient use of abdominal ultrasound; we found 57.9% appropriate use of the test. The difference between the two subgroups was not statistically significant.

The chi square test of independence was performed on the data to determine whether statistically significant deviations occurred from our expectation for the appropriate use of abdominal ultrasound in the two hospitals. We postulated a priori that at least 80% of the abdominal ultrasounds ordered in the two hospitals would meet the appropriateness criteria. Using this 80% level of appropriateness, the chi square test was performed. Results of the test are presented in Table 2. For the overall study group,  $\chi^2 = 51.05$ ,  $df=2$ ,  $P<0.0001$ . For TMC,  $\chi^2 = 23.23$ ,  $df=1$ ,  $P<0.0001$ , and for the VA,  $\chi^2 = 4.15$ ,  $df=1$ ,  $0.025<P<0.05$ .

... DISCUSSION ...

As physicians struggle to protect their autonomy in the changing atmosphere of medical practice in this country, appropriate use of diagnostic procedures or tests is essential to the future of medicine. Although steps are being taken to require physicians to justify laboratory tests,<sup>1</sup> it is incumbent on the physician to use all available tests and procedures appropriately. A test should be ordered only after an appropriate history and physical examination of the patient who presents with a complaint. If the history and physical examination do not provide obvious clues as to the cause, then we may proceed to consider a differential

**Table 2.** Results of Chi Square Test for Appropriateness

Hospital	Observed	Expected	Total
TMC	26	64.8	81
VA	47	63.2	79
<b>Combined</b>	<b>73</b>	<b>128.0</b>	<b>160</b>

Based on 80% expected  
 Overall:  $\chi^2 = 51.05$ ,  $df = 2$ ,  $P<0.0001$   
 TMC:  $\chi^2 = 23.23$ ,  $df = 1$ ,  $P<0.0001$   
 VA:  $\chi^2 = 4.15$ ,  $df = 1$ ,  $0.025<P<0.05$

diagnosis and the results of our history and physical examination. Diagnostic tests and/or procedures are then considered, but the test or procedure must be able to provide discriminatory or confirmatory information to help to narrow the list of differential diagnoses. After the test has been performed, the physician must review the results, document its findings in the patient's chart, and act on the information to either eliminate some possibilities from the list of differential diagnoses or confirm the diagnosis.

In a tertiary medical center in which future physicians are trained, one would expect appropriate use of a simple diagnostic test, such as abdominal ultrasound. In this study, we arbitrarily decided on an 80% appropriate use of abdominal ultrasound. Overall, we found that only 45.6% of all the abdominal ultrasounds met the appropriateness criteria. Although we lack data for comparison of the findings, our chi square test demonstrates a statistically significant difference from our expectation. Most physicians would agree that the appropriateness criteria incorporate most of the tenets of good medical practice. However, agreement did not translate into actual practice performance, as demonstrated by the results of our study. The poor performance found in this study, although surprising, should direct our attention to the waste that is embedded in our clinical practices. We must improve on these numbers if we are to successfully retain our autonomy in the changing climate of medical practice. Above and beyond any threat to physician autonomy, it simply makes good sense that we should be able to meet the intuitive steps stipulated by these criteria.

When we examined the data for the two hospitals separately, we found that TMC had a surprising 32.1% appropriate use of this test, compared with 59.5% for the VA hospital. Because similar studies have not been carried out in any other medical center, we are unable to compare our results with other data. Vosburgh and Kopta<sup>2</sup> reported a 78% appropriate use of magnetic resonance imaging of the spine and extremities in Oklahoma City.

The lack of documentation of a history and physical examination, for example, may not preclude the existence of a compelling reason to order the abdominal ultrasound. As a result of the retrospective nature of the design, we cannot address this issue. However, others may argue that without such documentation, there is really no documentation of the need and therefore justification for the test. In this era of medicolegal-conscious medical practice, lack of documentation would likely be interpreted as having no reason for the test. Another issue is that some physicians may in fact have reviewed the results of the test but, in the hustle and bustle of a real-world hospital practice, failed to

document this fact. We have no estimation of this occurrence, but again, documentation of results is expected in many cases for reimbursement of many tests.

Another issue that can be raised in interpreting these results is that the reason for ordering the test was invalid. Nevertheless, if the physician fails to review the test, then the purpose of the test cannot be justified for reimbursement purposes.

Finally, even though abdominal ultrasound has no known adverse effects to the patient, the cost of this test is significant to the cost of medical practice. At the two hospitals studied, abdominal ultrasound costs \$194 to perform, and a charge of \$140 is added for skilled interpretation of the findings. Thus, we can estimate the cost of the inappropriate use of abdominal ultrasound in the hospitals studied. If the total cost of an abdominal ultrasound is \$334 per ultrasound and 275 abdominal ultrasounds per month are performed in the two hospitals, then the total cost of all abdominal ultrasounds is \$1.102 million per year. If only 45.6% of the tests are ordered appropriately, then 54.4% are ordered inappropriately; therefore, 54.4% of the cost will be the cost of the inappropriate use of abdominal ultrasound. That cost is \$599,597. We examined the contribution from each hospital as well. An average of 150 abdominal ultrasounds are performed per month in each of the two hospitals, at a cost of \$601,200 per year per hospital. At TMC, \$408,215 (67.9%) of the money can be considered a waste. The inappropriate use of abdominal ultrasound costs the VA hospital \$243,486 per year.

In conclusion, the criteria cited seem to perform well when applied to the evaluation of this simple, noninvasive test. The robustness of the test will be determined if and when it is applied to other medical tests, and when we can show the same utility when compared with other tests of appropriateness. Adopting these criteria as requisites for test ordering would likely also improve documentation and this issue also deserves further study.

#### Acknowledgment

The authors express appreciation to Ms. Sylvia Delsa of the radiology department at the Veterans Administration Hospital of New Orleans, and the staff radiology department at the Tulane University Hospital.

#### ... REFERENCES ...

1. Gianelli DM. HCFA proposal: Doctors must show lab tests are necessary. *AMA News* January 2, 1995.
2. Vosburgh CL, Kopta JA. Appropriate use of magnetic resonance imaging of the spine and extremities. *South Med J* 1994;87:801-804.