

# Feedback: Computed Tomography for Subarachnoid Hemorrhage

## Feedback: Computed Tomography for Subarachnoid Hemorrhage

Robert L. Prosser, Jr., MD  
Shawnee Mission Medical Center  
Shawnee Mission, KS

Jonathan A. Edlow, MD  
Harvard School of Medicine and the Department of Emergency Medicine  
Beth Israel Deaconess Medical Center  
Boston, MA

Peter C. Wyer, MD  
EBEM Series Editor  
Columbia University College of Physicians and Surgeons  
New York, NY

Reprints not available from the authors.

**Address for correspondence:** Jonathan A. Edlow, MD, Department of Emergency Medicine, Beth Israel Deaconess Medical Center, Finard 202, 330 Brookline Avenue, Boston, MA 02215; E-mail [jonathan\\_edlow@hms.harvard.edu](mailto:jonathan_edlow@hms.harvard.edu).

Copyright © 2001 by the American College of Emergency Physicians.

0196-0644/2001/\$35.00 + 0

47/1/115847

doi:10.1067/mem.2001.115847

## Which Review Should We Believe Regarding the Diagnostic Power of Computed Tomography for Ruling Out Subarachnoid Hemorrhage?

[Prosser RL Jr. Which review should we believe regarding the diagnostic power of computed tomography for ruling out subarachnoid hemorrhage? *Ann Emerg Med.* June 2001;37:679-680.]

### To the Editor:

A superficial reading of the recent review of false-negative computed tomographic (CT) scans for the diagnosis of subarachnoid hemorrhage could leave the impression that the false-negative rate is 6% to 7%.<sup>1</sup> In the 2 studies selected by the authors, the negative predictive value for third-generation CT scans was 97% (36/37 and 56/58). These data confirm the information that 1 of the authors reviewed earlier this year in another journal.<sup>2</sup> That article cited a retrospective study that reported that third-generation CT scans were close to 100% sensitive (80/80) in identifying subarachnoid hemorrhage in the first 12 hours.<sup>3</sup>

The 6% to 7% referenced by Edlow and Wyer<sup>1</sup> in the “clinical bottom line” may refer to the upper CI of the likelihood ratio in Table 3. Are the authors trying to show how likely it is to have a negative scan in a patient with proven subarachnoid hemorrhage, or do they want to show the probability of finding a patient with a positive lumbar puncture in the whole cohort of patients with sudden, severe headache and a negative scan? The negative predictive value of the CT scan is the statistical computation that reflects the latter answer.

Because of the devastating consequences of missing the diagnosis of subarachnoid hemorrhage, many clinicians would agree that a lumbar puncture is often indicated when the CT scan is negative. However, it is important to understand that the data reviewed by the authors

in this article<sup>1</sup> and the earlier one published in the *New England Journal of Medicine*<sup>2</sup> set the chance of having a false-negative scan at 0% to 3% and not 6% to 7%, as the summary of the *Annals of Emergency Medicine* article seemed to suggest.

Robert L. Prosser, Jr., MD

1. Edlow JA, Wyer PC. How good is a negative cranial computed tomographic scan result in excluding subarachnoid hemorrhage? *Ann Emerg Med.* 2000;36:507-516.
2. Edlow JA, Kaplan LR. Primary care: avoiding pitfalls in the diagnosis of subarachnoid hemorrhage. *N Engl J Med.* 2000;342:29-36.
3. Sidman R, Connolly E, Lenke T. Subarachnoid hemorrhage diagnosis: lumbar puncture is still needed when the computed tomography scan is normal. *Acad Emerg Med.* 1996;3:827-831.

### In response:

[Edlow JA, Wyer PC. Feedback: computed tomography for subarachnoid hemorrhage—don't throw the baby out with the bath water [response]. *Ann Emerg Med.* June 2001;37:680-685.]

We wish to respond to 2 challenging and useful commentaries regarding our review, “How good is a negative cranial computed tomographic scan in excluding subarachnoid hemorrhage?”<sup>1</sup> Dr. Prosser<sup>2</sup> suggests that we underestimated the utility of the negative scan in our “clinical bottom line.” Dr. Hoffman,<sup>3</sup> although agreeing with our conclusion, argues that we cannot provide the patient with any valid estimate about the risk of having a subarachnoid hemorrhage (SAH) when the computed tomographic (CT) scan is negative. These commentators raise issues that pertain to the methods and to the philosophy underlying evidence-based reviews. We will deal with both of these areas of concern in the course of our response.

### ISSUES OF EVIDENCE

Dr. Prosser<sup>2</sup> raises the question of which of several potentially relevant studies should be chosen as the best source of information suitable for counseling and for advising patients regarding important decisions pertaining to their presenting complaints. This issue has been discussed previously in this series.<sup>4,5</sup> Dr. Prosser's comments also pose the question of which of the commonly reported measures of performance of diagnostic tests best serve the needs of patients and clinicians.

We will begin by attempting to provide some aids to sorting out the alphabet soup of measures of diagnostic test performance and will then deal with our choice of studies. It is best to keep in mind some essential features

of the diagnostic process. When diagnostic tests are used judiciously, clinicians and patients want to know how likely the condition of interest is to be present once a particular test result has been obtained (in this case, a negative CT scan for SAH within 12 hours of onset of symptoms). Such information will help them to decide whether additional testing will be useful. Prerequisites to this assessment include knowledge of how the diagnostic test performs in this setting and also an estimate of how high the suspicion of disease was before the test.

The effect of diagnostic test results on clinical decision making may be termed “probability revision.” A pretest probability (ie, suspicion) of disease is transformed, by virtue of the test result, into an estimate of the posttest probability.<sup>6</sup> When the performance of a diagnostic test in a given setting is known, the posttest probability may be derived from the pretest estimate by using a likelihood ratio, which can easily be calculated from the reported sensitivity and specificity of the test.<sup>6</sup> Determination of the posttest estimate can be accomplished through the use of a well-known nomogram or through an exact calculation. For reasons of accuracy, we used the latter approach in our review. For further clarification of the relationships between the measures of test performance relevant to our original clinical question, we refer the reader to the Figure.

When a test has only 2 possible results, and the result is positive, the posttest probability of disease is identical to the positive predictive value of the test in that population. It is the percentage of patients with positive results who truly have the disease or condition in question. Unfortunately, the progenitors of commonly used nomenclature in this area have not been kind. The negative predictive value, although related to the posttest probability of disease when a test result is negative, is not the measure we are seeking but rather its complement. The negative predictive value is the likelihood that a patient with a negative test result does not have the condition rather than the likelihood that she does.

In our review of CT for suspected SAH, when the results of the most clinically appropriate study were taken as the best evidence for the query we explored, we found that the observed sensitivity of CT during the first 24 hours, and presumably during the first 12 hours, was 93%, yielding a false-negative rate of 7%.<sup>7</sup> In other words, 7 of 100 patients with SAH would have negative CT scans. Because of the clinically necessary and, in our opinion, appropriate decision to accept positive CT scans as true-positive results, our initial assumption was that no patients without true SAH would have positive scans and

that 100% would have negative scans. The resulting likelihood ratio for a negative CT result was therefore 7% divided by 100% or 0.07, for which we determined an upper limit of the 95% CI of 0.4. When we applied this upper limit of the negative likelihood ratio to our patient's estimated pretest probability of disease of 15%, the resulting worst-case posttest probability of SAH was 7%.

Did we choose the right study? In the earlier narrative review<sup>8</sup> cited by Dr. Prosser, it was appropriately observed that other studies, such as that by van der Wee et al,<sup>9</sup> have suggested an even better performance of CT in ruling out SAH. As discussed in our article, we considered the results of the study by Morgenstern et al<sup>7</sup> to be the most appropriate to apply to patient care. Sixty-eight percent of the patients in the study by van der Wee et al were ruled in for SAH. The CT may well have performed more accurately in

such a high-risk population. Similarly, we excluded the study by Sidman et al.<sup>10</sup> Although Sidman et al reported 100% sensitivity for CT scanning, the overall severity of SAH within the study population could not be determined and was likely very high. The choices and preferences we made throughout our review were guided by a consideration that we believed most clinicians would appreciate, namely the decision to err on the side of least risk for the patient.

Turning to Dr. Hoffman's commentary,<sup>3</sup> he argues that we pretend to have enough evidence to intelligently counsel the patient in our scenario when, for reasons that he elaborates in detail, we really do not. Dr. Hoffman's concerns reflect issues both of science and philosophy. Let us look at his concerns systematically, beginning with those referable to science (ie, knowledge) in this area.

Regarding the incidence or likelihood of SAH among patients similar to the one in our article, Dr. Hoffman correctly states that if our value for the prior probability of SAH is incorrect, then all subsequent calculations will be inaccurate. We disagree that there are "many published case series" yielding importantly different estimates that would provide equally appropriate bases for estimating the prevalence of SAH in unselected outpatients presenting with acute severe headache and a normal neurologic examination. The series by van der Wee et al<sup>9</sup> involved only admitted patients. As noted previously, 68% had SAH. Although this may reflect a difference in clinical practice in The Netherlands, we believed that the restriction of the study to admitted patients was likely to have resulted in a proportionally more severe patient population and that very few practitioners in North America would report 2 of every 3 patients they evaluated for possible SAH as actually having the disease. Other series observing importantly higher rule-in rates for SAH might be cited but are less applicable because of the same issue of appropriateness of patient spectrum.<sup>11</sup> On the other hand, in addition to the study of Morgenstern et al,<sup>7</sup> one other series observed remarkably similar findings regarding the incidence of SAH in this population (outpatients with severe acute onset headache and a normal neurologic examination), which was about 12%.<sup>12</sup>

Dr. Hoffman<sup>3</sup> challenges the criteria we used to select the articles we included in our analysis and refers to "numerous small case series of missed SAH after a false-negative CT scan." We disagree that there are numerous such series that we should have included. Although there are some series that reported patients with SAH missed by CT, they are not of early presenting cases. For example, in the series of Lledo et al,<sup>11</sup> the median delay from onset of

**Figure.** Numerical and conceptual definitions of measures of diagnostic test performance for a negative test result.

| Performance Measure       | Numerical Definition* | Meaning   |
|---------------------------|-----------------------|---|
| Pretest probability       | $a+c/a+b+c+d$         | Percentage of patients having the index disease   |
| Sensitivity               | $a/a+c$               | Percentage of patients with disease having positive test result   |
| False-negative rate       | $c/a+c$               | Percentage of patients with disease having negative test result   |
| Specificity               | $d/b+d$               | Percentage of patients without disease having negative test result  |
| Likelihood ratio†         | $c/a+c:d/b+d$         | Probability of negative test result in patients with disease compared with that in patients without disease |
| Negative predictive value | $d/c+d$               | Probability that a patient with a negative test result is free of index disease                             |
| Posttest probability†     | $c/c+d$               | Probability that a patient with a negative test result has the index disease                                |

\*Referring to letters in "box."  
 †For negative test result.

Schematic representation of patient categories for a diagnostic test having only 2 possible results.

|       | Disease+ | Disease- |
|-------|----------|----------|
| Test+ | a        | b        |
| Test- | c        | d        |

The letters in each box correspond to the number of patients in each category.

headache to the presentation to the ED was 61 hours (72 hours for the patients with SAH). More importantly, Lledo et al did not report the time delays for individual patients. Therefore, many, if not most, of the patients with SAH and a negative CT scan in this report presented late, which was probably a major factor leading to false-negative CT scans.

Dr. Hoffman<sup>3</sup> in fact appears to ignore the crucial importance of the time delay from onset of headache to CT scanning in influencing the performance of the scan, even though he acknowledges the potential importance of the time factor in connection with an entirely different example, the use of human chorionic gonadotropin (hCG) in confirming pregnancy. Dr. Hoffman seems to have equated the issue of the time dependence of the performance of a diagnostic test with an entirely independent issue, that of patient spectrum. We will deal with this latter crucial issue below.

Regarding time sensitivity of diagnostic test performance, Dr. Hoffman's choice of hCG levels in pregnancy is useful and appropriate and one that emergency physicians will find intuitively satisfying. Studies designed to measure the sensitivity and specificity of hCG in diagnosing pregnancy in women whose last normal menses occurred 6 months before the test should not be used to infer the performance of the same test in evaluating women who present with dizziness and a normal period less than 28 days before presentation. We used this same principle in restricting our review of CT for SAH to studies reporting data on the use of CT to evaluate patients for SAH during the first 12 hours of symptom onset. An important difference between these examples is that CT scanning of patients with suspected SAH, unlike hCG testing in pregnancy, is more sensitive the earlier the patient presents. We know this from numerous studies,<sup>10,13-15</sup> and it is consistent with the pathophysiology of what happens to blood in the subarachnoid space.

Turning now to the issue of patient spectrum, we agree with Dr. Hoffman that this is one of the important things for clinicians to understand correctly in evaluating studies of the performance of diagnostic tests that they use in their practices. We therefore will take this opportunity to attempt to explain it in some detail.

Although Dr. Hoffman's parallel example of hCG in diagnosing pregnancy is not entirely unrelated to the issue of spectrum of disease, it ultimately breaks down as a vehicle for illustrating the latter principle. We suspect that most emergency clinicians are familiar with the clinical significance of the epithet that an appropriately selected female patient cannot be a little bit pregnant. She either is

or is not pregnant. The issue of how fast you can accurately relay the joyful (or unhappy) news to such a patient is not the issue of spectrum of disease. It is a different issue. In the case of SAH, the size of the initial hemorrhage is quite independent of the issue of how soon the patient presents to the emergency physician. Spectrum is inherently an issue of severity and not one of time.

The issue of spectrum is defined in an upcoming revised compendium of the *Users' Guides to the Medical Literature* as follows:

A diagnostic test is useful only to the extent it distinguishes between conditions or disorders that might otherwise be confused. Almost any test can differentiate the healthy from the severely affected; this ability tells us nothing about the clinical utility of a test. The true, pragmatic value of a test is therefore established only in a study that closely resembles clinical practice.<sup>16</sup>

Some years ago, it was customary to design studies of tests used in clinical diagnosis by performing the tests on 1 study population consisting of patients with unequivocal and severe disease (cases) and on another population of healthy volunteers with no clinical symptoms (control subjects). A case-control study of CT for diagnosing SAH, following this design, might draw the cases from patients admitted to a neurosurgical ICU after clipping of documented bleeding aneurysms and the control subjects from a group of college student volunteers with no history of headache. In such a study, no subject would actually require the test in question for purposes of clinical decision making, and almost any test, even a nonspecific test (eg, peripheral WBC count), might appear to be effective in discriminating between the 2 populations. The case-control design is the single most important factor in studies of diagnostic test performance likely to importantly overestimate the predictive power of the tests under investigation.<sup>17</sup>

Moving beyond the example of the case-control study design, it has also been suggested that within a study population selected to be clinically appropriate targets for a diagnostic test, the tests may be more effective on higher risk than on lower risk patients. This was reported in a study in which urine dipstick was tested on patients estimated by clinicians as having a greater than 50% likelihood of urinary tract infection and on patients estimated as having a less than 50% likelihood.<sup>18</sup> An observed sensitivity of the dipstick of 92% in the higher risk group declined to 56% in the lower risk group when a colony count of  $10^5$  was used as the criterion standard. However, the study by Lachs et al<sup>18</sup> might actually be interpreted as illustrating a different point (ie, the main point regarding

choice of patient group for a diagnosis study). Fewer than 7% of the “less than 50%” patient group turned out to have a urinary tract infection, and almost 1 of 4 low-risk patients who would have been treated on the basis of the dipstick result were ultimately found to have negative results by using the culture criterion. The rule-in rate for the “greater than 50%” group was actually slightly less than 50%. For many clinicians, this high-risk group might well constitute the truly appropriate spectrum of subjects for a study of the performance of the dipstick for purposes of clinical decisionmaking.

Perhaps Dr. Hoffman’s most provocative scientific quarrel with our review is his suggestion that not only the other studies we considered but also that by Morgenstern et al<sup>7</sup> should have been rejected because of an inappropriate patient spectrum. We believe this assertion to entail a serious distortion of the concept of spectrum in evaluating studies of diagnostic test performance. To be sure, the clinical manifestations of a disease may reflect its severity and contribute to the assessment of the appropriateness of a diagnostic test. Dr. Hoffman incorrectly states that 40% of patients in the Morgenstern study had abnormal mental status. In Table 1 of that study, “lethargy” is noted as a patient-reported symptom. However, as stated in the text and also verified by us with the authors, all patients had normal mental status on physical examination. The only abnormality allowed on examination was neck stiffness. Therefore these patients were not very different from our hypothetical patient.

Dr. Hoffman suggests that the presence of stiff neck (meningismus) among fewer than 30% (not “almost half”) of the patients included in the study by Morgenstern et al<sup>7</sup> for whom full data was available importantly conflicts with the appropriate target spectrum of disease. He overlooks the fact that we do not in fact know how many of the 51 patients in the subgroup of interest to us (<24 hours from symptom onset) had this finding. Although a very modest and nonstatistically significant increase in the likelihood of SAH among patients with stiff neck was observed by Morgenstern et al, the supposition of an important difference in diagnostic spectrum of SAH solely on the basis of presence or absence of a single potential predictor, such as meningismus, is unknown and unsupported by the literature. The fact that the only 2 patients with SAH in the Morgenstern study who were missed by CT scanning had nuchal rigidity simply underlines the fact that we have no real evidence that the presence or absence of meningismus importantly alters the performance of CT in detecting blood in the subarachnoid space within the first 12 hours of its appearance. We therefore

believe, and suspect that our clinician readers will agree, that it was indeed clinically appropriate to order a CT scan on patients meeting the inclusion criteria of the study by Morgenstern et al, including those with and without stiff neck. In other words, the spectrum of the patients in this study was appropriate and does not invalidate its results from a reasonable clinician’s standpoint.

Dr. Hoffman is correct in pointing out that, as specifically addressed in our review, there is no agreed on specific cutoff point for determining when the results of a spinal tap should be considered positive. Morgenstern et al<sup>7</sup> stated that of the 20 patients in the study whose cerebrospinal fluid findings were abnormal but who were thought not to have SAH, none were found to have subsequent SAH over 2 years’ follow-up. Thus, the cutoff point for cerebrospinal fluid used by Morgenstern et al, although possibly arbitrary, seems to have worked clinically.

We agree with 2 further points made by Dr. Hoffman regarding the issue of 100% specificity of the CT scan and variability of interpretation. Once again, we explicitly addressed this in the next to last paragraph of our review, when we consciously tilted the calculations toward the most conservative figures precisely because of these issues.

---

#### ISSUES OF PHILOSOPHY

Having dealt with the matters of science raised by Dr. Hoffman, let us deal briefly with those pertaining to the philosophy of clinical practice. This could be and has been the basis of extended discussions and controversies concerning evidence-based medicine in the medical literature.<sup>16</sup> As a means of demonstrating that there is a philosophical disagreement underlying Dr. Hoffman’s criticisms and comments, it is interesting to see where each of us ended up.

In our conclusion, we wrote:

You explain to the patient that even after the CT scan, she could still have as much as a 7% [not 4% as Hoffman states] chance of having an SAH, and because of the quality of the medical literature that is available, it could possibly be even higher. Based on your analysis of the data and the potential for a bad outcome if an SAH is missed, you strongly recommend to the patient that she undergo a LP [lumbar puncture] but leave the final decision with her.<sup>1</sup>

Dr. Hoffman concluded that, “The only way we could provide any real assurance that she does not have SAH, a condition which could soon maim or kill her if we fail to find it, but which with identification and treatment has an excellent prognosis, is by doing an LP [lumbar puncture].”<sup>3</sup>

Dr. Hoffman has arrived at the same conclusion as us: this patient needs a spinal tap. This belies the content of his detailed arguments without negating their instructional utility. Dr. Hoffman's overriding difficulty with our review pertains to philosophical issues regarding the appropriateness of being as concrete and precise as possible in communicating to patients what is and is not known about the options we are offering them and about the consequences of making decisions on the basis of that knowledge. It is the process by which conclusions are reached that is at issue and not the conclusions themselves.

Although it may not be entirely legitimate to attribute a single and uniform philosophical outlook to evidence-based medicine, one philosophical tenet implied in much of the primary literature on evidence-based medicine, including the *JAMA* "Users' Guides" series, is the commitment to involve patients in decisionmaking about their own health care to the extent that it is possible under given clinical circumstances. Dr. Hoffman objects to our having the patient in our clinical scenario set a threshold of 1% as the maximum risk she would be willing to take for the sake of avoiding the unwanted lumbar puncture. Certainly patients (and physicians) may have difficulty with the idea of being precise about their action thresholds. The details of physician-to-patient communication, both what is said and how it is said, are vitally important. Clinicians need to be sensitive to differences within patients' comfort levels regarding precise numbers and even regarding being directly involved at all in the clinical decisionmaking process.

Dr. Hoffman's statement that "it is clear that CT scanning is not extremely sensitive for SAH overall," is ambiguous. What is "extremely sensitive?" What is the significance of "SAH overall?" The fact remains that CT is more sensitive in patients who present early. But is it sensitive enough? Certainly this is what both our patient and her doctor wished to know. Is it not a physician's scientific and ethical obligation to help guide the patient (in a fashion appropriate to the patient's educational background, interest, and capacity to engage in detailed discussion in the context of acute illness) through the process of making this determination? An evidence-based philosophical orientation, as articulated by the leading proponents of evidence-based clinical practice, answers this question with an emphatic yes.

Evidence alone is rarely, if ever, the sole basis for patient-centered decisionmaking. Clinical experience, context, and patient values and preferences codetermine every important medical decision.<sup>16</sup> An evidence-based approach to patient care is a matter of bringing all of the relevant parties

together to make such decisions. This crucially important point has been elaborated in depth elsewhere.<sup>5</sup>

Suppose that at the point that we explained to the patient that a lumbar puncture will be advised if the CT scan is negative, our patient decided to sign out against advice, without having either CT or lumbar puncture? Dr. Hoffman proposes that we tell her, "We simply don't know how good the CT is in this situation." Is that a true statement, and will we be satisfied making it? To say that we can provide no real assurance does not help the patient make an informed decision. We do not believe, either scientifically or philosophically, that it is correct or useful to tell a patient presenting within the first 12 hours of headache onset that we have no more information after CT than before CT. Rather, we are quite confident in our assertion that the risk of SAH in this patient is very likely greater than her own identified decision threshold (1%) but probably less than 10%.

Emergency physicians frequently operate with ambiguity. Some deficiencies in information are due to not having needed data, such as present or past history, available in real time. Other deficiencies are due to the state of medical knowledge itself. Although the state of our current knowledge with respect to diagnosis of SAH is imperfect, to approach the patient as if we cannot offer any guidance whatsoever is neither useful to the patient nor consistent with what facts we do know.

Finally, we wish to address 1 further important comment of Dr. Hoffman. Clearly, the extent of inquiry presented in our published review of CT in SAH is beyond the scope of a busy emergency physician in the course of a typical clinical shift. Throughout the evidence-based emergency medicine clinical question installments, an instructional artifice has been adopted modeled on that of the original *JAMA* "Users' Guides."<sup>19</sup> The artifice might be interpreted to suggest that a busy clinician in any specialty or practice setting could routinely perform reviews such as this one on the fly. The answer and solution to this seeming discrepancy was provided in the first installment of the Skills for Evidence-Based Emergency Care subseries.<sup>20</sup> It involves choosing important questions, researching them when the time is available and creating practice and practitioner-specific databases of retrievable clinically pertinent information, sometimes called "critically appraised topics." The *Annals'* Evidence-Based Emergency Medicine section has been framed throughout to help the reader develop the skills required to begin to accomplish such a goal.

Ultimately, the matter boils down to issues of philosophy and outlook toward the emerging world of clinical practice. All dimensions of the care of patients, from our

own clinical assessment skills to our estimates of the value of the options we offer to our patients, frequently suffer from fallibility and imprecision. We ought not throw out the baby with the bath water.

*Jonathan A. Edlow, MD*

*Peter C. Wyer, MD*

1. Edlow JA, Wyer PC. How good is a negative cranial computed tomographic scan result in excluding subarachnoid hemorrhage? *Ann Emerg Med.* 2000;36:507-516.
2. Prosser RL Jr. Which review should we believe regarding the diagnostic power of computed tomography for ruling out subarachnoid hemorrhage? *Ann Emerg Med.* 2001;37:679-680.
3. Hoffman JR. Computed tomography for subarachnoid hemorrhage: what should we make of the "evidence?" *Ann Emerg Med.* 2001;37:345-349.
4. Schriger DL. One is the loneliest number: be skeptical of evidence summaries based on limited literature reviews. *Ann Emerg Med.* 2000;36:517-519.
5. Rowe BH, Klassen T, Wyer PC. One is the only number that you'll ever need! *Ann Emerg Med.* 2000;36:520-523.
6. Hayden SR, Brown MD. Likelihood ratio: a powerful tool for incorporating the results of a diagnostic test into clinical decision making. *Ann Emerg Med.* 1999;33:575-580.
7. Morgenstern LB, Luna-Gonzales H, Huber JC, et al. Worst headache and subarachnoid hemorrhage: prospective computed tomography and spinal fluid analysis. *Ann Emerg Med.* 1998;32:297-304.
8. Edlow JA, Caplan LR. Pitfalls in the diagnosis of subarachnoid hemorrhage. *N Engl J Med.* 2000;342:29-36.
9. van der Wee N, Rinkel GJ, Hasan D, et al. Detection of subarachnoid haemorrhage on early CT: is lumbar puncture still needed after a negative scan? *J Neurol Neurosurg Psychiatry.* 1995;58:357-359.
10. Sidman R, Connolly E, Lemke T. Subarachnoid hemorrhage diagnosis: lumbar puncture is still needed when the computed tomography scan is normal. *Acad Emerg Med.* 1996;3:827-831.
11. Lledo A, Calandre L, Martinez-Menendez B, et al. Acute headache of recent onset and subarachnoid hemorrhage: a prospective study. *Headache.* 1994;34:172-174.
12. Linn FH, Wijdicks EF, van der Graaf Y, et al. Prospective study of sentinel headache in aneurysmal subarachnoid haemorrhage. *Lancet.* 1994;344:590-593.
13. Sames TA, Storrow AB, Finkelstein JA, et al. Sensitivity of new-generation computed tomography in subarachnoid hemorrhage. *Acad Emerg Med.* 1996;3:16-20.
14. Adams HP, Kassell NF, Torner JC, et al. CT and clinical correlations in recent subarachnoid hemorrhage: a preliminary study of the Cooperative Aneurysm Study. *Neurology.* 1983;33:981-988.
15. van Gijn J, van Dongen KJ. The time course of aneurysmal haemorrhage on computed tomograms. *Neuroradiology.* 1982;23:153-156.
16. Guyatt GH, Rennie D, eds. *User's Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice.* Chicago, IL: The Evidence-Based Medicine Working Group, American Medical Association; 2001.
17. Lijmer J, Mol BW, Heisterkamp S, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA.* 1999;282:1061-1066.
18. Lachs MS, Nachamkin I, Edelstein PH, et al. Spectrum bias in the evaluation of diagnostic tests: lessons from the rapid dipstick test for urinary tract infection. *Ann Intern Med.* 1992;117:135-140.
19. Guyatt GH, Rennie D. User's guides to the medical literature. *JAMA.* 1993;270:2096-2097.
20. Wyer PC. The critically appraised topic: closing the evidence-transfer gap. *Ann Emerg Med.* 1997;30:639-640.