AJNR

Pontine Reversible Edema: A Newly Recognized Imaging Variant of Hypertensive Encephalopathy?

Sean O. Casey and Charles L. Truwit

This information is current as of May 5, 2025.

Foreword

Two editorials were sought on the subject of diskography based on the paper by Schellhas et al in this issue of the *AJNR.* The suspicion was that there would be divergent opinions concerning the value of diskography in general and on its value at the C2–C3 level in particular. The reliance that many spine surgeons place on diskography should make us aware of the potential value of diskography in certain well-defined clinical situations. The journal encourages the submission of papers on this subject, particularly those that are evidence-based and deal with clinical outcomes.

Robert M. Quencer, MD
Editor-in-Chief

Diskography: Science and the Ad Hoc Hypothesis

In this issue of the *AJNR* (page 269), Schellhass et al prompt this reader to reflect on the nature of science, and the concept of the hypothesis. In many ways, one has to admire this article as a technical tour de force. Nevertheless, it builds upon a hypothesis that has never been accepted. Even proponents of diskography will admit that instead of being evaluated and proven under strict scientific guidelines, such as those that apply to new drugs, diskography was popularized and adopted before its validity was determined (1).

Any proposed explanation in science is regarded initially as a hypothesis that is more or less probable on the basis of the available facts and relevant evidence. As a hypothesis, the question of its truth or falsehood is open to scrutiny, and there should be a continual search for more and more evidence to decide the question. The hypothesis in this case is the authors' stated claim that ''cervical diskography is a clinically useful test.'' It is certainly likely that the disk, under certain conditions of derangement, is a source of a patient's pain. This is supported by the observation that certain patients have reproduction of their symptom complex after disk stimulation. The question that arises, however, is whether this information is enough to support the entire hypothesis. In particular, whether the information available supports diskography as a ''clinically useful test.''

Unfortunately, this question of clinical usefulness is mired in controversy with both pro and con forces disparaging the scientific methods of the opposing camp. To date, both the critics and proponents of diskography have attempted to advance their positions by methodologically flawed scientific studies or editorial position papers. These publications have biased entry criteria and cohorts, and lack reasonable control groups. The end results are therefore difficult, if not impossible, to generalize.

Why is a diagnostic test performed in the first place? The purpose of a diagnostic test is twofold; it must provide reliable information about a patient's condition and influence the physician's plan for managing the patient

Concerning the first purpose, a diagnostic test, the following is known: 1) the morphologic characteristics of a disk, as revealed with diskography, is essentially irrelevant (1); 2) the negative studies by Holt (2) are sufficiently flawed as to be of no scientific value; and 3) the study by Walsh (3) suggests that the false-positive rate for a positive painful response to stimulation of the disk is very low. Normal disks have been shown not to produce pain with stimulation. Abnormal disks will produce pain with stimulation; however, not all abnormal disks are painful when stimulated. To employ a diagnostic test successfully, one needs to define the clinical problem in such a way that we understand the characterization of the patient population, the natural history of the disease process, including prevalence, incidence, and behavior over time, as well as the various diagnostic therapeutic options available. This underscores the complex morass that diskography seeks to inhabit; ie, the concept of ''internal disk disruption'' as a symptom-producing complex. Understanding the disease or symptom-producing complex purported to be the stimulus for diskography is a daunting challenge. We do not understand the natural history of this putative disorder, whether it always results in pain, whether it is one entity or many, or whether it represents a surgical objective.

There is tremendous pressure on caregivers to provide an explanation for patient symptoms, because patients and physicians are more comfortable with clear-cut relationships between symptoms and disease, cause and effect. There is also a sense that the more clearly defined the pathophysiologic process of the disease is, the better use is made of the available treatment options. It is contended that diskography is an informational tool only, a test designed to obtain information about the source of a patient's pain. How that information is used or abused is not the responsibility of the proponents of the test. Accepting diskography as a valid diagnostic tool, however, may lull surgeons into considering surgical treatment for ''discogenic pain'' and potentially may lead to inappropriate surgery. This brings us to the second pur-

pose of a diagnostic test—to influence the physician's plan for managing the patient.

A diagnostic test that has no impact on treatment choice is unlikely to benefit the patient, except through the reassurance of the physician. Others would contend that knowledge about a disease, even in the absence of effective treatment options, has been described as an important part of the healing process (4). In the case of diskography, however, these assertions clearly need to have some type of utility assessment before they can be accepted.

Let us return to the most important question inherent in accepting the hypothesis that diskography is clinically useful. Does it affect the choice of treatment so that it has a positive impact on patient outcome? Proponents would argue that patient management may be improved by excluding invasive therapy such as surgery if one finds multiple painful disks or obtains indeterminate results. This argument presupposes that the surgical intervention that is held in the balance has a proven efficacy, which is not the case.

Given that 1) the symptom complex and source of disease are poorly understood; 2) the treatment techniques vary widely in their theoretical approach and efficacy; and 3) the diagnostic test that serves as a cornerstone in the decision-making process has not been subjected to well-controlled studies, one must conclude that the use of diskography is predicated on flawed logic and science, and vulnerable to abuse and misuse. I do not believe it is worthwhile to pursue diskography as an "informational tool" for the purposes of establishing a diagnosis for which there is no proven therapy. In fact, any use of diskography as an informational tool (outside of well-controlled and designed trials to establish basic efficacy) must be seriously questioned from an ethical standpoint.

Diskography is not going to go away, despite wishful thinking. Clearly, prospective investigations of diskography are warranted. Although it seems likely that disk stimulation can enable identification of a symptomatic disk, it remains to be shown that this information has prognostic value. This must be done. What might these clinical trials look like? The following is one example. A patient population could be stratified based on symptoms/ working diagnosis and identification of patients likely to benefit from diskography (probably those

with chronic pain). These patients could be randomly placed into groups that are tested with either MR imaging or MR imaging plus diskography, and followed thereafter. The surgical rates, costs, number of days worked, complications, patient anxiety, and sense of well-being could be tabulated, with the primary endpoint being functional status. These two groups could then be compared to determine if diskography had a positive predictive value or a positive influence on therapeutic thinking.

The beliefs that a false test is useful or that a useful test is false are equally reprehensible. There is a need for continued research into the pathophysiologic mechanisms of disk stimulation. The real challenge lies in identifying a patient who would benefit from this test. Until these decisions can be based on well-controlled clinical trials, there is no basis for the performance of diskography in clinical medicine.

The authors' final statement is perhaps a finer use of the English language than they intended: "Provocative cervical diskography, including [diskography of] C2–3 if possible, can be employed to evaluate head and neck pain of suspected cervical discogenic origin." The question is, however, not can it be employed, but rather should it be employed.

We do no justice to the test or our patients by refusing to follow good clinical science. If diskography is a valuable tool, let's prove it. If it is not, let's discard it quickly once and for all.

MICHAEL T. MODIC, M.D.

Chairman, Division of Radiology
Cleveland Clinic Foundation
Cleveland, OH

References

1. Bogduk N, Modic MT. **Controversy: lumbar discography.** *Spine* 1996;21:402–404
2. Simmons JW, Aprill CN, Dwyer AP, Brodsky AE. **A reassessment of Holt's data on: "the question of lumbar discography."** *Clin Orthop* 1988;237:120–124
3. Walsh TR, Weinstein JN, Spratt KF, et al. **Lumbar discography in normal subjects.** *J. Bone Joint Surg* 1990;72A:1081–1088
4. Guyer RD, Ohnmeiss DD. **Contemporary concepts in spine care—lumbar discography.** *Spine* 1995;20:2048–2059

Cervical Diskography: Analysis of Provoked Responses at C2–C3, C3–C4, and C4–C5

As radiologists, we depend heavily on imaging as our primary diagnostic tool. We have thus come to believe that what we see with imaging is objective truth: only what we see is real; what we do not see does not exist. Most radiologists have little direct patient contact, so there is a tendency to discount or ignore the patient's history and physical examination findings. These aspects of diagnosis, which are dependent upon patient responses, are considered to be too subjective.

Is there any facet of radiology that has led to more controversy than diskography, an examination depen-

dent more on a patient's responses than on imaging? Some radiologists feel there is no scientific basis for diskography, because it relies heavily on a patient's subjective responses. Others further complain that not enough reproducible studies have been performed to assure scientific credibility.

Radiologists may find it hard to believe that what we see is not always clinically significant, and what we cannot see may exist nonetheless. MR imaging is one of the most sensitive and specific tools radiologists have, yet studies have shown that many spine abnormalities we see may not be clinically significant (1, 2). Therefore, it would be folly to treat a patient solely on the basis of imaging findings, and without regard to clinical context. Furthermore, there have been studies showing that imaging may fail to reveal the cause of disease, whereas the cause may be revealed by a more subjective test, such as a physical examination or diskography (3).

Schellhas et al attempt to bridge the gap between what radiologists do or do not see with imaging, and what might be a clinically relevant cause for a patient's pain. It is an elegant study. Their science begins to provide us with an objective, valid rationale for the use of upper cervical diskography as an additional diagnostic tool, particularly when imaging fails to correlate with clinical findings. In practical terms, diskography can be a valuable diagnostic tool when MR findings are too numerous to interpret. It provides us with a tool to narrow the diagnosis to what is clinically relevant. Furthermore, in cases where the imaging findings fail to reveal the cause of pain, diskography may enable the radiologist to detect the cause.

Those of us who directly treat patients afflicted with chronic pain soon realize that pain is a complex, multi-faceted phenomenon. Sometimes overwhelming pain masks lesser pains. We see examples of painful vertebral compression fractures where, after successful vertebroplasty, the dominant pain disappears; but then the patient gradually

notices the appearance of a different pain (eg, degenerative facet disease, diskogenic pain, or spinal stenosis) that becomes more and more bothersome. We see this as an unmasking of lesser pains that were present, but were overshadowed by a much more intense pain. If we treat only the acute compression fracture that is obvious on image, but ignore all other patient concerns, we fail as physicians.

The study by Schellhas et al is a step toward scientifically validating the use of upper cervical diskography. We hope the authors will continue, as more needs to be done prospectively to verify this technique's clinical significance. If upper cervical diskography proves to be a clinically valid test that helps in the selection of patients for treatment, then it will become an extremely valuable tool for the radiologist, particularly in those cases when imaging does not clearly show the abnormality.

We congratulate the authors for a scientific work that elucidates a most difficult and subjective patient problem, pain, and for tackling an emotional issue, diskography.

WADE WONG, D.O.
CHARLES KERBER, M.D.
Interventional Neuroradiology
University of California
San Diego, CA

References

1. Healy J, Healy B, Wong W, Olson E. **Cervical and lumbar MRI in asymptomatic older male lifelong athletes: frequency of degenerative findings.** *J Comput Assist Tomogr* 1996;20:107–112
2. Jensen M, Brant Zawadski MN, et al. **Magnetic resonance imaging of the lumbar spine in people without back pain.** *N Engl J Med* 1994;331:69–73
3. Shellhas KP, Smith MD, Cooper R, et al. **Cervical discogenic pain: prospective correlation of magnetic resonance imaging and discography in asymptomatic subjects and pain sufferers.** *Spine* 1996;21:300–312

Pontine Reversible Edema: A Newly Recognized Imaging Variant of Hypertensive Encephalopathy?

Although the typical imaging findings of hypertensive encephalopathy have been well described in the medical literature, the diagnosis can occasionally present a challenge. This may be related to the infrequency of patients presenting with hypertensive encephalopathy, relatively vague clinical symptoms (headache, visual disturbances, and seizures), and failure to communicate the patient's elevated systemic blood pressure to the radiologist. Imaging findings in mild cases of hypertensive encephalopathy are those of edema, usually within the cortex and subcortical white matter of the parietal, occipital, temporal, and to a lesser degree, the posterior frontal lobes, typically with bilaterality, although not with

perfect symmetry. More severe cases tend to have flagrant involvement of the subcortical white matter and may extend to the frontal, posterior temporal, cingulate, and central sylvian regions, as well as to the cerebellar white matter. The most severe cases can have various degrees of thalamic, insular, and pontine involvement.

In most cases, the changes of hypertensive encephalopathy appear to represent reversible vasogenic edema, which is seen on T2-weighted images and can sometimes be shown by diffusion-weighted imaging (DWI) and apparent diffusion coefficient (ADC) maps. Most investigators believe that hypertensive encephalopathy begins with progressive hyperten-

sion, eventually leading to failure of autoregulation and cerebral hyperperfusion. This in turn results in breakdown of the blood-brain barrier with vasogenic edema, which begins in the cerebral cortex and later accumulates in the subcortical white matter.

The symptoms and imaging findings of hypertensive encephalopathy have been found to be remarkably similar, if not identical, to changes associated with a number of other acute illnesses, including eclampsia/pre-eclampsia, cyclosporin-A (CSA) neurotoxicity, and other more obscure illnesses (1). It is now believed there is a distinct clinico-neuroradiologic syndrome that encompasses these various conditions, reported recently in the *New England Journal of Medicine* by Hinchey et al (1), wherein the name ''reversible posterior leukoencephalopathy syndrome'' was suggested. With the increasing numbers of organ and marrow transplant patients on immunosuppressive agents, this syndrome may be seen quite frequently in some tertiary transplant centers. Animal models and our recent experience with fluid-attenuated inversion-recovery imaging suggest that the early changes in this syndrome occur within the cerebral cortex, rather than in the white matter. Thus, we prefer the term ''posterior reversible encephalopathy (or edema) syndrome (PRES),'' as a more appropriate descriptor of the syndrome. In recent years, this syndrome has been known by several names, including ''hypertensive encephalopathy,'' ''hyperperfusion encephalopathy,'' ''reversible leukoencephalopathy,'' ''occipito-parietal encephalopathy,'' and ''reversible posterior cerebral edema syndrome.'' Additional risk factors include renal failure, systemic lupus erythematosis, and numerous drugs including other immunosuppressants, such as FK506 and high-dose corticosteroids, as well as various chemotherapeutic agents (1). Often patients have a combination of risk factors and may have been placed on one or more of these medications, putting them at risk for PRES. Most patients have some hypertension, although many of the patients, especially children and those with CSA neurotoxicity, do not have systemic blood pressure levels as high as are typically encountered with ''pure'' hypertensive encephalopathy. It is likely that, outside the category of hypertensive encephalopathy, there are various superimposed pathophysiologic mechanisms leading to the common imaging appearance of PRES among its different entities. The physiologic changes of PRES are dynamic, and pathologic correlation is fortunately rarely obtained. In past reports, pathologic analysis has yielded only evidence of fibrinoid necrosis within the arteriole walls, interstitial edema, and petechial microhemorrhages, but little or no evidence of infarction (1).

At least two unusual variants of PRES may exist. The first is a syndrome of cerebral autoregulation disturbance that occurs as an uncommon complication in chronically hypoperfused portions of the brain following revascularization procedures. After carotid endarterectomy or stenting, this phenomenon is termed the ''postcarotid endarterectomy hyperperfusion syndrome'' and results in reversible edema uni-laterally in the affected cerebral hemisphere on the same side as the endarterectomy.

It appears that de Seze et al, in this issue of the *AJNR* (page 391), report a second imaging variation of PRES. They describe two patients with hypertensive encephalopathy manifested as reversible increased signal on T2-weighted images essentially isolated to the pons. The MR findings described by de Seze et al in these two cases are indeed unusual, but may not be as rare as one might think. The differential diagnosis for such pontine T2 hyperintensity includes pontine glioma, ischemic and radiation changes (generally irreversible conditions), as well as central pontine myelinolysis (CPM), often a devastating condition except in mild cases. The absence of abnormal serum sodium levels, the clinical recovery of their patients, and the resolution of the MR abnormalities make the diagnosis of CPM less likely. It is interesting to note that there is an association of CSA neurotoxicity, one of the better known etiologies of PRES, with CPM in patients after liver transplantation (2). Some of these patients have undergone severe serum sodium flux and truly have CPM. Most of these patients, however, are reported to have good clinical recovery. In reality, there is no pathologic proof that all of these cases represent reversible CPM, as opposed to a central variant of PRES, or a combination of the two conditions. Perhaps some neuroradiologists may be able to recall a case with similar pontine imaging findings to those shown by de Seze et al, but in which a diagnosis of CPM was unlikely from the absence of the typical clinical course and the absence of serum sodium alterations. Some of these cases may, in retrospect, have been secondary to this newly described imaging variation of hypertensive encephalopathy or PRES, perhaps associated with less than obvious elevations of systemic blood pressure. For instance, we recently encountered the following case in point.

A 44-year-old man presented with a decreased level of consciousness and a history of recent cocaine use and was admitted to our emergency department. CT revealed a punctate hemorrhage in the right thalamus, but the diagnosis of hypertensive hemorrhage was complicated by the presence of bilateral thalamic and pontine hypoattenuation. MR imaging was performed, verifying a small intraparenchymal hemorrhage within the right thalamus, but also showing extensive edema within the thalami bilaterally, the pons, and the midbrain. Upon inquiry, it was learned that the patient's systemic blood pressure was extremely elevated. Nevertheless, none of the typical posterior cerebral cortical and subcortical T2 hyperintensities of PRES were evident on the MR images. DWI and ADC maps in this patient showed increased ADC in the thalami, midbrain, and pons, therefore excluding an acute ischemic phenomenon and indicating, instead, a process of reversible vasogenic edema. After institution of antihypertensive therapy, the patient had clinical resolution of his symptoms. The imaging changes, although not completely resolved, were greatly improved on follow-up MR imaging 3 days later. This case would thus appear to be a similar

example of brain stem involvement of hypertensive encephalopathy in the absence of peripheral cerebral lesions.

In patients with such pontine T2 hyperintensities, DWI is useful for two reasons. First, it is helpful in refining the differential diagnosis because hyperintensity on DWIs might favor other processes, such as acute ischemia or acute CPM (we have seen two such autopsy-proven cases of CPM), although it may still be possible in severe cases of PRES. Although in the majority of cases the lesions of PRES can be assumed to be reversible with treatment, we have seen a small number of severe cases with ischemic complications heralded by restriction of fluid movement on ADC maps. The use of DWI in these cases also provides prognostic information regarding the likelihood of reversibility.

As we learn more about PRES, we recognize the wider spectrum of imaging appearances of this condition. This spectrum has already been reported in the uremic encephalopathies, a group of conditions including hemolytic-uremic syndrome, hepatorenal syndrome, and thrombotic thrombocytopenic purpura. Some reports have demonstrated bilateral lesions in the parieto-occipital regions identical in appearance to those seen in PRES, whereas other reports have demonstrated predominant involvement of relatively central structures, such as the basal ganglia or the brain stem. It now appears that the uremic encephalopathies represent additional etiologies of PRES, which, for reasons unknown, have a greater tendency for central distribution.

Future research on PRES should be directed at perfusion imaging. To date, there have been a number of conflicting reports using perfusion imaging in these entities, with some reporting hyperperfusion and others reporting hypoperfusion. It is likely that these results depend on the time of imaging relative to the onset of therapy in patients with PRES. Present data tends to favor the theory that the condition begins with hyperperfusion, resulting in failure of autoregulation, and breakthrough accumulation of vasogenic edema. We believe that overly aggressive antihypertensive therapy, in the setting of disturbed cerebral blood flow autoregulation, can result in hypoperfusion, even with apparently normal blood pressures. In some severe cases, this can lead to infarction predominantly in the posterior border zones. Ischemia may also result from status epilepticus and hypoxic complications. Over the years, this mix of transient hyperperfusion and the infrequent ischemic complication has created much confusion in our attempt to elucidate the pathophysiologic mechanisms contributing to the various etiologies of PRES. Given the dynamic nature of brain perfusion in PRES, it may be useful to perform perfusion imaging in selected patients who are responding poorly to therapy as a guide for a more "personalized" titration of antihypertensive therapy.

In conclusion, it is important for neuroradiologists to become aware of the spectrum of imaging findings in the acute presentation of PRES. The diagnosis of PRES is one of the more satisfying diagnoses made in our practice, as it is often unsuspected by clinicians, and relatively dramatic changes on MR imaging can be predicted to be predominantly, if not completely, reversible. The neuroradiologist will often be the first physician to suggest the appropriate diagnosis in the hope of averting unnecessary biopsies and initiating appropriate therapy. Clinicians also need to become more familiar with this syndrome and the treatment issues in order to minimize underlying risk factors and to avoid potential ischemic complications.

SEAN O. CASEY, M.D.
CHARLES L. TRUWIT, M.D.
University of Minnesota
Minneapolis, MN

References

1. Hinchey J, Chaves C, Appignani B, et al. **A reversible posterior leukocephalopathy syndrome.** *N Engl J Med* 1996;334:494–500
2. Fryer JP, Fortier MV, Metrakos P, et al. **Central pontine myelinolysis and cyclosporin neurotoxicity following liver transplantation.** *Transplantation* 1996;61:658–61

Radiographic Screening for Orbital Foreign Bodies Prior to MR Imaging: Is It Worth It?

The article by Seidenwurm et al in this issue of the *AJNR* (page 426) addresses questions that are faced daily by any radiologist who performs MR imaging. These are: Does a patient have intraorbital metal that would be a contraindication for having an MR examination? Which patients should be screened radiographically? When is radiographic screening cost-effective? Because these are common dilemmas, the impact of any recommendations based on this study is potentially important. It is therefore imperative that a cost-effectiveness analysis (CEA) be rigorous and complete.

Guidelines have indeed been developed for conducting and evaluating CEAs (1). We reviewed Seidenwurm et al's article based on 10 points that all readers should consider when evaluating such studies (Table, page 246).

Cost-effectiveness can be determined by comparing the resources consumed by a given strategy (the costs) with the improvement in health that results from that strategy (the consequences). The consequences are measured in units most relevant to the strategy under study. This results in ratios such as "dollars per year of life gained."

Criteria for evaluating a cost-effectiveness article

10 Criteria According to Drummond (1)	How Well Seidenwurm et al Fulfilled Criteria
1. Was a well-defined question posed in an answerable form?	Yes
2. Were all the important and relevant costs and consequences for each alternative identified?	Probably
3. Were costs and consequences measured accurately in appropriate physical units?	Partially
4. Was a comprehensive description of the competing alternatives given?	Partially
5. Was there evidence that the program's effectiveness had been established?	Yes
6. Were costs and consequences valued credibly?	Cannot tell
7. Were costs and consequences adjusted for differential timing?	Partially
8. Was an incremental analysis of costs and consequences of alternatives performed?	Yes
9. Was a sensitivity analysis performed?	Yes
10. Did presentation and discussion of study results include all issues of concern to users?	Yes

Most researchers think that, in general, quality of life should be incorporated into these analyses. Seidenwurm et al chose to use utilities, which are the most widely accepted measures of quality of life. Utilities refer to preferences for a particular level of health status. These preferences may be those of either an individual or society for a particular health outcome, and they result in a "quality weighting" factor. The denominator of a cost-utility analysis is therefore quality adjusted life years (QALY), rather than simply life years.

This is the type of analysis that Seidenwurm et al have performed. The first and most important step in the design of any study is the formation of a focused, answerable question. This question must describe the alternatives being compared as well as the viewpoint of the analysis. Seidenwurm et al do an admirable job by clearly stating that the purpose of their study was to compare the cost-effectiveness of clinical versus radiologic screening for orbital foreign bodies. They describe the clinical screening in adequate detail, but they could have provided more information for the radiologic screening, such as the number of views obtained. To their credit, the authors unambiguously declare that the analysis is from the societal viewpoint. This viewpoint takes into account the widest possible range of costs and consequences and is most appropriate for policy decision making.

Seidenwurm et al are somewhat unconventional in their identification of costs and consequences and in how they organize their economic model. Economists generally categorize costs as direct and indirect. The costs of organizing and operating a service are the direct costs, and they include health professionals' time, supplies, equipment, power, capital costs, and out-of-pocket expenses for the patient. Time lost from work is an indirect cost. In their classic paper on cost-effectiveness analysis, Weinstein and Stason (2) present the following equation for determining the net healthcare costs of an intervention:

$$\Delta C = \Delta C_{Rx} + \Delta C_{SE} - \Delta C_{Morb} + \Delta C_{Rx\ \Delta LE}$$

where ΔC_{Rx} includes all direct medical and healthcare costs, ΔC_{SE} are the costs associated with adverse effects of the intervention, ΔC_{Morb} are the savings due to prevention or alleviation of disease, and $\Delta C_{Rx\ \Delta LE}$ are the costs of treating diseases that would not have occurred if the patient had not lived longer because of the intervention. Because length of life is probably not affected significantly by the intervention (orbital screening), $\Delta C_{Rx\ \Delta LE}$ can be ignored. Similarly, there are probably no adverse effects of orbital screening, so this term can be ignored as well. ΔC_{Morb} needs to be estimated because this is the cost benefit of screening. The authors account for this with their variables A and M, both of which they assume to be $0 in their base case. Although one can question this base-case assumption, their sensitivity analysis demonstrated that these were not influential variables. The authors ignore direct out-of-pocket costs, and while it is difficult to be certain that these are insignificant, assuming they are negligible is conservative.

Seidenwurm et al chose to measure consequences in terms of QALY, which is the most appropriate measure for a cost-utility analysis. Drummond identifies two other categories of consequences: 1) changes in functional status (physical, social, and emotional functioning); and 2) changes in future resource use. Neither of these categories is addressed by Seidenwurm et al's article, but this is true of many economic analyses.

The authors state that the cost of screening was "culled from the medical literature on screening for orbital foreign body, Medicare fee schedules for various examinations, and usual, customary and reasonable charges fee schedules for various examinations." Using Medicare fee schedules is probably appropriate in this setting, because they reflect a resource-based relative-value scale. Nevertheless, the authors remain vague as to how exactly they arrived at their base-case estimate of $173, an amount they indicate represents the charge of the examination rather than a true cost. Numerous authors have emphasized why it is important to distinguish between costs and charges, with a recent example being an editorial by Picus in *Radiology* (3). Seidenwurm et al state in their discussion that the Medicare allowable fee for a single view screening examination is $25. How do they account for the difference between this amount and their base case? They state that $25 does not cover the costs of radiography. This may be true, but needs justification. After all, their analysis demonstrated that the cost of the radiographic screening was a critical variable, and if the cost was as low as $25, then screening might be cost-effective.

Using QALY as a metric for consequences implies accounting for preferences, either on the individual or societal level, for given health states. The authors estimate the degree of disability from monocular blindness using two separate sources. Both of these, however, probably use functional status and not preference-based measures, and thus are not true utilities. Nonetheless, their base-case estimate of the utility for monocular blindness being 0.24 is probably quite conservative. We recently collected a cohort of 142 patients who completed a time trade-off for monocular blindness, and the mean utility was 0.82.

It is impossible to tell from the authors' methods exactly how various types of disability were converted to QALYs. The authors include in their cost-effectiveness equation the variable "D," which is the degree of disability associated with injury. They use disability rating guides to assess disability due to ocular injury, but do not provide essential details. QALYs describe a preference for a given health state, and not just the functional status within that health state.

The alternative to radiologic screening, clinical screening, is reasonably well described in their methods. Enough details are supplied so that a different provider could carry out the clinical screening. The radiologic screening is less thoroughly described, with no details provided as to whether one or more views were obtained, or if costs assumed digital or film-screen systems.

One of the most compelling aspects of the article is the last paragraph of the discussion section in which they describe their experience using the proposed screening protocol. Although limited to a single practice, this experience is a true measure of effectiveness (how a protocol performs in real life).

The authors appropriately use a range of discount rates for costs in their sensitivity analysis. They do not discount consequences. This is a somewhat controversial area, but for the most part, people agree that it should be done.

With respect to costs, the authors account for both the costs of radiologic screening, for which they use charges as a proxy, and the costs of clinical screening, which they argue are negligible. They also look at the incremental improvement in the detection of ocular foreign bodies and thus the incremental improvement in QALY of radiologic versus clinical screening.

Sensitivity analysis is a method to determine the degree of uncertainty associated with economic analyses. It is in many ways the equivalent of defining confidence intervals. A sensitivity analysis is performed by varying the value of a particular variable across a range of clinically relevant values. If large changes in the value of this variable do not substantially affect the cost-utility ratio, then the confidence in the original results is high. If certain variables do greatly affect the ratio, then greater precision is needed in defining the value of these variables. A one-way sensitivity analysis varies one variable at a time. Two-way and greater sensitivity analyses can be done, although the difficulty of interpreting the analysis increases as the number of variables increases.

The authors performed multiple one-way sensitivity analyses, and determined that cost of screening, expected life span, and prevalence of foreign bodies were all critical variables. This means that their model is not robust along a realistic range of values for these variables. The authors discount the importance of the cost of screening, asserting that the point at which screening becomes effective ($25) is so low as to be unrealistic. As I've stated, they need to justify that costs are significantly greater than $25. Similarly, if patients can be preselected to increase the prevalence of foreign bodies to 2.5%, then screening becomes cost-effective.

In their discussion, the authors touch on aspects of the analysis that required them to make critical assumptions, such as the average length of life, or the utility associated with blindness. One aspect of the decision-making process the authors do not address, but which may be the most critical variable, is the question of liability and the legal costs associated with ocular injury. This is an indirect cost, and therefore is not accounted for in their analysis. The fear of litigation, however, may be the driving force in current screening protocols.

As Drummond (1) states, the "... intent in offering a checklist is not to create hypercritical users who will be satisfied only by superlative studies... [but rather to] help users of economic evaluations to identify quickly the strengths and weaknesses of studies." Although Seidenwurm et al fall somewhat short of the rigorous and complete standard set by Drummond, they have made a commendable effort, and their conclusions are probably correct.

JEFFREY G. JARVIK, M.D., M.P.H.
SCOTT RAMSEY, M.D., PH.D.
University of Washington
Seattle, WA

References

1. Drummond MF. *Methods for the Economic Evaluation of Health Care Programmes* 2nd ed. Oxford; New York: Oxford University Press; 1997 (Oxford medical publications)
2. Weinstein MC, Stason WB. **Foundations of cost-effectiveness analysis for health and medical practices.** *N Engl J Med* 1977; 296:716–721
3. Picus D. **Comparing competing medical procedures: costs or charges–what should it matter? [comment].** *Radiology* 1996; 199:623–624; discussion 624–625.