Clinical Efficacy of SPECT Bone Imaging for Low Back Pain

Benjamin Littenberg, Alan Siegel, Anna N.A. Tosteson and Thomas Mead

Departments of Medicine, Radiology and Community and Family Medicine and Dana Biomedical Library, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire

SPECT has been advocated as an accurate and useful diagnostic tool for patients with low back pain. We sought to answer three questions:

1. What is the accuracy of SPECT in diagnosing the cause of low back pain?
2. What is the clinical effect (influence on management or patient outcomes) of SPECT in low back pain?
3. What is the cost-effectiveness of SPECT in low back pain?

Methods: We conducted a comprehensive structured review of the literature, analyzing 940 citations from 1966 through September 1993 and completed a narrative review. We also attempted quantitative synthesis of the accuracy of SPECT evaluation of low back pain. **Results:** We found thirteen reports on accuracy. Only three provided a reasonable gold standard reference test and allowed the calculation of sensitivity and specificity. There is weak evidence that SPECT is useful in: (a) detecting pseudarthroses after failed spinal fusion, (b) evaluating young patients with back pain and (c) distinguishing benign from malignant lesions in cancer patients. SPECT has not been sufficiently studied in any other setting. We found no reports on the clinical outcome of SPECT or its cost-effectiveness. **Conclusion:** The decision to use SPECT in most patients with low back pain cannot be supported by clinical trials. Its effect on clinical management and cost-effectiveness are unknown. The medical community should mount a large-scale, prospective evaluation of SPECT in low back pain.

Key Words: single-photon emission computed tomography; low back pain; radionuclide bone imaging; meta-analysis

J Nucl Med 1995; 36:1707–1713

SPECT has been promoted as an accurate and useful diagnostic tool for patients with low back pain. In evaluating this exciting new imaging technology, we searched the medical literature to document SPECT's diagnostic accuracy, clinical usefulness (effect on patient management) and cost-effectiveness. We applied the tools of structured literature review and meta-analysis to minimize bias in our assessment. We prepared a formal research protocol before beginning work. The protocol posed research questions, detailed the inclusion and exclusion criteria for the study, the procedures for obtaining data and the analytic methods in an effort to minimize bias and error.

Received June 27, 1994; revision accepted Sept. 20, 1994.

For correspondence contact: Benjamin Littenberg, MD, Program in General Medical Sciences, Campus Box 8005, Washington University School of Medicine, 660 S. Euclid, St. Louis, MO 63110

Reprints are not available.

MATERIALS AND METHODS

This work sought to address three issues by systematically reviewing the literature on bone SPECT in low back pain. First, we attempted to summarize any valid clinical trials that estimated the accuracy of SPECT in low back pain. Second, we searched for published literature that demonstrated the clinical effect of SPECT in low back pain. Third, we looked for published literature that estimates the societal benefits, particularly cost-effectiveness data, of SPECT in low back pain.

Literature Search

Eligible reports included published work that estimated the diagnostic accuracy of SPECT in humans with low back pain. Articles were excluded if they did not study SPECT, did not study back pain in humans, were not in English, did not study at least 10 subjects, did not provide specific counts of true-positive, true-negative, false-positive and false-negative results, duplicated previously published data, did not address low back pain or did not report an adequate reference test (surgical results or long-term follow-up). We also searched for articles reporting the cost-effectiveness or clinical effect of SPECT in patients with low back pain, even if they did not provide specific accuracy data.

We searched MEDLINE on the MEDLARS system at the National Library of Medicine in Bethesda for articles published from 1966 through September 1993. We also used two non-MEDLINE computer resources (Biological Abstracts and Excerpta Medica) to search for relevant studies in a few, selected, non-*Index Medicus* journals.

Each citation was reviewed by an investigator. If the citation (including the title, key words and abstract when available) was clearly irrelevant, it was coded as "not eligible." Citations that were possibly relevant or not clearly excludable on the basis of the computer record were marked "to obtain." References marked "to obtain" were found in the library or departmental collections or were acquired by interlibrary loan.

Each obtained report was reviewed in full by an investigator. If it was still eligible after full review, it was coded "eligible" and subjected to analysis. The reference lists of all eligible reports

were added to the database to broaden our capture of potentially eligible reports.

Analysis

We subjected each eligible article to a narrative review that sought to highlight its strengths and weaknesses. We abstracted the number of subjects, the patient sources (to assess referral bias), inclusion and exclusion criteria (to assess generalizability), the clinical problem under study, the authors' conclusions, the presence of a reference test (including surgical findings or long-term clinical follow-up), independence (blinding), the presence of any primary data indicating economic or fiscal outcomes or any data indicating follow-up of patients and whether the data appeared to be published in duplicate. We did not distinguish among various technical variations in performing SPECT. For instance, studies using older equipment were not analyzed separately from newer reports. For each report that provided counts of true-positive cases, false-positive cases, etc., we calculated the true-positive rate (TPR or sensitivity) as the number of true-positive cases divided by the number with disease (true-positives plus false-negatives). We calculated the false-positive rate (FPR or 1-specificity) as the number of false-positive cases divided by the number of cases without diseases (false-positives plus true-negatives). We also determined the likelihood ratio as a convenient single measure of the test's discriminating power. It was calculated as the ratio of TPR-to-FPR. A likelihood ratio of 20.0 means that the odds of disease are 20 times higher after the test than was thought before the test was performed. A likelihood ratio of 1.0 occurs when a test has no ability to discriminate patients with disease from healthy subjects.

In addition, we planned quantitative meta-analyses using the methods of Littenberg and Moses (*1,2*) and a formal review of cost-effectiveness and clinical affect. We would be willing to pool even two or three eligible articles on the same clinical problem using similar imaging techniques and epidemiologic methods. We found no such groupings, however, in the SPECT literature.

RESULTS

Literature Found

The literature search returned 940 citations. After review, three full reports met all inclusion criteria and provided interpretable accuracy data, three abstracts, and seven full reports provided some partial information. No reports contained any information on clinical affect or cost-effectiveness. The characteristics of the 13 reports are described in Table 1 and summarized below.

Eligible Reports (n = 13)

Full Reports with Complete Data (n = 3). Bodner et al. (*3*) reported on 15 young patients with low back pain. They did not indicate the source for the patients. A reference diagnosis was provided from medical records or telephone conversation. Two were diagnosed with spondylolysis, one with spondylolisthesis, one with lumbar Scheuerman disease and eight with fractures. SPECT detected 11 of these 12 (TPR = 0.917) and was normal in all three cases of "mechanical back pain" (FPR = 0.0). Because there were no false-positives, the likelihood ratio for a positive SPECT diagnosis cannot be calculated. The authors did not provide any data on the clinical importance of the 12 positive cases or on the clinical outcome of SPECT imaging.

Even-Sapir and colleagues (*4*) provided the most recent data. Although their main goal was to analyze individual lesions, they did provide some patient-level data on 233 subjects who had undergone SPECT of the lumbar spine. Among 75 patients with a known malignancy but no known spinal metastases, 74 were evaluated by some combination of biopsy, CT, MRI, follow-up planar bone scan, follow-up plain films or clinical examination over 8-17 mo. One patient was never evaluated. Twenty-nine had metastases and 45 were thought to have benign tumors (prevalence = 39.2%). SPECT was positive (meaning not completely normal) in nearly all cases, including the 28 metastases (TPR = 0.966) and 43 benign cases (FPR = 0.956). Because so many SPECT images were positive by this very broad criterion, the test did not appear to have discriminating power. A positive test increased the probability of metastases from 39.2% to 39.4%. The likelihood ratio of a positive test was 1.01.

Even-Sapir and colleagues provided some interesting evidence on another important point. Although nearly all the studies were "positive," the pattern of positivity may be helpful in discriminating benign from malignant lesions. In particular, they found that lesions in the pedicle were more commonly cancerous while those confined to the body of the vertebra tended to be benign. Unfortunately, although they demonstrated this finding on a lesion-by-lesion basis, they did not provide adequate data to determine if it is useful in discrimating patients with metastases from those without metastatic spread.

The same report (*4*) provided data on 158 patients with back pain but no known primary malignancy. Seven cases of sacroiliitis were confirmed by MR or CT. Sixty-eight of 73 normal SPECT images were confirmed by a variety of reference tests; five cases are unknown. Unfortunately, the final diagnoses of 78 patients with positive SPECT images were not recorded. The accuracy of the test in the clinical setting of low back pain without malignancy cannot be calculated from the data provided. The authors did not comment on the clinical or economic effects of SPECT.

Slizofski and colleagues (*5*) reported a series of 26 patients after spinal fusion. Fifteen patients had back symptoms at the time of study and underwent SPECT to detect possible pseudarthroses. Of the 15, 11 were confirmed by repeat surgery and 4 by clinical follow-up. SPECT correctly identified seven of nine pseudarthroses (TPR = 0.778) and five of six nonpseudarthroses (FPR = 0.167). The ratio of TPR-to-FPR resulted in a likelihood ratio of 4.7. The remaining 11 patients were not fully described. The authors concluded that SPECT and planar bone scanning together are "a highly sensitive screening test to detect painful lumbar pseudarthrosis."

This study may have bias from lack of diagnostic independence. The operating surgeons were probably not blind to the scintigraphic images when making their diagnoses and the authors do not explain why repeat surgery failed in three of nine patients. If these three patients did not really have a pseudarthrosis, then SPECT detected four of six

TABLE 1
Literature Summary

First author and year of publication	Population and clinical problem	Reference tests	Provides full data on		
			Accuracy	Clinical effect	Cost-effectiveness
Bellah RD 1991 (*6*)	162 young adults with LBP. R/O pars fracture.	None. Compared to PBS and CT in some.	No	No	No
Bodner RJ 1988 (*3*)	15 young patients with LBP. Find any abnormality.	Clinical f/u × 7 mo	Yes	No	No
Buscombe JR 1990 (*13*)	28 adults with chronic low back pain. Clinical problem or goal of testing not specified.	None. Compared to CT, PBS and planar x-ray	No	No	No
Collier BD 1985 (*7*)	19 adults with current or prior LBP. Find painful spondylolysis or spondylolisthesis.	Presence of pain. No independent diagnosis.	No	No	No
Even-Sapir E 1993 (*4*)	74 adults with LBP and cancer not known to involve the spine. Rule out metastases.	Combined biopsy, autopsy, imaging and clinical f/u × 6 mo.	Yes	No	No
Gates GF 1988 (*8*)	100 adults with suspicion of a spine or pelvis lesion Find any lesion.	None. Compared to PBS.	No	No	No
Nagele-Wohrle B 1989 (*9*)	70 unspecified patients. Find any lesion.	None. Compared to PBS.	No	No	No
Onsel C 1992 (*10*)	753 adults with LBP. Find sacroiliac uptake.	Clinical f/u of some patients.	No	No	No
Ryan PJ 1992 (*11*)	34 adults with LBP and normal blood tests and unchanged radiographs. Find any lesion.	CT	No	No	No
Ryan RJ 1992 (*12*)	80 consecutive LBP patients. Find any lesion.	None. Compared to PBS.	No	No	No
Ryan PJ et al. 1990 (*14*)	33 adults with LBP. Clinical problem or goal of testing not specified.	None. Compared to PBS.	No	No	No
Ryan PJ et al. 1992 (*15*)	10 patients with LBP and increased facet activity by SPECT. Direct injection therapy.	Response to facet injection.	No	No	No
Slizofski WJ 1987 (*5*)	15 adults with LBP after spinal fusion. Find pseudarthrosis.	Repeat surgery in 11. Not specified in 4.	Yes	No	No

LBP = low back pain; PBS = planar bone scan; f/u = follow-up.

cases (TPR = 0.667) and missed four of nine noncases (FPR = 0.444) for a likelihood ratio of 1.5. If these errors occurred, then the corrected prevalence of pseudarthrosis is 40% (rather than 60% as published) and is more consistent with other published series.

Full Reports with Partial Data (n = 7). These reports provide a partial assessment of the performance of SPECT in low back pain. Because they did not provide an external reference test or gold standard, we cannot determine if the SPECT results are accurate. Accordingly, we cannot cal-

culate TPR, FPR or a likelihood ratio from most of these reports.

Bellah et al. (*6*) presented a case series of 162 young athletes with low back pain referred for scintigraphy in Boston by orthopedic surgeons. Ninety-one had both normal planar bone scans and normal SPECT. Seventy-one (44%) had abnormal SPECT images of the spine. Planar bone scan was positive in only 32 of the 71 patients. In other words, only 45% of SPECT abnormalities were detected by planar bone scan. The authors provided no external validation of the scintigraphic diagnoses and no information on the clinical outcome of SPECT imaging.

Collier et al. (*7*) studied a series of 19 adults with current or prior low back pain. Thirteen had current pain and six were asymptomatic. All had spondylolysis and/or spondylolisthesis and all had breaks of the pars interarticularis on plain film radiographs or CT. All underwent both SPECT and planar bone scanning. The reference test was the presence of pain at the time of scintigraphy. The scintigrams were read without knowledge of the clinical situation. SPECT was positive in 11 of 13 symptomatic subjects (TPR = 0.846) and 1 of 6 pain-free subjects (FPR = 0.167). Planar bone scanning had a lower sensitivity (TPR = 0.632) than SPECT. SPECT identified 68% of pars defects compared to 42% for planar bone scanning. This report demonstrated that SPECT tends to be positive in patients with painful pars fractures and negative in asymptomatic fractures. In other words, it appears that SPECT may be useful in distinguishing which radiographically detected fractures are associated with pain. The authors suggest that this information may be useful in directing clinical management, but do not provide clinical outcome data.

Gates (*8*) studied 100 patients with both SPECT and planar bone scintigraphy. SPECT was positive in 21 patients with normal planar bone scans, but no confirming diagnostic data were presented. This report demonstrated that SPECT is more sensitive that planar bone scanning but does not address the issue of false-positives by providing an independent reference test.

Nagele-Wohrle et al. (*9*) used a similar design in their study of 70 patients. Thirty-nine cases had "identical" results with both planar bone scintigraphy and SPECT. In 31 cases, SPECT provided information that planar bone scans did not. Apparently, no SPECT images were normal. This study did not confirm the accuracy of SPECT because it did not provide a reference test diagnosis for any of the patients.

The report from Onsel and colleagues (*10*) presented data on 43 patients with sacroiliac uptake. Although no details of the diagnostic methods were provided, the authors declared that a definitive diagnosis was made in all but seven patients. No patients developed cancer. These data do not allow calculation of sensitivity, but the FPR can be estimated (assuming that all the "definitive diagnoses" were clinically important) as 7/43 = 0.163. In other words, at least 16% of patients with sacroiliac uptake on SPECT have no significant diagnosis.

Ryan et al. (*11*) studied 34 patients with low back pain who were referred from a rheumatology clinic for SPECT imaging. They were selected because they all had normal blood tests, stable radiographs and stable CT scans of the lumbar spine. The authors provided comparative data with plain radiography and CT but did not provide any nonimaging data to confirm the final diagnoses. CT was used as the reference test and SPECT had a TPR of 24/28 = 0.857 and a FPR of 3/6 = 0.500. Of course, one might argue that the SPECT images are more likely to be correct than the CT studies. This issue cannot be resolved, however, without a true reference test.

Ryan et al. also studied 80 consecutive patients with low back pain (*12*). Sixty percent had positive SPECT studies compared to 35% for planar bone scintigraphy. No reference test data were provided. (*N.B.*: Although this article was written by "RJ Ryan," we believe it is the same author as "PJ Ryan.")

Abstracts (n = 3). Buscombe et al. (*13*) studied 28 patients with chronic low back pain, excluding those with nerve root compression or malignancy. All patients received planar radiography, CT, planar bone scintigraphy and SPECT. SPECT and CT were more commonly positive than planar bone scans or planar radiography. The authors concluded that "CT and SPECT are superior" but do not provide any information on the true state of the patient. In other words, there was no reference test and we do not know how many of the positive tests were false-positives.

The remaining two abstracts are from the Guy's Hospital group. In the first (*14*), the authors reported the relative rates of positive planar bone scans and SPECT in 33 adults with low back pain. SPECT was more commonly positive (45% versus 24% for planar bone scans), but there were no reference test data and therefore, no way to tell if the index tests were accurate.

The last abstract (*15*) presented a pilot study of facet joint injection to relieve chronic low back pain in 10 adults with increased facet activity. Two patients were "cured," four had some improvement and four had no change in their symptoms. If we take relief of symptoms to be a reference test, we can calculate that SPECT has a positive predictive value (PPV) of either 20% or 60% (depending on how one classifies the partial responders). Eliminating the four partial responders gives a PPV of 2/6 = 33%. No patients with back pain and normal facet activity by SPECT were studied. Therefore, we cannot calculate sensitivity or specificity.

Accuracy

The published literature provide very few data on the accuracy of SPECT in patients with low back pain. Only three full reports met even minimal standards of methodologic rigor. One small series (*5*) addressed patients with failed back surgery and suggested that SPECT is highly accurate. It had little protection against diagnostic incor-

TABLE 2
Accuracy of SPECT as Reported in Three Studies

Author	TP	FP	FN	TN	TPR	FPR	Prev	PPV	NPV	LR
Bodner et al. (*3*)	11	0	1	3	0.917	0.000	0.800	1.000	0.750	UN
Even-Sapir et al. (*4*)	28	43	1	2	0.966	0.956	0.392	0.394	0.667	1.01
Slizofski et al. (*5*)										
As published	7	1	2	5	0.778	0.167	0.600	0.875	0.714	4.66
Corrected	4	4	2	5	0.667	0.444	0.400	0.500	0.714	1.50

TP = true-positive cases; FP = false-positive cases; FN = false-negative cases; TN = true-negative cases; TPR = true-positive rate; FPR = false-positive rate; Prev = prevalence; PPV = positive predictive value; NPV = negative predictive value; LR = likelihood ratio; UN = undefined.

poration bias and may have seriously inflated estimates of TPR and FPR.

One report (*4*) indicated that SPECT was of little value in detecting spinal metastases in patients with known primary cancer. In this study, SPECT failed to discriminate mainly because it was almost always read as positive. Only 4% of evaluable patients had a negative SPECT test. The data suggest, however, that a different interpretation policy (based on the pattern of uptake) might have resulted in better performance. In other words, the authors provide evidence that SPECT can distinguish benign from malignant lesions, but not that it is useful in diagnosing low back pain in patients.

The third report (*3*) was also quite small. In this retrospective series, SPECT correctly classified 14 of 15 young patients with low back pain. The reference test is not well described and it is unclear if the images were read independently of (without knowledge of) the final diagnosis.

Although the abstract reporting on facet injection (*15*) provides a reference test of sorts, the population is small and selected on the basis of a positive index test (SPECT). Therefore, it cannot contribute to our estimates of accuracy.

Because these reports studied such different patient populations, we did not combine the individual estimates of accuracy. Furthermore, it is not possible to estimate accuracy in other important populations such as adults with no known cancer or patients with suspected osteomyelitis.

Patient Outcomes and Cost-effectiveness

We were unable to find any published data on the cost-effectiveness of SPECT imaging of low back pain or on the role of SPECT in clinical management, patient outcomes or resource use.

DISCUSSION

Quality of the Published Evidence

A physician, a patient or a policy-maker seeking to understand the accuracy, clinical effect or cost-effectiveness of SPECT imaging for low back pain will be disappointed by the published literature. We found only three small studies, with only 104 subjects in three different clinical settings, that met the most minimal requirements for quality control. Unfortunately, these reports suffer from many potential biases and other inadequacies as outlined in the Results. Although the technical aspects of delivering SPECT technology and the advantages in sensitivity over planar bone scintigraphy have been fairly well studied, data to evaluate the clinical accuracy, effect or cost-effectiveness of SPECT for low back pain are scarce.

We suspect that SPECT is a useful tool in at least some groups of patients with low back pain. This suspicion is partly due to the groundbreaking work reviewed here. The authors of these pioneering studies should be congratulated for leading the effort to investigate this important technology. Unfortunately, because the supporting evidence is so scant, this potentially valuable technology may not be offered to some patients who would benefit from it. Likewise, SPECT may be used for some patients who do not benefit from it because the test is inaccurate or otherwise useless. No amount of biologic theory or technical accomplishment can substitute for solid evidence of a test's value (*16*). Physicians cannot possibly make good judgements about using SPECT given the current state of knowledge.

Characteristics of an Ideal Study

What should we have found? An ideal study would reduce the effect of random error ("noise") by enrolling a large cohort. The precision of an accuracy estimate is driven, in large part, by the size of the smallest cell in the 2 x 2 table. One cannot make a stable estimate of the specificity if the sample does not include at least a few false-positive subjects. As a rough rule of thumb, 5 or 10 subjects in the smallest cell are usually adequate (*17*). Table 2 shows how none of the studies we reviewed met this standard. Depending on prevalence, TPR and FPR, this could require hundreds or thousands of subjects.

Other sources of variation that should be minimized include technical factors in test performance and interpretation standards of the index and reference tests. Good studies will prospectively provide training and quality control at every step of the process.

An ideal accuracy study should avoid selection bias or spectrum-of-disease bias by systematically enrolling all subjects with a particular diagnostic problem (*18*). In this case, we would require a series of patients with back pain for whom it is reasonable to suppose that SPECT could be helpful. They should be enrolled at their original site of

care in primary care or orthopedics rather than after referral to radiology or nuclear medicine services.

Incorporation bias occurs when the index test is interpreted with some knowledge of the results of the reference test or vice versa. To avoid these problems, the two tests should be performed and interpreted independently. Furthermore, all subjects should receive both tests, regardless of the results of the first. In other words, a normal index test should not prevent the use of the reference test.

Finally, the reference test should be a clinically meaningful gold standard. This is difficult to do in imaging in general and is particularly difficult in SPECT for low back pain. For pseudarthrosis, Slizofski et al. (*5*) chose their reference test quite well. Surgical exploration seems like a reasonable reference test. It could be improved upon by explicitly documenting the criteria the surgeons used to diagnose pseudarthrosis and by ensuring that they had no access to the SPECT images until after surgery. If the nuclear medicine physician and the surgeon and the patient believe that SPECT will be helpful, they may be reluctant to prescribe something as invasive, costly and dangerous as surgery while blinded to the SPECT images. Given that we could find reports of only 11 patients who have ever had their SPECT diagnosis of pseudarthrosis confirmed surgically, and that SPECT was far from perfect in these cases, these decision makers may be too quick to endorse this technology!

Surgery is not a reasonable reference test in all clinical situations. Often, another test must be sought. Imaging procedures such as planar bone scintigraphy, radiography, CT, MRI and myelography are not convincing reference tests. They all clearly have substantial error rates. Sometimes the only reference test worth performing is long-term clinical follow-up. Although their study is too small to provide stable estimates of specificity and sensitivity, Bodner et al. (*3*) demonstrated this method to some advantage. It is important to be studying a clinical situation in which "time will tell" and effective treatment is unlikely to obscure the final diagnosis: the detection of metastases comes to mind. Even-Sapir et al. (*4*) did just this and also incorporated autopsy results when available. Again, it is valuable to provide a protocol for using clinical information to make a reference test diagnosis and for ensuring that the index test does not contaminante this process.

These study design suggestions have been discussed in the literature (*18–20*) and will strengthen an investigation that seeks to measure accuracy. Accuracy is, however, an intermediate outcome (*16*). We would prefer to know that patients who undergo SPECT have superior long-term outcomes. Do they live longer than those who forgo the test? Do they have better physical function? Are the costs lower? Better satisfaction? To answer these questions (and the associated question of cost-effectiveness), we advocate a prospective, randomized, controlled trial in which both groups receive the best possible care, but one group has a SPECT test to facilitate diagnosis and the other does not.

Limitations of This Analysis

The major potential weakness of this literature review or meta-analysis is missing data: Three areas of possible missing data are particularly troublesome. First, foreign literature may have escaped our attention if it were not available in English. We have no evidence that such literature available, but we cannot be sure. Second, literature that is not cataloged and indexed in MEDLINE is often difficult to find. We searched several other indices and the entire run of several non-MEDLINE nuclear medicine journals. We also communicated with several senior nuclear medicine physicians to protect against this possibility. Third, unpublished data do not appear in this review. In view of the long lag time between study completion and publication, it is possible that more eligible studies will soon appear. It is also conceivable that someone has done one or more eligible studies but has failed to publish in peer-reviewed journals. We undertook a careful and thorough search of the literature, assisted by several experienced nuclear medicine specialists. Nonetheless, if we failed to find a large, well-designed and carefully executed analysis of SPECT in low back pain, it could influence our views. If such a trial were found, however, it would be unlikely to address more than one or two of the many potential clinical indications for SPECT. For example, a superb trial of SPECT to diagnose degenerative disease in patients with radicular pain would offer little evidence on the effectiveness of SPECT in possible osteomyelitis.

This analysis is limited to the application of SPECT technology for the evaluation of low back pain. We did not evaluate its use in any other clinical setting. Nor do we have data to offer on other imaging modalities.

Some might argue that this systematic review of the literature should not bear the label "meta-analysis." After all, there was no pooling or statistical summary. We believe that the salient points of a meta-analysis were all present: an explicit question, a written protocol with specific inclusion and exclusion criteria and analytic methods, protection against bias and systematic error and a comprehensive search of the literature. Only the final step of a meta-analysis, pooling or summarizing, was omitted for lack of data.

CONCLUSION

There is some flawed (but interesting) evidence that SPECT is useful in two specific clinical circumstances: detecting pseudarthroses after failed spinal fusion and evaluating young patients with back pain. SPECT does not appear to be as clinically helpful, as evidenced by Even-Sapir et al. (*4*), in diagnosing patients who have back pain. It may, however, be able to discriminate benign from malignant lesions. It has not been well studied in any other setting. Therefore, the decision to use SPECT in most patients with low back pain cannot now be supported by clinical trials. In all settings, its effect on clinical management and cost-effectiveness are unknown.

Given that SPECT consumes resources (including money, facility space, physician and technologist time and patient time and travel) and has potential adverse effects (exposure to radiopharmaceuticals, claustrophobia and the deleterious cascade after a false-positive test), it is imperative that the medical community mount a well-designed, prospective evaluation of SPECT in low back pain.

ACKNOWLEDGMENTS

The authors are grateful to the members and staff of the Society of Nuclear Medicine and the American College of Nuclear Physicians who generously provided information for this review and especially to Drs. Henry D. Royal, William H. McCartney, B. David Collier and John W. Keyes for their assistance.

This work was supported by the SPECT Project Foundation of the Society of Nuclear Medicine and the American College of Nuclear Physicians. Dr. Littenberg is an American College of Physicians George Morris Piersol Teaching and Research Scholar.

REFERENCES

1. Littenberg B, Moses LE. Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method. *Medical Decision Making* 1993;13:313–321.
2. Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: Data-analytic approaches and some additional considerations. *Statistics in Medicine* 1993;12:1293–1316.
3. Bodner RJ, Heyman S, Drummond DS, Gregg JR. The use of single photon emission computed tomography (SPECT) in the diagnosis of low-back pain in young patients. *Spine* 1988;13:1155–1160.
4. Even-Sapir E, Martin RH, Barnes DC, Pringle CR, Iles SE, Mitchell MJ. Role of SPECT in differentiating malignant from benign lesions in the lower thoracic and lumbar vertebrae. *Radiology* 1993;187:193–198.
5. Slizofski WJ, Collier BD, Flatley TJ, Carrera GF, Hellman RS. Painful pseudarthrosis following lumbar spinal fusion: detection by combined SPECT and planar bone scintigraphy. *Skel Radiol* 1987;16:136–241.
6. Bellah RD, Summerville DA, Treves ST, Micheli LJ. Low back pain in adolescent athletes: detection of stress-injury to the pars interarticularis with SPECT. *Radiology* 1991;180:509–512.
7. Collier BD, Johnson RP, Carrera GF, et al. Painful spondylolysis or spondylolisthesis studied by radiography and single-photon emission computed tomography. *Radiology* 1985;154:207–211.
8. Gates GF. SPECT imaging of the lumbosacral spine and pelvis. *Clin Nucl Med* 1988;13:907–14
9. Nagele-Wohrle B, Nickel O, Hahn K. SPECT bone scintigraphy of benign and malignant lesions of the spine. *Neurosurg Rev* 1989;12:281–283.
10. Onsel C, Collier BD, Kir KM, et al. Increased sacroiliac joint uptake after lumbar fusion and/or laminectomy. *Clin Nucl Med* 1992;17:283–287.
11. Ryan PJ, Evans PA, Gibson T, Fogelman I. Chronic low back pain: comparison of bone SPECT with radiography and CT. *Radiology* 1992;182:849–854.
12. Ryan RJ, Gibson T, Fogelman I. The identification of spinal pathology in chronic low back pain using single photon emission computed tomography. *Nucl Med Commun* 1992;13:497–502.
13. Buscombe RJ, Summers B, Edgar M, Hogg P, Ell PJ. A comparison of SPECT and CT in the investigation of chronic low back pain [Abstract]. *Nucl Med Commun* 1990;11:199.
14. Ryan PJ, Gibson T, Fogelman I. Planar and SPECT imaging in the lumbosacral spine in patients with low back pain [Abstract]. *Nucl Med Commun* 1990;11:198–199.
15. Ryan PJ, DiVadi L, Gibson T, Fogelman I. Facet joint injection in patients with low back pain and increased facetal joint activity on bone scintigraphy with SPECT: a pilot study [Abstract]. *Nucl Med Commun* 1992;13:401.
16. Littenberg B. Technology assessment. *Acad Med* 1992;67:424–428.
17. Kendall M, Stuart A. *The advanced theory of statistics*, 4th ed. vol. 2. New York: Macmillan; 1979:587.
18. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med* 1978;299:926–930.
19. Hulley SB, Cummings SR. *Designing clinical research*. Baltimore: Williams and Wilkins; 1988.
20. Begg CB, McNeil BJ. Assessment of radiologic tests: control of bias and other design considerations. *Radiology* 1988;167:565–569.