AJNR

CT, myelography, and phlebography in the detection of lumbar disk herniation: an analysis of the literature.

J W Kardaun, J Schipper and R Braakman

AJNR Am J Neuroradiol 1989, 10 (5) 1111-1122
http://www.ajnr.org/content/10/5/1111

CT, Myelography, and Phlebography in the Detection of Lumbar Disk Herniation: An Analysis of the Literature

Jan W. P. F. Kardaun[1]
Jaap Schipper[2]
Reinder Braakman[3]

Despite the large number of reports on the relative usefulness of various radiographic procedures for the diagnosis of lumbar disk herniation, there has been no consensus of opinion on the best imaging procedure. Different study designs, including criteria for patient selection and retrospective consideration of patients who underwent surgery only, hamper direct comparisons between studies. A major drawback is the common use of "accuracy" as a measure of quality. We reviewed the CT, myelographic, and phlebographic findings in lumbar disk herniation published since 1970. After the reports were systematically classified and assessed for quality, the results became more coherent.

Many results tend to be sensitive and not very specific. We found there was no clear difference in the overall diagnostic quality of phlebography, myelography, and CT.

***AJNR* 10:1111–1122, September/October 1989**

The two most common causes of radicular compression syndrome are a herniated intervertebral disk and stenosis of the spinal canal. Over the years the ability of several radiologic investigations like myelography, epidural phlebography, CT, and more recently MR imaging to detect either of these abnormalities has been studied. However, there has been no consensus of opinion on the best imaging procedure. Differences in study design and methods of analysis have hindered comparisons.

The objective of this article was to assess the extent to which the capability of CT, myelography, and phlebography to visualize lumbar disk herniation can be derived from the literature and to explain the discrepancies in results. To this aim, special attention was paid to the way the results in the different studies were obtained.

Materials and Methods

Selection of Literature

The advent of safe water-soluble contrast media limits the literature review to 1970 and later. The first CT articles appeared in 1976 [1] and 1977 [2]. CT developments have been very rapid, especially since 1983.

There is another X-ray-based procedure for diagnosing herniated disks, phlebography, that, despite its history of about 30 years [3], never became as popular as myelography. When catheterized [4, 5], selective [6, 7], and later double-sided catheterized [8] phlebography were introduced, the results were claimed to be equal or superior to myelography. However, since 1980, CT has had a more dominant place in the literature than phlebography. Though it seems that phlebography has been overtaken by the newer technologies, it is included in this review to determine its relative diagnostic accuracy. Because MR is still evolving, it is not included in this review, unless it is included in an article that is cited for other reasons.

A minimum number of 20 patients in a study was required for inclusion of an article in this analysis, unless the article announced an innovation. Only publications in regular journals

Received September 15, 1988; revision requested November 18, 1988; revision received December 30, 1988; accepted January 10, 1989.

[1] Department of Public Health and Social Medicine, Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands. Address reprint requests to J. W. P. F. Kardaun.

[2] Department of Radiology, Leyden University Hospital, Rijnsburgerweg 10, 2333 AA Leyden, the Netherlands.

[3] Department of Neurosurgery, University Hospital Dijkzigt, Dr Molewaterplein 40, 3015 GD Rotterdam, the Netherlands.

0195–6108/89/1005–1111

were considered—no proceedings, textbooks, or monographs were searched.

There is an important difference between assessing the accuracy of one department or one study and trying to assimilate the results from several publications into a more general statement. Therefore, before presenting the concrete results of our review, we will discuss the requirements of a study if it is to play a useful part in comparative literature.

Requirements of the Literature

The requirements of the reviewed investigations can be considered from different viewpoints. Several aspects are evaluated: (1) the selection criteria of the patient; (2) the pro- or retrospective design; (3) the evaluation criterion of the test, that is, the verification of the diagnosis; and (4) the measure for the results. These points are discussed in reverse order, as the reasons for the first steps in the design of a study are best understood by knowing the aim of the investigation.

The measurement of the predictive power of diagnostic procedures.—In the reviewed literature, the usual measure was "accuracy," that is, the number of all correct diagnoses divided by the total number of diagnoses. A correct diagnosis can be a suspected herniation at myelography that is also found at operation. A correct diagnosis can also be the assumption of absence of herniation on a radiologic procedure, which is confirmed either by operation (e.g., on clinical grounds or because of a positive CT) or by the transitory and noniterative character of the complaints. Accuracy is a sufficient measure to determine whether a particular physician or department makes many or few errors, but lumping together these two types of correct results hides important and interesting information: We do not know whether the accuracy pertains to a population with many or only a few diseased persons.

Assume, for example, that a population has a 90% prevalence of herniated disk (as is realistic after screening by clinical signs and other means, e.g., complaints that are resistant to a "lege artis" conservative therapy), then it would be a good guess when a myelogram was completely uninformative (e.g., was lost) to pretend that a herniated disk was seen. In fact, an accuracy of 90% could be achieved (and a lot of time saved) by completing the radiographic reports for all patients as having a positive myelogram, before the myelogram was seen. (This is only a hypothetical example, of course.)

We need to introduce a few symbols and the concept of the two-by-two table to determine what would be a better measure of the predictive power of a procedure. Assuming that a disease (herniation, stenosis) is either present or absent (D+ and D−), and that a diagnostic procedure did or did not predict the same disease (T+ and T− for a test, i.e., myelography, CT, or phlebography), there are four possible combinations for each test.

	T+	T−	
D+	a	b	p
D−	c	d	q
	m	n	N

In a clinical context a, b, c, and d are usually called the true-positive, false-negative, false-positive, and true-negative test results. The accuracy of the procedure is usually defined as $(a + d)/N$, the prevalence of the disease is p/N, the specificity is d/q, the sensitivity a/p.

We can now easily see why the accuracy is only a limited measure: it does not tell us anything about sensitivity and specificity. It makes all the difference whether the erroneous myelography reports are concentrated in b or in c, for in the first case we have lower specificity and in the other a lower sensitivity. The quadruple specificity, sensitivity, prevalence, and N completely describe the two-by-two table.

There are more ways to look at this table: usually a practicing clinician is confronted with the question: If my test is positive (or negative), what is the probability of disease or no disease? The probabilities a/m and d/n are usually referred to as the predictive value of a positive test (PV+) or a negative test (PV−), respectively. In order to describe the table completely, we need in addition to the PV+ and PV− (and N) a fourth quantity, such as the ratio m/n. We are more interested in the prevalence (p/N) of the disease than in this m/n because it has a more direct interpretation.

For all these ways of assessing the effectiveness of a radiographic investigation, all four cells in the two-by-two table must be known. This not only follows the above simple calculations, but it is also common sense. If we want to know the conclusions that can be drawn from a "positive" or "negative" radiograph, then both categories of patients should be followed to see what is really the (likely) cause of the radicular syndrome.

This leads immediately to two other requirements: the evaluation of the patients and the pro-/retrospective study design.

Evaluation of the patient.—All patients who had myelography or CT should have verification of their diagnosis. For those patients likely to have herniation, surgical verification often will be available, but for others at least an attempt should be made to see whether their follow-up gives an indication of the presence or absence of herniation. Of course, this is only an approximation of the anatomic situation, but it is better than ignoring the issue at all, and it often will give a fair clue to the right answer.

Pro- or retrospective study design.—The above leads naturally to the requirement that studies for evaluating diagnostic test procedures ought to be "prospective," that is, they should assess all patients who present themselves to the study. (Retrospective studies are to be used to determine etiology when we want to know whether some event/exposure caused a disease.) The question is not what to do with patients in whom herniation was found, but to know what to do with a patient with signs and history that are likely to be due to a herniation.

Patient selection.—There is a wide range of indications for studying patients with CT, myelography, or phlebography because of radicular compression complaints. In some centers, all patients with suspected herniation will have one of these investigations. In other centers, only patients with clear symptoms who are to undergo surgery routinely have a CT scan or myelogram. In yet other centers, patients with very clear symptoms undergo surgery on clinical presentation only. To compare information about history and clinical signs in reports from several centers, it is useful to classify each patient into one of a few groups varying from slight to clear clinical evidence of herniation.

Blind vs clinical information.—It is important to know whether the radiographic reports are completed with or without the use of clinical information and other tests. Depending on the goal of the analysis, both ways can be defended. The combined use is good if one wants to know whether this combination is sufficient for making the decision to operate (and where) or not to operate, and if one is interested in an analysis that should be more applicable in a clinical setting. The blinded approach is preferable if one wants to know the contribution of each test in the decision making.

Classification of the Articles

The articles selected for review were classified with respect to the following points (Table 1): (1) whether they included, in some way, the use of clinical, myelographic, CT, phlebographic, and surgical data; (2) what selection criteria were used for inclusion of patients in the study; (3) whether there was an apparent bias in the selection; (4) whether the verification criteria for the diagnosis were strict or lax; (5) whether the data were, either in the text or in a table, presented in a complete way, so that a complete two-by-two table could be (re)constructed; (6) if a table could be reconstructed, the prevalence, sensitivity, and specificity of the test; (7) year of publication; (8) number of patients in the study (if some group of patients was not analyzed at all, but excluded right away, then the number in the table may be lower than that mentioned in the article or title); (9) whether the test was evaluated blindly or together with clinical information; (10) whether the analyses were done per patient, or per level; and (11) how equivocal tests were handled. Many articles analyzed several groups of patients. In most cases the smaller group was a subgroup of the larger.

Results

Comments on Selected Reports

The relevant literature is summarized briefly and evaluated relative to our research question. Though almost all reports did provide at least some measure of prediction, mostly accuracy, that often was not the main point of interest of the articles. In articles with an emphasis on technical, anatomic, or differential diagnostic aspects of a procedure, or on complications, rare cases, etc., the presentation of the numeric data and the accuracy of the radiologic procedures was often secondary. By looking mainly at the presentation of the data and reproducibility of the predictive power of the diagnostic method, it is not possible to give full credit to the overall quality and importance of the articles. The articles are summarized chronologically highlighting new information not included in earlier papers.

In 1970, Hudgins [9] analyzed 490 patients admitted for low back pain or leg pain and 102 patients who had lumbar myelograms for other reasons (controls). Hudgins asserted that most articles on this subject use the wrong approach, that is, they study a population of (surgically) proved herniations and do not assess the predictive value of a radiographic procedure. Besides having controls, Hudgins followed all operated patients. In addition to giving the predictive value of myelography and giving complete data, several interesting questions were discussed that are outside the scope of this review.

In 1974, Gargano et al. [6] analyzed 32 patients who had both myelograms and phlebograms and who underwent surgery for herniated disks. The emphasis of this article was on anatomy of and indications for phlebography, and much of the older literature was reviewed. A tabulation of the data showed better results for phlebography, but their population consisted of patients who previously had "clinical lumbar disk disease, but negative or equivocal myelograms." This gave a bias in favor of phlebography.

In 1976, MacNab et al. [11] studied a group of 110 patients with symptoms of herniation who underwent phlebography. They considered the myelograms of the 50 patients without prior disk surgery and found a diagnostic accuracy of 98% and 90% for phlebography and myelography, respectively. Their emphasis was more on anatomy and phlebographic technique than on numeric evaluation. Their data could not be reconstructed.

In 1977, Mohsenipour et al. [13] reviewed the myelograms of 500 patients. Three hundred seventy-four patients with a clearly positive myelogram underwent surgery, as did 27 of 85 patients with an equivocal myelogram. It is unclear whether none of the 41 patients with a clearly negative myelogram underwent surgery on clinical grounds. A few complications were reported, and a low threshold was favored for obtaining a myelogram, arguing that clinical signs cannot predict the level of a herniation well. The interpretation of a partial or complete block was elaborated with instructive images. Their data are not really complete, but a reconstruction of their accuracy is 72%. From the context, it is assumed that they applied a strict criterion as to the level of the herniation.

Also in 1977, Moringlane et al. [14] reviewed 140 patients who underwent operations for herniated disks. They described the population by specifying that all of the patients had had (several) conservative treatments, except for the patients with emergency symptoms. The patients agreed before myelography to undergo surgery if an abnormality were to be found. This was probably a population with a high prevalence. Complete data were presented, and the accuracy was 95% (using strict criteria as to level, side, and cause). The authors mentioned the use of the prior probabilities for discriminating between different levels and mentioned for the first time in a numerical analysis the findings at myelography and/or operation of multiple herniations (46% of their patients). They described the myelographic and surgical techniques and provided useful illustrations and differential diagnoses of the myelographic findings.

In 1978, Roland et al. [15] described phlebographic anatomy, technique, and diagnostic criteria. For 111 patients (chosen from 240 with unstated indications for selection) with "mainly . . . previous ambiguous or normal myelography," they compared the diagnostic performance of phlebography and myelography. Both per patient and per level analyses were provided, and all the data were presented. This comparison was heavily biased against myelography, but the authors correctly concluded that additional phlebography is useful for the five types of cases they mention.

In 1979, Cook and Wise [16] compared the findings of plain radiography and myelography in 50 operated patients, 49 of whom had lumbar disk protrusions. It is unclear how the patients were selected. Some patients had more than one herniation. There were some negative explorations, but it is not clear whether this was caused by multiple-level laminectomies. Of 36 patients who had myelography but no surgery, complications were reported, and related to the use of oil- and water-based media. No follow-up was undertaken to find additional signs of herniation. There was a detailed discussion of the discrepancies in their material, and the literature was reviewed, with some emphasis on technique. They elaborated on how to interpret discrepant findings in terms of clinical usefulness and pointed out that equivocal findings should be considered incorrect as far as clinical use is concerned. With these strict criteria they had an accuracy of 92% (based on levels).

TABLE 1: Summary of Literature Assessing Imaging Methods in Lumbar Disk Herniation

Year of Publication Reference	Sources Considered	No. and Units of Observations	Direct Comparisons	Blinded Evaluations	Apparent Favoring Bias	Selection Criteria	Treatment of Equivocal Results	Severity of Diagnostic Criteria	Prevalence	Sensitivity	Specificity
1970											
(1) Hudgins [9]	M,CI,S,FU	309 pts	M/S+FU	Not known	NA	Combination[a]	Not stated	Medium	–	–	–
		135 pts	M/S	Not known	NA	Operated pts	Not stated	Medium	0.79[b]	0.75	0.90
1974											
(2) Gargano et al. [6]	M,P,S	32 pts	M/S	Not known	P	Clinical findings[c]	Not stated	Medium	0.94	0.63	0.50
		32 pts[d]	P/S	Not known	P	Clinical findings[c]	Not stated	Medium	0.94	0.93	0.50
1976											
(3) Drasin [10]	P,S	19 levels	P/S	Not known	NA	Operated pts[e]	Own category	Medium	0.84	0.94	0.33
(4) MacNab et al. [11]	M,P,S	50 pts	M/S	Not known	None apparent	Operated pts[f]	Not stated	Medium	0.74	0.95	0.77
		50 pts[d]	P/S	Yes	None apparent	Operated pts[f]	Not stated	Medium	0.74	0.97	0.85
(5) Miller [12]	P,S	38 pts	P/S	Not known	NA	Operated pts[e]	Not stated	Medium	0.92	0.86	1.00
1977											
(6) Mohsenipour et al. [13]	M,S	401 pts	M/S	Yes	None apparent	Operated pts[g]	Not stated	Medium	0.93	0.95	0.36
(7) Moringlane et al. [14]	M,S	140 pts	M/S	Not known	NA	S candidates[h]	Not stated	Severe	0.96	0.98	0.20
1978											
(8) Roland et al. [15]	M,P,S	111 pts	M/S	Not known	P	Unclear	Not stated	Medium	0.91[i]	0.40	0.40
		111 pts[d]	P/S	Not known	P	Unclear	Not stated	Medium	0.91[i]	0.93	0.60
1979											
(9) Cook and Wise [16]	M,S,PR	62 levels	M/S	Not known	NA	S verified dx	Own category	Severe	0.84	0.96	0.70
(10) Meyenhorst [8]	M,P,S[j]	339 levels	P/M	Not known	None apparent	Unclear	Own category	Several	–	–	–
		151 levels	M/S	Not known	None apparent	Operated pts	Own category	Several	0.52[k]	0.81	0.94
		151 levels[d]	P/S	Not known	None apparent	Operated pts	Own category	Several	0.52[k]	0.97	0.87
1980											
(11) Lotz et al. [17]	M,P,S	37 levels	M/S	No	None	Operated pts	Own category	Medium	0.86[l]	0.81	1.00
		37 levels[d]	P/S	No	None	Operated pts	Own category	Medium	0.89[l]	0.94	0.50
		37 levels[d]	M/S	Yes	None	Operated pts	Own category	Medium	0.85[l]	0.71	0.91
		37 levels[d]	P/S	Yes	None	Operated pts	Own category	Medium	0.84[l]	0.71	0.42
		37 levels[d]	M/S	No	None	Operated pts	Own category	Medium	0.85[l]	0.75	0.82
		37 levels[d]	P/S	No	None	Operated pts	Own category	Medium	0.84[l]	0.74	0.67
(12) Thijssen et al. [18]	M,S	104 pts	M/S	No	NA	Operated pts[m]	Own category	Medium	0.91[b]	0.98	1.00
1981											
(13) Gulati et al. [19]	PCT,MCT,M	15 levels	PCT/M	Not known	M	CI + M signs	Own category	Medium	–	–	–
		15 levels[d]	CTM/M	Not known	M	CI + M signs	Own category	Medium	–	–	–
1982											
(14) Anand and Lee [20]	PCT,MCT,M	25 pts	PCT/M	Yes	None apparent	Unclear	Not stated	Medium	–	–	–
		75 pts	CTM/M	Yes	None apparent	Unclear	Not stated	Medium	–	–	–
(15) Claussen et al. [21]	M,CT,S	41 pts	CT/M	Not known	None apparent	Susp LDH	Not stated	Medium	–	–	–
		26 pts	CT/S	Not known	None apparent	Operated pts	Not stated	Medium	0.92	0.88	0.50
		23 pts[d]	M/S	Not known	None apparent	Operated pts	Not stated	Medium	0.91	0.81	0.50
(16) Fries et al. [22]	M,CT,S	192 levels	M/S	Not known	None apparent	S verified dx	Not stated	Medium	0.91[i]	0.87	0.89
		227 levels	P/S	Not known	None apparent	S verified dx	Not stated	Medium	0.90[i]	0.92	0.78
(17) Haughton et al. [23]	CT,M,S	55 levels	M/S	Not known	None	Back pain/sciatica	Own category	Medium	0.55	0.93	0.64
		55 levels[d]	CT/S	Yes	None	Back pain/sciatica	Own category	Medium	0.55	0.97	0.68
(18) Jepson et al.	M,S,CI[n]	55 pts	M/S	Not known	NA	Operated pts	None	Medium	0.89	0.90	0.83

[24]		55 pts[d]	M/S	No	NA	Operated pts	Not stated	Severe	0.89	0.61	0.83
(19) Sachsenheimer [25]	M,CT	35 pts	M/CT[o]	Not known	None apparent	S verified dx	Own category	Medium	–	–	–
1983											
(20) Moufarrij et al. [26]	M,CT,FU	50 pts	CT/S[p]	Not known	None apparent	Operated pts[q]	Not needed[r]	Medium	0.90	0.62	0.80
		46 pts[d]	M/S[p]	Not known	None apparent	Operated pts[q]	Not needed[r]	Medium	0.85	0.82	0.43
1984											
(21) Bell et al. [27]	M,CT,S	122 pts	M/S,CT/S[s]	Yes	None apparent	S verified dx	Weighted	Mixed[t]	–	–	–
(22) Bosacco et al. [28]	Cl,CT,M,S	134 pts	Cl/CT/M/S	No	None apparent	Susp LDH[q]	Not stated	Medium	–	–	–
		52 pts[d]	CT/S	No	None apparent	Susp LDH[q]	Not stated	Medium	0.96	0.92	1.00[u]
		52 pts[d]	M/S	No	None apparent	Susp LDH[q]	Not stated	Medium	0.96	1.00	0.50[u]
(23) Valat et al. [29]	P,S	104 pts	P/S	Not known	NA	Operated pts	Not stated	Medium	0.95[j]	0.99	0.20
1985											
(24) Kampmann et al. [30]	M,CT,S	36 levels[v]	CT/M	Not known	None apparent	Sciatica[h]	Not stated	Medium	–	–	–
		31 levels	M/S	Not known	None apparent	Operated pts	Not stated	Medium	0.90	0.96	0.33
		122 levels[v]	CT/S	Not known	None apparent	Operated pts	Not stated	Medium	0.91	0.99	0.57
1986											
(25) Modic et al. [31]	M,CT,MR	151 levels	CT/MR	Yes	None apparent	Susp LDH[w]	Not stated	–	–	–	–
		218 levels	M/MR	Yes	None apparent	Susp LDH[w]	Not stated	–	–	–	–
		42 pts	CT/S	Yes	None apparent	Operated pts[w]	Not stated	Mixed[x]	1.00[y]	0.83	–
		45 pts	M/S	Yes	None apparent	Operated pts[w]	Not stated	Mixed[x]	1.00[y]	0.72	–
		62 pts	MR/S	Yes	None apparent	Operated pts[w]	Not stated	Mixed[x]	1.00[y]	0.83	–

Note.—M = myelography; Cl = clinical; S = surgery or surgically; FU = follow-up; P = phlebography; PR = plain radiography; PCT = plain CT; MCT = metrizamide CT; MR = MR imaging; pts = patients; NA = not applicable (only one comparison in the study); dx = diagnosis; Susp LDH = suspected lumbar disk herniation.

[a] Combination of controls (e.g., patients having myelography for other reasons) and surgically and conservatively treated sciatic patients.

[b] Data reconstructed by reviewers.

[c] Patients with clinically suspected herniated disk, but "negative or equivocal" myelograms.

[d] Same as or included in patient or level group immediately above.

[e] With normal or equivocal myelograms.

[f] Operated patients with myelograms and phlebograms, but without prior surgery.

[g] Includes only patients with positive or dubious myelograms; it is not stated whether patients with negative myelograms had surgery.

[h] Most patients had had unsuccessful conservative treatment.

[i] Data also specified by level.

[j] Distinction made between single- and double-sided phlebography; the data in the table reflect double-sided phlebography.

[k] Data reconstructed by reviewers. Dubious and nondiagnostic cases split up by marginals.

[l] Detailed data (specified by level and observer) were aggregated by the reviewers for all variants of Lotz et al. Equivocal results were divided, according to marginals, into positive and negative categories. Comparisons 1 and 2 = original clinical interpretation; comparisons 3 and 4 = blind interpretation; comparisons 5 and 6 = reinterpretation with clinical information.

[m] "This paper presents . . . the findings of a series of lumbar myelographies of adequate quality."

[n] Data also available for patients who had surgery on clinical grounds only. This group was quite similar to that with myelography; those patients are not included here.

[o] No distinction between true positive and true negative.

[p] Extracted from four-way table.

[q] Patients with previous surgery or spinal stenosis were excluded.

[r] An equivocal category was not very important since several diagnoses were considered.

[s] Data were not tabulated because there was no distinction between true positive and true negative, nor between false positive and false negative.

[t] Severe, medium, and relaxed.

[u] Based on two cases.

[v] There were no errors in CT levels.

[w] Patients with prior surgery or known diagnoses were excluded.

[x] Severe and medium.

[y] There were no negative explorations, making the specificity unknown.

Also in 1979, Gershater and St. Louis [32] analyzed 1200 patients suspected of having lumbar disk herniations who underwent phlebography. The emphasis was on anatomy, technique, differential diagnosis, and complications of phlebography. It is unfortunate that the data in this very large series were insufficient to allow calculations of the test qualities, for example, the results of 243 phlebograms of nonoperated patients, the selection criteria for surgery (only 50% of all patients were operated), and the correctly negative findings.

Another 1979 publication, by Meyenhorst [8], gave a detailed analysis of anatomy and methods for phlebography, and provided extensive results of 63 surgically verified cases studied with both myelography and phlebography. In 113 patients, the relationship between the myelographic and phlebographic findings was assessed. There was a good analysis of which questions are important for the evaluation of the diagnostic value of radiographic procedures and a good literature review. Tabular material was ample, but it was hard to synthesize the many single aspects of the comparisons. The accuracies were 94% and 84% for double-sided phlebography and myelography, respectively.

In 1980, Lotz et al. [17] argued that many previous reports comparing myelography and phlebography concerned biased populations and did not provide surgical findings. They obtained examination results from 50 patients with clinical signs of herniated disks at L4–L5 or L5–S1 and evaluated the results with and without clinical information. There was no follow-up of nonoperated patients. The data presented were concise and complete. Not counting unsuccessful and equivocal radiographic findings, accuracies ranged from 69 to 89% for myelography and 59 to 89% for phlebography. Considering complications and economic factors, they considered myelography to be the procedure of first choice.

Also in 1980, Thijssen et al. [18] analyzed the myelograms of adequate quality of 104 patients who underwent surgery and partially analyzed the data of 143 patients who did not undergo surgery. It was not stated how many inadequate myelograms were discarded. These authors argued that myelography can be equal to phlebography, provided the technique is of high quality. Attention was given to the inconclusive results of myelography and phlebography, both in their findings and in the literature. An accuracy of 93–98% was stated, depending on the treatment of the inconclusive group. A complete tabulation of the data was provided.

Another 1980 article, by Williams et al. [33], analyzed 16 patients studied with CT who had herniated disks at operation, and mentioned 21 patients who had CT but did not undergo surgery. Since they had no false positives, nor "did false negatives come to their attention," their analysis was oriented toward technical details, exact localization of the herniation, differential diagnosis, and advantages or disadvantages of CT vs myelography.

In 1981, Gulati et al. [19] described 10 patients with clinically herniated disks and at least one positive level at metrizamide myelography. All these patients had a high-resolution CT study. This is one of the first articles on CT for lumbar herniated disk that provided a numeric analysis. The findings were given for eight operated patients; the other two did not have follow-up. All operated patients had positive CT findings (with perhaps one exception that they classified as an artifact).

Also in 1981, a short, but clear article by Hanson [34] described 22 patients with clinically suspected herniated disks who underwent CT. Although only 22 patients were studied, the study was complete in that the findings of both the seven operated patients and the follow-up of the 15 nonoperated patients were provided. The merits of CT over myelography were discussed.

Another 1981 article, by Harley et al. [35], reviewed 81 patients who had CT because of suspected disk herniation and did not have prior disk surgery. Twenty-six patients underwent surgery, 16 of whom had positive and 10 of whom had negative CT findings. No follow-up was provided for the others. These authors stressed the importance of fragmented disks and made a plea for a distinction between normal, bulging, and herniated disks or equivocal results for the procedures. Their data could not be reconstructed or recalculated.

In 1982, Anand and Lee [20] analyzed 100 patients with suspected lumbar disk disease. All of the patients underwent myelography and plain or metrizamide CT. In their comparison of low-dose metrizamide CT and plain CT, the former was found to be superior. Criteria for the CT diagnosis of herniated disk were refined. All the radiographs were blindly reviewed by two observers, and complete data were provided on the comparison of CT with myelography (75% agreement), but the analysis of the 53 operated patients was not very clear. There was no follow-up of the other 47 patients. Because the surgical findings were all positive, it is hard to assess the false-positive rate of the procedures. When they stated that "in our study, when CT and myelography agreed . . . there was no problem in diagnosis, and either test could have been used," this cannot be applied to three patients in both the table and text who had negative CT, negative myelography, and positive operations.

Also in 1982, Claussen et al. [21] presented the CT data of 77 patients, of whom 41 had myelograms and 26 underwent surgery. Twenty-three of the operated patients had two investigations, for which complete data are presented. There was a general introduction on the research question, and attention was given to CT anatomy and criteria.

In another 1982 article, Dublin et al. [36] reviewed the results of plain film metrizamide myelography, CT metrizamide myelography, and plain CT. They found 106 patients with both plain film and CT metrizamide myelograms out of 736 with spinal CT scans. In about half the 106 patients, thoracic or cervical levels were studied. The plain film and CT metrizamide myelograms were evaluated blindly and compared with each other and with clinical/surgical findings. The authors found that CT metrizamide myelography was superior to plain film metrizamide myelography in 40%, equal in 50%, and inferior in 10% on the basis of a decision-oriented criterion, that is, the findings should change the surgical approach or the clinical management of the patient. They stressed the importance of a high-resolution unit and provided several reasons for problems of interpretation. One of the most interesting remarks was that with CT the patient is investigated in the position most likely to relieve pain. Their data were not included in Table 1 because their problem was so different (spinal disease in general) from that in the other studies that their classifications were not comparable, and their selection of patients for the investigations was not clear. Though they provided a schema for the management of spinal neurologic problems, they did not indicate whether CT or plain film metrizamide myelography should be performed first and in what conditions the other investigation should be performed. At least for the plain CT vs CT metrizamide myelography analysis, it was suggested that CT metrizamide myelography be performed only after nondiagnostic plain CT. Their high proportions of osseous hypertrophy and neoplasms suggest a different population from that in the other studies.

A fourth 1982 article, by Eldevik et al. [37], compared the reports of CT and myelography with and without clinical information. For 107 previously reported patients (52 were operated), they completed their data with clinical and blind reports for CT and myelography, respectively. There was special emphasis on the difficult diagnostic evaluation of patients with prior surgery. Both CT and myelography were interpreted more accurately without the clinical information. For all 107 patients the differences between the blind and clinically reported studies were compared. Blind studies were better in detecting herniation (for CT, marginally significant; for myelography, not significant). For the 52 operated patients, the operative findings were compared, and it was found that both CT and myelography had better results without clinical information, as CT had three more false-positive findings with clinical information and myelography had two more false-positive findings; however, they did not account for the fact that myelography with clinical information played a role in the indication for operation so that false-negative findings resulting from clinical information were less likely to be noted. This was not true for CT, but in the casuistic they accounted for five of the six differences, allowing us to classify those as three more false positives and two more true positives, which is not impressive. Their data did not allow for a complete reconstruction.

In 1982, Fries et al. [22] selected 188 patients (244 intervertebral spaces) with both CT-diagnosed and surgically confirmed herniated disks. The false-negative and false-positive findings consequently were limited to multiple-level patients. However, there was a detailed analysis of the distribution of herniation sites over levels and in levels. The capability of CT to diagnose central and migrated herniations was assessed. Techniques, interpretation of images, and an explanation for erroneous CT and myelograms were elaborated. Complete data were provided for CT and myelograms vs operations, and the problems of multiple locations and multiple diseases were addressed, but their claim of a complete lumbar spine examination (above L4–L5) is not supported by the use of clinical indications for the levels.

In another 1982 article, by Haughton et al. [23], data were collected on 107 patients referred for low back or sciatic pain. CT and myelographic findings were available for all patients, and operative findings were available for 52. Complete data were provided about the correlation and discrepancies of the diagnosis. Their population did not include patients with normal findings at surgery. The CT findings were evaluated blindly, and myelography was evaluated with clinical information. Strict criteria (site, level, and pathology) were applied, but good accuracies were achieved (84% and 88% for CT and myelography, respectively), which can be related to the probably strict criteria for operation. They argued that the most common differential diagnostic problem is formed by herniation vs bulging anulus and spondylosis, but argued that perhaps a nuclear fragment lodged behind the posterior ligament can be mistaken for a bulging anulus at surgery.

Also in 1982, Herkowitz et al. [38] compared myelography and phlebography in 30 patients with a surgically verified diagnosis of herniated disk or spinal stenosis. They clearly defined what constitutes herniation and stenosis and stated the correct (i.e., usual) definitions for sensitivity and specificity, but their tables used a less obvious interpretation of these rates, so their figures cannot be compared. Fortunately, they provided a table presenting their raw data, so that accuracies and other measures could be reconstructed. They discussed the variations in accuracies of phlebography and the usefulness and complications of metrizamide. They outlined a precise range of indications for phlebography.

Jepson et al. [24] studied two groups of patients, both undergoing surgery because of suspected lumbar disk lesions; one of these groups underwent myelography routinely and the other (1964–1968) did not. The authors argued that myelography did not contribute much to the indication for operation. In their discussion they reviewed this problem and pointed out that many articles with high accuracies are not applicable in clinical use. Their own data, which are complete, mentioned accuracies of 64–89% for myelography depending on the strictness of the criteria. They did find fewer herniations in the group with myelograms than with clinical signs only (>90% in the latter), though this might reflect a relaxed threshold for operations.

Nelson and Gold [39] reported 10 patients with negative myelograms and positive operative findings. Clinical histories and a detailed tabulation of data are included. Depending on the criteria, all 10 or eight of the 10 patients had a correct CT finding. However, the population was heavily biased against myelography, and there was no indication about the selection as to considering all similar cases.

Raskin and Keating [40] addressed the problem of whether CT or myelography should be performed first. In this study with both metrizamide and Pantopaque myelograms, 106 patients had both procedures within 6 weeks (apparently for sciatica). Thirty-nine percent had surgery; in the other patients CT was compared with myelography. There is a good description of CT criteria for herniated disk. They explained that, due to the different way of visualizing, CT and myelography often will not provide exactly the same results (as to precise localization), but often will give the same conclusions as to operation. They made a distinction between major and minor discrepancies but had a rather large number of inconclusive investigations. As for myelography, they argued that most of the inconclusive cases were caused by anatomic variations.

We believe these should be counted in the category "major discrepancies." They gave rather low accuracies, which was caused partially by applying strict criteria, but also by the low prevalence of herniation in their group. However, their data did not allow a complete construction.

Stoeter et al. [41] verified CT findings in 106 patients by either surgical or CT findings, using a rather large doubt category and plural diagnoses (stenosis and other osseous causes, postoperative status). Consequently, their data were not comparable. There was no distinction between correctly positive or negative findings. Their moderate accuracy (74%) likely could be explained by the rather strict criteria applied. Attention was given to localization (medial vs lateral). They offered a number of practical considerations as to when to use CT or myelography as the first investigation.

Tchang et al. [42] analyzed 52 patients with surgically confirmed disk herniations; 45 of them also had myelography. They make a distinction between original interpretation (actually used for patient management), interpretation with the use of clinical knowledge and other tests, and blinded evaluation. In their classification, the patients apparently had only a single abnormality. Their data were complete; however, their selected population did not contain operations with negative findings or patients without herniations. Their accuracies were 94% and 87% for CT and myelography, respectively. They described their technique and discussed the differential diagnosis of CT findings.

Also in 1983, Griebel et al. [43] reviewed 100 patients who underwent surgery because of herniated disks; all had CT scans. The authors found better results for CT than for myelography, bit did not mention the role of either in the decision to operate. Their classification of three cases as neither correct, false positive, or false negative, but as misinterpreted by the radiologist, is somewhat unusual. Their data did not allow a reconstruction.

Moufarrij et al. [26] selected 50 patients who underwent surgery for herniated disks, with exclusion of previously operated patients and those with suspected stenosis. CT findings were available for all patients and myelographic findings were available for 46. The patients were divided into five categories based on CT diagnosis, including spondylosis and stenosis. From their four-way table (CT vs myelography vs surgical findings vs outcome), the test qualities were reconstructed by the reviewers. CT appeared most accurate if the only finding was herniation.

The 1983 article of Pythinen et al. [44] reviewed a series of 214 patients who (apparently) had both myelography and CT (in this order) for different indications (177 for herniated disk). They found that CT gave additional information in 12%, and in 7% (16 patients) the therapy was changed by CT. Detailed data were provided about these 16 cases, but there was no follow-up of the nonoperated cases (61%). Their data did not allow a reconstruction.

In 1984, Bell et al. [27] reviewed the CT and myelographic findings in 122 patients with surgically proved pathology of herniated lumbar disk or stenosis or both. In this excellent article the herniated disk and the stenosis were considered for the first time together in a numeric evaluation of radiographic procedures, and the idea was developed of a local and global diagnosis. The images were evaluated blindly, which probably gave a lower accuracy. There was a range of accuracies, depending on strict or relaxed criteria. The literature review was extensive. It was not mentioned whether there were patients with the same diagnosis who had to be excluded because neither CT nor myelographic findings were available. The indications for operations were not mentioned.

In another 1984 study, Bosacco et al. [28] described 134 patients with suspected lumbar disk herniations (excluding patients with previous spinal surgery or likely stenosis). The CT and myelographic findings in these patients were evaluated with clinical information, but separately. For 61%, conservative treatment was chosen, reflecting either a different patient selection or a different indication for surgery from most of the other studies. For these patients, only a correlation between CT and myelography was presented, without specifying the fraction of correct positives and negatives. For 52 patients, the results of CT and myelography were compared with surgical findings. The data were complete for reconstruction of a two-by-two table. The study provided a description of the clinical symptoms and related them to CT, myelographic, and surgical findings. The important role of workmens' compensation in negative cases was illustrated. The relative merits of the procedures were discussed.

Also in 1984, Valat et al. [29] presented phlebographic data from 104 operated patients in comparison with surgical findings, in both a per patient and per level analysis (complete data). If clinical presentations and plain radiographs were consistent as to level and site, phlebography did not add much information.

In 1985, Kampmann et al. [30] analyzed the data of 158 patients with sciatica in light of a number of questions: CT and myelography were compared for 36 nonoperated patients (90% correspondence); for the operated patients the CT (and part of the myelographic) diagnoses were compared with surgical findings. The criteria for correct diagnosis were rather strict: the analysis was level-specific, but no errors were found in level or side of localization. The erroneous radiologic diagnoses were discussed, and the authors maintained that a myelogram would not have helped overcome the shortcomings of CT (based on three cases); they concluded that myelography is indicated only if clinical signs and CT disagree.

In 1986, Greenough et al. [45] followed 22 patients who had had CT because of clinically suspected lumbar disk herniations with negative or equivocal myelograms. Although the prospective approach was completed for all patients, the numbers were small and the selection bias (against myelography) clear, so we did not include this study in Table 1. It is interesting to note that only eight of the 12 CT-positive patients had surgery.

Also in 1986, Kratzer and Hipp [46] analyzed the CT and myelographic findings in 133 patients with a clinical suspicion of a herniated disk, of whom 93 had surgery. The relative advantages of CT and myelography were discussed in detail, with special attention to discrepancies between the clinical presentation and the radiologic findings, including recurrent herniation. The numeric data did not allow reconstruction of a two-by-two table.

In another 1986 article, Modic et al. [31] compared the MR, CT, and myelographic findings (without clinical information) in 60 patients with suspected lumbar disk herniation or spinal stenosis, but without prior known diagnosis or imaging results. Forty-eight of these patients had surgery, five were normal, and seven had other disease. There were no negative explorations, so the test qualities in this review are not applicable. The lack of negative explorations suggests a rather high threshold for operation or for referral to the radiologist, but this was not confirmed. The fact that all normal MR studies were also normal on CT and myelography contradicts this assumption. Because the criteria for correlation of the radiologic procedures were strict (two diagnoses, level-specific stenosis), there were still discrepancies between the procedures, which were treated and commented on for the various combinations. In brief, the agreement of CT and myelography with MR was 87% and agreement with surgical findings was 83, 83, and 72% for MR, CT, and myelography, respectively. The relative merits of the procedures were discussed in detail. The authors concluded that MR (with surface coil) is equal to myelography or CT and that no single procedure alone is adequate.

In 1987, Schoedinger [47] described the CT, myelographic, phlebographic, and electromyographic findings in 100 patients with surgically verified lumbar disk herniations. The various combinations of these tests have accuracies between 56 and 100%, but no single procedure predominated. For the most part, discrepancies in levels were disregarded. The possibility of false-positive results was mentioned only in the last section of the article, and there were not enough data to reconstruct the single test qualities.

Also in 1987, Voelker et al. [48] analyzed the metrizamide CT scans and myelograms of 80 patients who had either lumbar disk herniation or spondylosis at surgery. The images were evaluated blindly, which, together with the two surgical diagnoses under consideration, largely eliminated circular selection bias. The difference in accuracy between a per patient and a per level analysis was explored. Lumbar disk herniation and spondylosis did not differ in detectability between CT and myelography. Patients who had undergone surgery previously were shown to be measurably more difficult to diagnose. Their data did not allow a reconstruction of the test qualities, because the authors did not mention the existence or nonexistence of false-positive diagnoses and seemed to be aiming only at not missing a diagnosis. This also affected their treatment of combined use of the tests.

Analysis

The aspects described above are presented in Table 1 for those articles that contained data allowing calculation of sensitivity and specificity, or that contained otherwise complete numeric analyses. In the discussion that follows, the findings in the 25 articles listed in Table 1 are summarized.

The selection criterion was operated patients in 10 articles, operated patients with supplemental data in four articles, surgically verified diagnosis of herniation or stenosis in four articles, clinical presentation in five articles, and unclear in two articles. The selection criteria *operated patients* and *surgically verified patients* represent a retrospective, rather than a prospective selection. The number of available observations (either levels or patients) was less than 50 in seven articles, between 50 and 150 in 13, and more than 150 in five. Only one article had both a prospective selection criterion and more than 150 observations. Four papers had blind interpretation, two with a clinical interpretation and two with both types of interpretations on the same population. (The complement to Haughton et al. [23] is presented in a separate article [37] that is not included in Table 1.) The other 15 papers did not state whether they used clinical information. Four papers did apply more than one threshold for the diagnostic criterion.

Myelography and surgery were compared in 19 articles and phlebography and surgery in nine. CT and surgery were compared in seven articles.

ROC Analysis of Data

For those articles that contained a comparison of phlebography, CT, or myelography with surgery, the sensitivity and specificity, which are shown in Table 1, are also presented in a Receiver Operating Characteristic (ROC) space (Figs. 1 and 2) [49]. This allows us to see the difference (between studies) in both sensitivity and specificity, which must be always considered together.

The ROC space is divided into three regions (Fig. 2) to distinguish between relatively well and poor performing tests. The positions of the boundaries are rather arbitrary. The best and worse performing sixth parts are pragmatically called deviant. Though the number of observations in general is too small for the individual points (studies) to be compared, we can look at the total pattern that emerges by studying the extremes.

Lower-quality outliers.—Roland et al. [15] reported myelographic findings in "patients with a clinically suspected disk but negative or equivocal myelograms." Theirs is the only study with results that are below the diagonal D in Figure 2, which means that the test outcomes should be interchanged for optimal results. Gargano et al. [6] used the same biased selection criterion as Roland et al. In their fourth group, Lotz et al. [17] reported phlebographic results from a "blinded" experiment; the results were manipulated by the reviewers, in that the quite large "dubious" category was distributed over the cells according to the marginals. This degrades the results. The second comparison of Jepson et al. [24] is a variant with stricter criteria; their first study is in the medium region. Moufarrij et al. [26] and Claussen et al. [21] did not provide clear explanations for their (relatively) poor results.

High-quality outliers.—Thijssen et al. [18] considered only myelograms of "adequate quality," which renders the test results overly optimistic. The second comparison of MacNab et al. [11] was limited because patients without prior surgery were studied; in addition, the evaluation was blinded, making interpretation more difficult. In the studies of Miller [12] and Meyenhorst [8] (second and third comparisons), the selection criteria of the patients were not described in enough detail to attribute or deny any influence from it. The results in the third comparison of Bosacco et al. were quite possibly affected by artifacts, since specificity was based only on two cases.

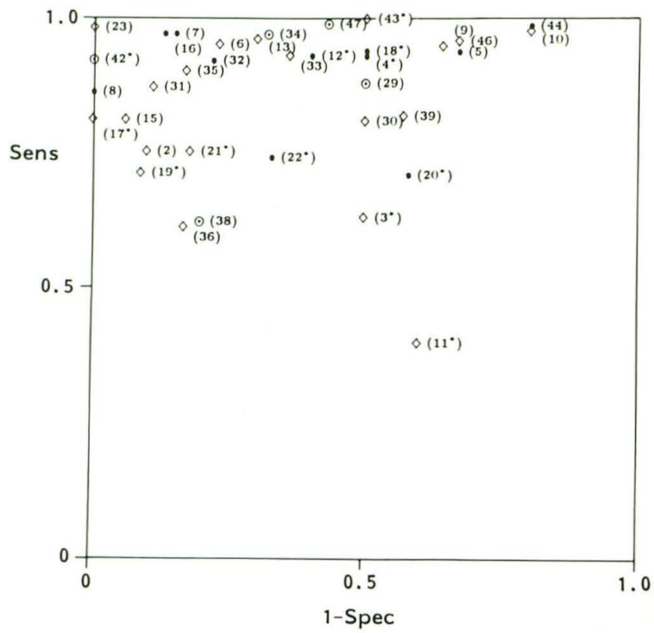

Fig. 1.—Test qualities from sensitivity (Sens) and specificity (1-Spec) of phlebography (*solid circles*), myelography (*diamonds*), and CT (*circles with dots*). Sensitivity (1—true-positive ratio) is on Y axis; specificity is on X axis. Numbers in parentheses correspond to 25 consecutive references listed in Table 1. Those with an asterisk denote studies with a marked bias or other reasons for being out of order.

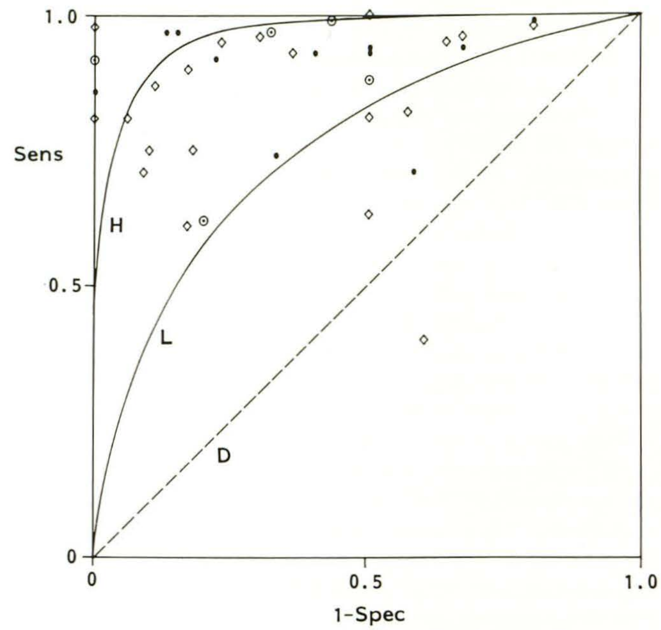

Fig. 2.—Test qualities from sensitivity (Sens) and specificity (1-Spec) of phlebography (*solid circles*), myelography (*diamonds*), and CT (*circles with dots*). Sensitivity (true-positive ratio) is on Y axis; specificity (1—false-positive ratio) is on X axis. Curves and broken line connect sensitivity/specificity pairs of tests with a constant quality, but with different cutoff points between normal and abnormal. Diagonal line (D) represents a complete uninformative test; low- (L) and high- (H) quality tests are represented by H and L curves.

If we exclude all articles with an apparent bias, and the article of Lotz et al. [17] (because of the manipulation by the reviewers), the remaining results show, despite their large differences in sensitivity or specificity, considerable coherence when regarded as a ROC curve. (Excluded articles are marked with an asterisk in Figure 1.)

In Table 1, it is hard to conclude that myelography is better or worse than phlebography. Only MacNab et al. [11], Meyenhorst [8], and Lotz et al. [17] provided direct comparisons between the two procedures on the same population and without apparent bias toward one. MacNab et al. and Meyenhorst found phlebography to be superior to myelography; in the Lotz et al. article, myelography was equal to or better than phlebography.

Discussion

Ever since its introduction in 1936 [50], myelography has been widely used for the detection of lumbar disk herniation. The introduction of alternative diagnostic tools did raise the question of whether myelography could be replaced by one of the newer methods and in what cases which procedure should be chosen.

Various studies in which different radiologic techniques are compared have tried to answer these questions. The conclusions of these studies, however, are sometimes conflicting. Because of this, the need became obvious for a review and analysis of the literature that tried to explain these discrepancies.

In the current review, we screened the available literature addressing the capabilities of myelography, epidural phlebography, and CT to detect lumbar disk herniation and spinal stenosis. Phlebography seems to have lost its place in daily practice with the introduction of CT (it was used especially for the detection of lateral and intraforaminal herniations). It is included in our study because of its importance in the spectrum of diagnostic reliability and its role in the comparative studies reported in the literature. MR imaging is not included because there were too few comparative studies with enough patients.

There are of course subjective elements in the analysis of the reviewed literature. Though we tried to maintain uniformity in our analysis, many articles include special circumstances or boundary-case presentations. Sometimes a different approach to what is clinically relevant prevented us from using data that were otherwise valid. It was not our intention to judge the quality of single articles, and, as has been mentioned, other qualities of these articles were overlooked.

Despite the necessarily subjective classification, most of the conclusions would remain the same if some papers were otherwise rated, because the pattern of this set of papers would not change. Though we tried to find all the relevant literature, it is possible that some data about myelography, CT, or phlebography are included in other articles, lengthy reports, or monographs with another scope; if so, these were not referenced in the reviewed papers.

Our analysis and review concentrated on the capability of the radiographic procedures to predict surgical findings, not whether one or more of these procedures are required before deciding to operate.

Only some of the literature was suitable for interstudy comparisons. An important problem was the definition of the study population. Some studies had only a small number of patients; very often it was not even mentioned whether the images were evaluated with our without clinical information or to what degree the results of the images influenced the decision to operate. It remains unclear how the results were influenced by important factors such as mild or strict criteria for correspondence of radiologic and surgical findings. Too often, emphasis was on accuracy as a measure of the results, which gives less information than the simultaneous consideration of sensitivity and specificity in a ROC space.

The interpretation of the points in a ROC space should be treated with caution: In the regular ROC analysis it is assumed that there is a means of verifying the disease, independent of the test. Consequently, *prevalence* here refers to the detected prevalence in a clinical subpopulation. This cannot be maintained for the investigations under study. In most of the articles, the radiologic diagnosis will have played some role in the decision to operate. However (especially in the investigations with two radiologic procedures), they will not have *determined* the decision. This difference can be seen between groups 4 and 5 and groups 2 and 6 of Lotz et al. [17].

Whatever other arguments played a role in the decision to operate, and whatever weights may have been attached to the test results, together they appear to have placed the cutoff points of the majority of these papers along one ROC curve, if articles with apparent bias are excluded. One possible interpretation of this phenomenon is that the values assigned by patients and physicians to false-positive and false-negative outcomes vary among the cited studies and that the tests are of the same order of quality, despite the differences in study design and technique; this cannot be ascertained.

The infulence of a positive radiologic procedure on the decision whether to operate will most likely will cause the stated accuracy to be too optimistic, and this effect will be more pronounced for the medium or poor tests. Poor tests, however, will tend to have less influence on the decision to operate.

In Figures 1 and 2 we see that there is more variation in the specificity (0.2–1) than in the sensitivity (0.6–1). Apparently in most of the studies there was a tendency not to miss a possible herniation. Although a threshold was never mentioned in the original publications, the existence of implied values attributed to false-positive or false-negative diagnoses is undeniable from the set of articles.

A literature analysis always excludes the most recent publications, which are always behind the most recent developments. This is more meaningful for CT, where developments have been very fast, than for myelography or phlebography. It is possible that if this analysis were repeated after a few years, CT would have a stronger position. However, based on the findings of the literature reviewed so far, there are no clear indications to consider myelography or CT or phlebography superior over the other studies in the diagnosis of herniated lumbar disk.

REFERENCES

1. Di Chiro G. computed tomography of the spinal cord after lumbar intrathecal introduction of metrizamide (computer-assisted myelography). *Radiology* **1976**;120:101–104
2. Coin CG, Chan YS, Keranen V, et al. Computer assisted myelography in disk disease. *J Comput Assist Tomogr* **1977**;1:398–404
3. Helander CG, Lindblom A. Sacrolumbar venography. *Acta Radiol* (Stockh) **1955**;44:410–416
4. Perey O, Lind J, Wegelius C. Phlebography of the intervertebral plexus. *Acta Orthop Scand* **1956**;25:228–233
5. Nathan MH, Blum L. Evaluation of vertebral venography. *AJR* **1960**;83:1027–1033
6. Gargano FP, Meyer JD, Sheldon JJ. Transfemoral ascending lumbar catheterisation of the epidural veins in lumbar disk disease. *Radiology* **1974**;111:329–336
7. LePage JR. Transfemoral ascending lumbar catheterization of the epidural veins. *Radiology* **1974**;111:337–339
8. Meyenhorst GCH. Transfemoral epidural double-catheter venography in the diagnosis of lumbar disc herniation [Thesis]. Nijmegen, the Netherlands: the Nijmegen Catholic University, **1979**
9. Hudgins WR. The predictive value of myelography in the diagnosis of ruptured lumbar discs. *J Neurosurg* **1970**;32:152–162
10. Drasin GF, Daffner RH, Sexton RF, Cheatham WC. Epidural venography: diagnosis of herniated lumbar intervertebral disc and other disease of the epidural space *AJR* **1976**;126:1010–1016
11. MacNab I, St. Louis EL, Grabias SL, et al. Selective ascending lumbosacral venography in the assessment of lumbar-disc herniation. *J Bone Joint Surg [Am]* **1976**;58-A:1093–1098
12. Miller MH, Handel SF, Coan JD. Transfemoral lumbar epidural venography. *AJR* **1976**;126:1003–1009
13. Mohsenipour I, Twerdy K, Pirker E. Eine Gegenuberstellung Rontgenbild-Operationsbefund bei lumbalen Bandscheibenerkrankungen. *ROFO* **1977**;127:540–543
14. Moringlane JR, Voigt K, Seeger W. Vergleich myelographischer und intraoperativer Befunde beim lumbalen Bandscheibenvorfall. *Neurochirurgia (Stuttg)* **1977**;20:199–208
15. Roland JR, Treil J, Larde D, et al. Lumbar phlebography in the diagnosis of disc herniations. *J Neurosurg* **1978**;49:544–550
16. Cook PL, Wise K. A correlation of the surgical and radiculographic findings in lumbar disc herniation. *Clin Radiol* **1979**;30:671–682
17. Lotz PR, Seeger JF, Gabrielsen TO. Prospective comparison of epidural venography and iophendylate myelography in the diagnosis of herniated lumbar disks. *Radiology* **1980**;134:127–132
18. Thijssen HOM, Rombouts JJM, Walder HAD. Diagnostic accuracy of lumbar myelography in the detection of lumbar disk herniations. *Diagn Imaging* **1980**;49:188–192
19. Gulati AN, Weinstein Z, Studdard W. CT scan of the spine for herniated discs. *Neuroradiology* **1981**;22:57–60
20. Anand AK, Lee BCP. Plain and metrizamide CT of lumber disk disease: comparison with myelography. *AJNR* **1982**;3:567–571
21. Claussen C, Grumme T, Treisch J, Lochner B, Katzner E. Die Diagnostik des lumbalen Bandscheibenvorfalls. *ROFO* **1982**;136:1–8
22. Fries JW, Abodeely DA, Vijungco JG, et al. Computed tomography of herniated and extruded nucleus pulposus. *J Comput Assist Tomogr* **1982**;6:874–887
23. Haughton VM, Eldevik OP, Magnaes B, et al. A prospective comparison of computed tomography and myelography in the diagnosis of herniated lumbar disks. *Radiology* **1982**;142:103–110
24. Jepson K, Nada A, Rymaszewski L. The role of radiculography in the management of lesions of the lumbar disc. *J Bone Joint Surg [Br]* **1982**;64-B:405–408
25. Sachsenheimer W, Hamer J, Muller HA. The value of spinal computed tomography in diagnosis of herniated lumbar discs. *Acta Neurochir* **1982**;60:107–114
26. Moufarrij NA, Hardy RW Jr, Weinstein MA. Computed tomographic, my-

elographic and operative findings in patients with suspected herniated lumbar discs. *Neurosurgery* **1983**;12:184–188
27. Bell GR, Rothman RH, Booth RE, et al. A study of computer assisted tomography. II. Comparison of metrizamide myelography and computed tomography in the diagnosis of herniated disk and spinal stenosis. *Spine* **1984**;9:552–556
28. Bosacco SJ, Berman AT, Garbarino JL, Teplick JG, Peyster R. A comparison of CT scanning and myelography in the diagnosis of lumbar disc herniation. *Clin Orthop* **1984**;190:124–128
29. Valat JP, Gatti P, Saindelle A. Phlébographie lombaire. Apport au diagnostic des lombosciatiques d'origine discale. *Sem Hop Paris* **1984**;60:539–542
30. Kampmann H, Schroedl P, Spranger M. Diagnostik lumbaler Bandscheibenvorfälle durch Computertomographie. Eine klinische Vergleichsstudie zwischen myelographischen, computertomographischen und operativen Untersuchungsergebnissen von 158 Patienten. *Rontgenblatter* **1985**; 38:387–391
31. Modic MT, Masaryk T, Boumphrey F, Goormastic M, Bell G. Lumbar herniated disk disease and canal stenosis: prospective evaluation by surface coil MR, CT, and myelography. *AJNR* **1986**;7:709–717, *AJR* **1986**;147:757–765
32. Gershater R, St. Louis E. Lumbar epidural venography. *Radiology* **1979**;131:409–421
33. Williams AL, Haughton VM, Syvertsen A. Computed tomography in the diagnosis of herniated nucleus pulposus. *Radiology* **1980**;135:95–99
34. Hanson RD. Computed tomography in the diagnosis of lumbosacral disc herniation. *Ariz Med* **1981**;38:839–842
35. Harley WD, Sava GF, Fleming RJ. Computerized tomography in the evaluation of lumbar disc herniation. *Conn Med* **1981**;45:349–352
36. Dublin AB, McGahan JP, Reid MH. The value of computed tomographic metrizamide myelography in the neuroradiological evaluation of the spine. *Radiology* **1983**;146:79–86
37. Eldevik OP, Dugstad G, Orrison WW, et al. The effect of clinical bias on the interpretation of myelography and spinal computed tomography. *Radiology* **1982**;145:85–89
38. Herkowitz HN, Wiesel SW, Booth RE, et al. Metrizamide myelography and epidural venography. Their role in the diagnosis of lumbar disc herniation and spinal stenosis. *Spine* **1982**;7:55–64
39. Nelson MJ, Gold LH. CT evaluation of intervertebral foramina lesions with normal or non diagnostic myelograms. Report of ten cases. *Comput Radiol* **1983**;7:155–160
40. Raskin SP, Keating JW. Recognition of lumbar disk disease: comparison of myelography and computed tomography. *AJNR* **1982**;3:235–241, *AJR* **1982**;139:349–355
41. Stoeter P, Schneider I, Bergleiter R, Ebeling U. Diagnostischer Wert der computertomographischer Untersuchung der Lumbosakralregion bei Patienten mit Lumboischialgien. *ROFO* **1982**;136:515–524
42. Tchang SPK, Howie JL, Kirkaldy-Willis WHK, et al. Computed tomography versus myelography in diagnosis of lumbar disc herniation. *J Can Assoc Radiol* **1982**;33:15–20
43. Griebel R, Tchang S, Khan M, et al. Correlation of computed tomography with surgical diagnosis in lumbar disc disease. *Can J Neurol Sci* **1983**;10:248–251
44. Pythinen J, Lahde S, Tamska EL, et al. Computed tomography after lumbar myelography in lower back and extremity pain syndromes. *Diagn Imaging* **1983**;52:19–22
45. Greenough CG, Dimmock S, Edwards D, Ransford AO, Bentley G. The role of computerised tomography in intervertebral disc prolapse. *J Bone Joint Surg [Br]* **1986**;68:729–733
46. Kratzer M, Hipp E. Stellenwert der Myelographie und Computertomographie bei Problemfällen in der lumbalen Bandscheibendiagnostik. *Z Orthop* **1986**;124:107–111
47. Schoedinger GR III. Correlation of standard diagnostic studies with surgically proven lumbar disk rupture. *South Med J* **1987**;80:44–46
48. Voelker JL, Mealey J Jr, Eskridge JM, Gilmor RL. Metrizamide-enhanced computed tomography as an adjunct to metrizamide myelography in the evaluation of lumbar disc herniation and spondylosis. *Neurosurgery* **1987**;20:379–384
49. Metz CE. Basic principles of ROC analysis. *Semin Nucl Med* **1978**;8: 283–298
50. Hampton AO, Robinson JM. Roentgenographic demonstration of rupture of the intravertebral disk into the spinal canal after injection of Lipiodol. *AJR* **1936**;36:782–803