AJNR

Effects of i.v. contrast administration on intraspinal and paraspinal tissues: a CT study. 2. Visual assessment.

J T Wilmink, J G Roukema and W van den Burg

AJNR Am J Neuroradiol 1988, 9 (1) 191-193
http://www.ajnr.org/content/9/1/191

Effects of IV Contrast Administration on Intraspinal and Paraspinal Tissues: A CT Study
2. Visual Assessment

Jan T. Wilmink[1]
Jaap G. Roukema[1]
Willem van den Burg[2]

Paired and matched CT scans, one before and one after IV injection of contrast material, from 30 spinal CT studies were presented in a blind fashion to six observers with varying degrees of radiologic experience. The observers were asked to indicate which section of each pair gave the best delineation of intraspinal soft tissues. Seventy-five percent of the verdicts indicated the postcontrast image was superior in this respect, but the six observers were unanimous on this in only 50% of the cases. At a follow-up viewing of the same cases several weeks later, some 20% of the verdicts were changed. Experience in interpreting spinal CT scans did not prove to influence rating performance or improve intraobserver consistency.

In summary, contrast material, in the dosage and injection mode used by us, appeared to be the most important factor in improving scan interpretation.

IV injection of a contrast medium produces an enhancement effect in intraspinal and paraspinal tissues. A previous study by us [1] showed a significant postcontrast increase in CT attenuation numbers in intraspinal and paraspinal scar tissue in a group of 30 individuals. Such measured mean increases have previously been mentioned [2–4]. In our study, however, individual variations in measured enhancement were so great that this measurement technique does not appear to be practically useful for identifying intraspinal scar tissue under clinical conditions. This finding is supported by others [5].

Another way in which the enhancement effect could be of use in postoperative spinal CT diagnosis is by the better delineation it affords of intraspinal structures. In a spinal CT scan of an unoperated individual, the dural sac, the extradural nerve roots, and sometimes also the epidural veins are clearly visible and sharply outlined against the surrounding epidural fat, which appears black with the usual window settings. After spinal surgery the epidural fat is often no longer visible, and structures such as dural sac and nerve roots are often indistinguishable from the now-isodense surrounding postoperative tissues.

Contrast enhancement usually makes it possible to distinguish between the CSF content of the dural sac, which shows no reaction, and the dura together with the surrounding tissues, which become relatively hyperdense (instead of being strongly hypodense like the preexisting epidural fat).

Visual inspection criteria for differentiating between postoperative epidural scarring and recurrent disk herniation rely on features such as contours with respect to the dural sac and the disk surface. Contrast enhancement, by improving the visibility of such features, can be helpful in the diagnosis of the postoperative spine [2, 6] and has been reported to increase the diagnostic accuracy and level of confidence in making the differentiation between recurrent herniated disk and scar [5]. Our present study was constructed to ascertain how often such an improvement in detail visibility could be discerned and how consistent different observers are in this type of assessment. With respect to the latter, we attempted specifically to ascertain first whether verdicts are more reliable when observers are experienced

Received January 30, 1987; accepted after revision June 25, 1987.

[1]Department of Diagnostic Radiology, Section of Neuroradiology, University Hospital, 59 Oostersingel, 9713 EZ Groningen, The Netherlands. Address reprint requests to J. T. Wilmink.

[2]Department of Neuropsychology, University Hospital, 9713 EZ Groningen, The Netherlands.

***AJNR* 9:191–193, January/February 1988**
0195-6108/88/0901-0191

in interpreting spinal CT scans and second whether the judgments contain a systematic idiosyncratic or observer-related component.

Materials and Methods

CT studies of 30 patients, performed before as well as after IV injection of 100 ml Telebrix 30 (30 g I), were scrutinized. The decision to inject contrast medium had been based on plain scan findings suggestive of epidural scarring. Of the 30 patients, 27 had undergone spinal surgery and two had chemonucleolysis. One individual had neither of these two procedures. All postcontrast scans were begun within 1–2 min of injection, which was delivered by hand as rapidly as possible through a 19-gauge needle. A previous study that used the scanning data from the same patients showed a peak in attenuation values within the blood vessels after IV injection of contrast material, followed by a mean decrease in enhancement over the next few minutes. For the extravascular compartment, however, no consistent decrease could be measured after an initial rise of attenuation values with IV injection of contrast material. This extravascular compartment, probably more representative of postoperative intraspinal tissues, could thus for practical purposes be regarded as remaining in a "steady state" of enhancement for some minutes after injection.

Pairs of pre- and postcontrast 4.5-mm-thick CT sections obtained with a Philips Tomoscan 310 machine were presented to a group of six observers. To ascertain the influence of experience, the group was composed of two neuroradiologists who interpret spinal CT scans daily (observers 1 and 2), two radiologists without special experience in this field (observers 3 and 4), and two clinicians with no formal training in radiology (observers 5 and 6) (Table 1). These observers were asked to state which of the two unmarked, paired CT sections gave the best delineation of intraspinal structures such as dural sac, nerve roots, and disk contour. The CT images had been trimmed to the anterior margin of the spine to prevent recognition of the postcontrast study by comparison of the iliac vessels. The matching of slice pairs was performed on the basis of bony landmarks to eliminate the possibility of motion between the pre- and postcontrast studies. A good match was achieved in all cases. One hundred eighty verdicts (section A more detail, section B more detail, or no distinction possible) were delivered on the 30 cases.

TABLE 1: Visual Observer Assessments of Enhanced and Unenhanced Spinal Tissue

Assessment: Observer No.	No. of Assessments (%)		
	More Detail After Contrast	No Distinction	More Detail Before Contrast
First:			
1	20	9	1
2	22	8	0
3	23	5	2
4	23	7	0
5	26	4	0
6	21	8	1
Total	135 (75)	41 (23)	4 (2)
Second:			
1	18	12	0
2	15	14	1
3	18	10	1
4	25	5	0
5	26	3	1
6	25	4	1
Total	127 (71)	48 (27)	4 (2)

Intra-observer variability was assessed by having the observers repeat their scrutiny and assessments several weeks later.

Results

Table 1 contains the basic results. In 135 (75%) of 180 verdicts the postcontrast study was identified as more clearly delineating intraspinal structures, as expected. In 41 instances (23%), however, no distinction could be made, while in four instances (2%) the precontrast study was considered superior. These four judgments were made in four different patients, and by three separate observers, and were delivered in cases where the measurements in fact indicated considerable enhancement. As Table 1 also shows, the performance of experienced observers did not differ from that of those less experienced. Table 2 illustrates how cases can be grouped according to verdicts. Thus, the top row refers to 15 cases in which verdicts were unanimous, the second to five cases in which there was near unanimity (five observers against one), and so forth. When these visual assessments were compared with the CT attenuation measurements performed in the same group of patients in a previous study, it can be seen that there is a degree of correlation between the two parameters (Table 3). Individual variations are very marked, however, as shown by the spread of attenuation coefficients. By using the "H coefficient" [7, 8] as an index of agreement, the judges proved to be reasonably consistent in their findings (Table 4). For the present study the H coefficient was reduced to the computational form rendered in Table 4. Consistency

TABLE 2: Cases Grouped According to Balance of Verdicts (First Assessment)

No. of Cases	More Detail After Contrast	No Distinction	More Detail Before Contrast
15	6	0	0
4	5	1	0
2	4	2	0
1	4	1	1
1	3	3	0
1	3	2	1
2	2	3	1
1	2	4	0
2	1	5	0
1	0	6	0

TABLE 3: Visual Assessment Compared with Attenuation Measurements

Visual Score (First Assessment)	Mean Attenuation Increase in H (Spread)
Observers unanimous: more detail postcontrast	36 (2[a]–60)
Majority observers: more detail postcontrast	25 (16–37)
Majority observers: no distinction	15 (6–26)
Observers unanimous: no distinction (one case)	4
Four verdicts: more detail precontrast	39 (9–54)

[a] One case showed only rim enhancement and was considered probable recurrent herniation.

was no greater in the most experienced subgroup of neuroradiologists, with an H coefficient of 0.70, than among the two general radiologists (0.75) or the nonradiologists (0.82). The six observers were unanimous in their verdicts in 16 of the 30 scan pairs (once with regard to "no distinction").

The follow-up assessment by the same group several weeks later is shown in Table 1. One study was now judged to be technically inadequate by one observer. Although performance was consistent for the whole observer group, there were considerable individual inconsistencies (Table 5).

To assess whether there was an idiosyncratic component related to individual observers in the consistency of judgments, we compared the intraobserver consistency with the interobserver consistency. For instance, the degree to which observer 1 was consistent with himself could be compared with the inconsistencies between the first verdicts of observer 4 and the second verdicts of observer 5 (matched-pairs sign test). Of the 120 comparisons possible, nine appeared to differ significantly at $p < .05$, which is only slightly more than can be expected by chance. Hence, the intraobserver consistency was not clearly larger than the interobserver consistency, and there thus appeared to be no significant idiosyncratic component in consistency.

Four pairs of scans that seemed to produce especially inconsistent ratings were scrutinized retrospectively. In one, the images were exceptionally noisy; in another the amount of scar tissue was small. In the other two scan pairs no explanation could be found for the inconsistencies.

TABLE 4: H Coefficient of Agreement

Observer 1	Observer 2		
	A > B	A = B	B > A
A > B	a	b	c
A = B	d	e	f
B > A	g	h	i

$$H = [(a + i) - (c + g)]/\sqrt{(a + b + c + g + h + i)(a + d + g + c + f + i)}$$

In this table, f symbolizes the number of times observer 1 finds sections A and B equally good and observer 2 finds B better than A. The other eight symbols are defined analogously.

TABLE 5: Intraobserver Discrepancy

Observer No.	First-order Discrepancy	Second-order Discrepancy	Total (%)[a]
1	7	0	7 (23)
2	8	1	9 (30)
3	7	0	7 (23)
4	4	0	4 (13)
5	3	0	3 (10)
6	6	1	7 (23)
Total	35	2	37 (21)

Note.—First-order discrepancy: verdict shifted from no distinction to more detail precontrast or more detail postcontrast or vice versa. Second-order discrepancy: verdict shifted from more detail postcontrast to more detail precontrast or vice versa.

[a] Percentages are derived from the number of discrepancies in the 30 assessments of each individual observer and the 180 assessments of the whole group of observers.

Discussion

The problem of observer variation in the interpretation of diagnostic data has been the subject of a great number of studies, and an extensive review was provided in two articles by Koran [9]. Comparison of observer agreement rates among different studies is difficult, but Garland [10], in a study design similar to ours, which required observers to indicate changes in serial chest films, found inter- and intraindividual discrepancy rates similar to those in our group. In addition, there proved to be no difference in performance between two categories of observers (radiologists and phtisiologists). Table 1 shows considerable intraobserver variability between the two assessments, but much more consistent results for the group as a whole.

The relative lack of correlation between measured attenuation increase and visual assessment indicates that better detail visibility depends not only on the differential attenuation increases of various intraspinal structures, but also on distribution factors. Edge enhancement, for instance, is much more easily perceived than measured with the aid of a light pen. It seems likely, however, that a high dose of contrast material, as has been advocated [6], will increase the discernibility of the enhancement effect.

In conclusion, IV administration of a contrast medium in the dosage and injection mode described by us appears to improve the visualization of intraspinal structures in about three-fourths of observations. In one-half of the cases the six observers were unanimous in this respect.

ACKNOWLEDGMENTS

We thank our colleagues who participated in the observer study, O. F. Brouwer, J. P. Klein, L. Penning, M. J. Staal, P. P. Taverne, and L. M. Vencken, and G. Weltevrede for help in manuscript preparation.

REFERENCES

1. Wilmink JT, Roukema JG. Effects of IV contrast administration on intraspinal and paraspinal tissues: a CT study. 1. Measurement of CT attenuation numbers. *AJNR* **1987**;8:703–709
2. Teplick G, Haskin ME. Computed tomography of the postoperative lumbar spine. *AJNR* **1983**;4:1053–1072
3. Mall JC, Kaiser JA, Heithoff KB. Postoperative spine. In: Newton TH, Potts DG, eds. *Computed tomography of the spine and spinal cord*. San Anselmo, CA: Clavadel, **1983**:187–204
4. Weiss T, Treisch J, Kazner E, Köhler D, Collmann H, Claussen C. CT of the postoperative lumbar spine: the value of intravenous contrast. *Neuroradiology* **1986**;28:241–245
5. Braun IF, Hoffmann JC, Davis PC, Landman JA, Tindal GT. Contrast enhancement in CT differentiation between recurrent disk herniation and postoperative scar. *AJNR* **1985**;6:607–612
6. Yang PJ, Seeger JF, Dzioba RB, et al. High-dose IV contrast in CT scanning of the postoperative lumbar spine. *AJNR* **1986**;7:703–707
7. Sjöberg L, Holley JW. A measure of similarity between individuals when scoring directions of variables are arbitrary. *Multivariate Behavioral Research* **1976**;2:377–384
8. Vegelius J. A G index generalization for trichotomized data. *Educ Psych Measurement* **1979**;39:23–27
9. Koran LM. The reliability of clinical methods data and judgements. *N Engl J Med* **1975**;293:642–646 (part 1), 695–701 (part 2)
10. Garland LH. The problem of observer error. *Bull NY Acad Med* **1960**:36:570–584