



## Discover Generics

Cost-Effective CT & MRI Contrast Agents



WATCH VIDEO

# AJNR

## Validation of Multisociety Combined Task Force Definitions of Abnormal Disk Morphology

C.H. Cho, L. Hsu, M.L. Ferrone, D.A. Leonard, M.B. Harris, A.A. Zamani and C.M. Bono

This information is current as of June 5, 2025.

*AJNR Am J Neuroradiol* 2015, 36 (5) 1008-1013

doi: <https://doi.org/10.3174/ajnr.A4212>

<http://www.ajnr.org/content/36/5/1008>

# Validation of Multisociety Combined Task Force Definitions of Abnormal Disk Morphology

C.H. Cho, L. Hsu, M.L. Ferrone, D.A. Leonard, M.B. Harris, A.A. Zamani, and C.M. Bono



## ABSTRACT

**BACKGROUND AND PURPOSE:** The multisociety task force descriptively defined abnormal lumbar disk morphology. We aimed to use their definitions to provide a higher level of evidence for the validation of MR imaging in the evaluation of this pathology in patients who have undergone diskectomy by retrospectively classifying their preoperative MRI.

**MATERIALS AND METHODS:** This retrospective, institutional review board–approved study included 54 of 86 consecutive patients (47 men; average age, 44 years) enrolled in an ongoing prospective trial of surgically treated lumbar disk herniation who had preoperative MRI and documented intraoperative classification of the abnormal disk as protrusion, extrusion, or sequestration by the treating surgeon. Preoperative MRI was classified by 2 blinded radiologists; discrepancies were resolved by a third reader. Statistical analysis of interobserver agreement and imaging compared with surgical findings was performed.

**RESULTS:** The readers disagreed on only 1 of the 54 cases. The third reader resolved the disagreement. Eight protrusions and 46 extrusions were found on imaging, with no sequestrations. At surgery, there were 13 protrusions and 40 extrusions, with 2 of the extrusions also containing sequestrations; the remaining case had only sequestration. There were 16 discrepancies between imaging and surgery, resulting in 70% agreement.

**CONCLUSIONS:** This study, which was intended to validate the multisociety combined task force definitions of abnormal disk morphology by using MR imaging with a surgical criterion standard, found 70% agreement between imaging diagnosis and surgical findings. Although reasonable, this finding highlights differences that often exist between intraoperative and preoperative imaging findings of lumbar disk herniation.

MR imaging of the lumbar spine with defined specific MR images has gained acceptance as the standard of care for the evaluation of degenerative disk disease.<sup>1</sup> However, the interpretation of these images continues to have much variability. An effort to standardize image reporting brought together multiple national medical societies including the North American Spine Society, American Society of Spine Radiology, and American Society of Neuroradiology. They produced and then updated a consensus document for image descriptions.<sup>2</sup>

One segment of the consensus document focuses on the mor-

phology of the lumbar disk as it relates to the location of abnormal disk content with respect to the outer annulus.<sup>2</sup> MR imaging studies evaluating the interobserver and intraobserver reliability of disk morphology by using the definition of the consensus document have been performed.<sup>3–5</sup> However, these studies lacked analysis of surgical findings for correlation with a criterion standard and, as such, do not provide the highest level of diagnostic evidence.

This study retrospectively classified MR imaging findings in a cohort of consecutive patients with surgically treated disk herniation by using the descriptive definition of abnormal lumbar disk morphology of the multisociety task force, with the aim of providing a high level of evidence for validation of MR imaging by using the combined task force definitions.<sup>2</sup>

## MATERIALS AND METHODS


A retrospective review was performed, with approval from the institutional review board of the study site, of the records of 86 consecutive patients (47 men, 39 women; average age, 44 years) who were enrolled between August 2009 and October 2013 in an ongoing prospective clinical trial evaluating the outcomes of sin-

Received September 8, 2014; accepted after revision October 24.

From the Departments of Radiology (C.H.C., L.H., A.A.Z.) and Orthopedic Surgery (M.L.F., D.A.L., M.B.H., C.M.B.), Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts.

Abstract of preliminary work previously presented at: Annual Meeting of the North American Spine Society, November 12–15, 2014; San Francisco, California.

Please address correspondence to: Charles H. Cho, MD, MBA, Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, 75 Francis St, Boston, MA 02115; e-mail: ccho81@partners.org

 Indicates article with supplemental on-line appendix.

<http://dx.doi.org/10.3174/ajnr.A4212>

gle-level lumbar disectomy. Inclusion criteria for entry into the prospective study were the presence of a symptomatic single-level lumbar disk herniation, failure of nonoperative treatment, primary radicular pain, and no prior lumbar surgery. Included in prospective data collection was documentation of the intraoperative classification of the herniation as a protrusion, extrusion, or sequestration. For the current study, the intraoperative findings were used as the diagnostic criterion standard because the surgeon is looking directly at the anatomy of the patient. These findings have been used as such in numerous other studies.<sup>6,7</sup> With continued improvement in imaging, the depiction of a patient's anatomy is becoming clearer and more accurate and is approaching equivalence to looking at it.

Surgical cases were from the clinical practice of 3 attending orthopedic spine surgeons from a single institution. Surgeons recorded operative findings that included classifying the disk herniation as a protrusion, extrusion, or sequestration intraoperatively. In the operative field, the outer margin of the annulus was visually identified. A defect of the annulus with disk material outside the annulus was classified as "extrusion." If there were no defect and no disk material outside the annulus, the classification of "protrusion" was made. "Sequestration" was determined if there was disk material separate from the annular defect with no visible attachment to the parent disk.

Of the 86 patients, 1 patient had no disk herniation noted during surgery, which was presumed to mean the disk had resorbed between imaging and surgery, and 3 patients had incomplete surgical data sheets. These cases were excluded. Of the remaining 82 cases, 54 had presurgical MR images accessible by the PACS of the study institution, so these 54 cases were included in the present study.

In preparation for the current analysis, a summary sheet (Online Appendix 1) was created and reviewed by 3 board-certified neuroradiologists, each with a minimum of 15 years of practice experience and a combined average of >26 years, based on the previously described multisociety task force definitions.<sup>2</sup> Sample cases from daily practice were used in a face-to-face setting to confirm the understanding of the summary sheet definitions before data collection.

The MR imaging of each subject (mean,  $81 \pm 75$  days before surgery) was reviewed by 2 board-certified radiologists who were blinded to the patient medical history and operative findings. In the evaluation, the disk level with the most substantial abnormality was classified as a protrusion, extrusion, or sequestration by using both axial and sagittal images. Classifications were confined to a single disk. In cases in which >1 disk level had abnormalities, only the more severe level was selected. In 1 case in which there was nearly equivalent severity at 2 disk levels, the surgical level was provided to the readers to create a final, imaging-based classification that could be compared with the surgical classification. Any discrepancy between the 2 readers was resolved by a third blinded reader, and a majority consensus evaluation became the final imaging interpretation. Of note, many of the MR images were from different, outside, referring hospitals. However, all cases included sagittal T1- and sagittal and axial T2-weighted images.

Most important, while the surgeons did have the MR imaging available at the time of surgery, they did not have the findings as

decided by the study radiologists. Additionally, the surgeons were instructed to make their classifications solely on the basis of their intraoperative observations. While having the MR imaging available to surgeons may have been a potential source of bias, surgeons always use MR imaging to plan for their surgery; thus, in practice, surgeons also have that potential source of bias. Research must be applicable to practice, and blinding surgeons to MR imaging would not mimic practice.

All statistical analyses were performed by using STATA/SE 13.1 (StataCorp, College Station, Texas). Agreement between imaging and surgical classifications was calculated by using the Cohen  $\kappa$ , as was agreement between imaging assessments. The weighted  $\kappa$  was not used because disk herniations do not always progress in a standardized manner. A standard  $2 \times 2$  table was used to determine sensitivity, specificity, and predictive values.

## RESULTS

### *Imaging Findings*

Of the 54 cases, the 2 readers disagreed in 3. One case had a final diagnosis discrepancy between protrusion and extrusion based on all images. In the other 2 cases, there was a disagreement on disk description as protrusion or extrusion on the axial plane, but both readers described an extrusion on the sagittal images. Thus, the final diagnosis of extrusion was the same for these 2 cases. All discrepancies were resolved by the third reader. Overall, 53 of 54 cases showed agreement in the final diagnosis between readers. Interobserver agreement between the 2 readers was 98% with a  $\kappa$  of 0.93. There were 8 protrusions, 46 extrusions, and no sequestrations based on imaging findings.

### *Surgical Findings*

Of the 54 surgical cases with MR images, there were 13 protrusions and 40 extrusions, with 2 of the extrusions each occurring in addition to sequestrations. The 1 remaining case had only sequestration and no extrusion.

### *Imaging and Surgical Agreement*

Four cases with an imaging diagnosis of protrusion were found to be extrusions in surgery. Nine cases with imaging diagnoses of extrusion were found to be protrusions in surgery. Whereas imaging data did not reveal any cases of sequestration, in 2 cases with imaging diagnoses of extrusion, surgery showed both extrusion and sequestration; these were considered discrepancies between imaging and surgical findings. In another case, imaging showed extrusion while surgery showed sequestration only. The total number of discrepancies was 16, calculated as an overall agreement of 70% ( $\kappa = 0.19$ ). Table 1 lists all the discrepancy cases and gives descriptions of the imaging findings.

Sensitivity and specificity of detection of a ruptured annulus were 0.90 and 0.31, respectively (Table 2). The positive predictive value was 0.80, while the negative predictive value was 0.50.

## DISCUSSION

Prior studies evaluating imaging interpretation of lumbar disk abnormalities have shown moderate interobserver agreement and substantial intraobserver agreement in cohorts of asymptomatic volunteers and symptomatic patients.<sup>3</sup> One study included expe-

**Table 1: Discrepancy cases between imaging and surgery, with a description of the imaging findings**

No.	Disk Level	Imaging	Surgical	Description
1	4/5	Extrusion	Protrusion	On sagittal images, the height of the abnormal disk is slightly greater than the disk height
2	2/3	Extrusion	Extrusion and sequestration	Although imaging suggests a sequestered fragment, there is a thin continuity between the disk fragments visible on MR imaging (Fig 1)
3	4/5	Extrusion	Protrusion	On sagittal images, the height of the abnormal disk is slightly greater and very close in distance to the maximum disk height
4	4/5	Extrusion	Protrusion	On sagittal images, the height of the abnormal disk is very close in distance to the maximum disk height (Fig 2)
5	4/5	Extrusion	Extrusion and sequestration	Sagittal images show a connection between disk fragments
6	4/5	Extrusion	Protrusion	On sagittal images, the abnormal disk appears contained by the outer annulus, but the height of the annulus is slightly greater than the disk height
7	4/5	Extrusion	Protrusion	On sagittal images, the height of the abnormal disk is slightly greater than the disk height
8	5/1	Extrusion	Protrusion	On sagittal images, the height of the abnormal disk is slightly greater than the disk height
9	5/1	Protrusion	Extrusion	On the sagittal view, the low signal of the disk and annulus is very difficult to differentiate (Fig 3)
10	4/5	Extrusion	Protrusion	On sagittal images, the height of the abnormal disk is very similar to the disk height
11	4/5	Protrusion	Extrusion	On the sagittal view, the low signal of the disk and annulus is very difficult to differentiate
12	4/5	Protrusion	Extrusion	On sagittal images, the entire annulus is not clearly defined
13	5/1	Protrusion	Extrusion	On sagittal images, the abnormal disk height is similar to the maximum disk height
14	4/5	Extrusion	Protrusion	On sagittal images, the abnormal disk height is slightly greater than the maximum disk height
15	5/1	Extrusion	Protrusion	On the sagittal view, abnormal disk height is much greater than the normal disk height and the outer annulus appears disrupted; in retrospect, the diagnosis is still extrusion by imaging criteria (Fig 4)
16	4/5	Extrusion	Sequestration	Imaging shows a thin connection between the parent disk and disk the fragment

**Table 2: 2 × 2 contingency table for a ruptured-versus-intact annulus as determined surgically and on imaging**

Imaging Assessment	Surgical Assessment (Criterion Standard)		Totals
	Annulus Ruptured <sup>a</sup>	Intact Annulus <sup>b</sup>	
Annulus ruptured <sup>a</sup>	37	9	41
Intact annulus <sup>b</sup>	4	4	13
Totals	46	8	54

<sup>a</sup> Extrusions and sequestrations.<sup>b</sup> Protrusions only.

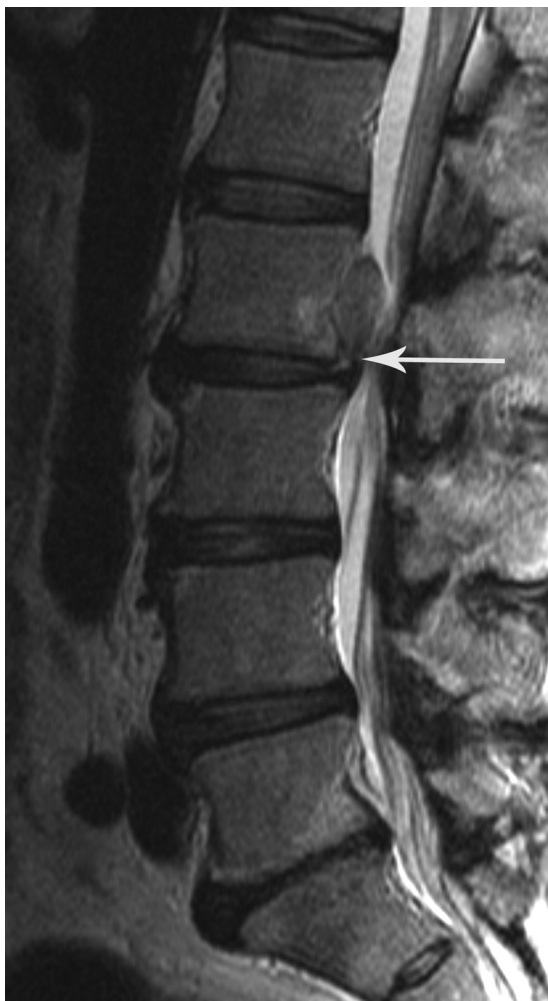
rienced readers in a single academic institution; interpretation included descriptions of normal and bulge as well as protrusion and extrusion, with the most common discrepancy occurring between normal and bulge.<sup>3</sup> Another study, by using readers from 1 academic center with a prestudy effort to define the review criteria, found near-perfect interreader agreement (summary  $\kappa = 0.81$ ) for normal/bulge, protrusion, and extrusion/sequestration categories.<sup>4</sup> Other reports replicating conditions close to those of clinical practice by using radiologists from different hospitals with no prior set diagnostic criteria or training showed only moderate interobserver agreement for herniation and fair-to-moderate agreement for disk contour.<sup>5</sup> Most important, all of the above studies lacked a surgical criterion standard.

In the current study, substantial effort was made to decrease interobserver variability, because the primary goal was to evaluate the correlation with surgical findings. Before data collection, the 3 readers discussed the nomenclature by using daily clinical cases from the previous few months. There was extensive discussion about the definitions of protrusion and extrusion, especially on the sagittal images. In the end, the critical point of interpretation among the readers was the exact location of the outer annulus insertion on the vertebral body, which was crucial in discriminat-

ing protrusion from extrusion. To better understand this anatomy, the readers reviewed studies in the literature, including pathologic and traumatic conditions, in terms of how they affect the outer annulus insertion site and how it is localized above the bony endplate on the sagittal view.<sup>8-11</sup> These studies showed that the insertion of the annulus is beyond the bone edge of the disk; this finding creates discrepancies among readers on the insertion site. With this understanding, we went back to the protrusion definition from the combined task force report<sup>2</sup> and noted that the insertion site of the annulus cannot be greater than the maximum height of the disk space in the sagittal plane. For uniformity, we therefore decided to use the disk space height to define the difference between protrusion and extrusion. Thus, any disk morphology less than the maximum sagittal disk height was designated as protrusion. This distinction is depicted in On-line Appendix 2.

After this effort to optimize uniform understanding and agreement of the definitions among the radiologists, there was consensus in the final imaging diagnosis of 53 of 54 cases (agreement, 98%;  $\kappa = 0.93$ ). If the 2 other cases of discrepancy in the axial or sagittal description (which did not affect the final diagnosis) are included, the agreement was 51 of 54, or 94%. These values are at least as good as, if not better than, those previously reported.

Initially the surgical level was not provided to the readers because transitional vertebral levels may have directed the attention to an incorrect surgical level by causing the readers to number the vertebrae differently. Thus, the readers were instructed to report the most severe disk level during the imaging analysis. In 1 case, L5-S1 had moderate central protrusion and L4-L5 had a small foraminal protrusion. The latter, despite appearing less severe, was in fact at the symptomatic level of radiculopathy. In this case, the correct level was indicated to the readers at the time of imaging data collection and before the analysis.



**FIG 1.** Case 2. Sagittal T2-weighted image shows a thin connection (arrow) between the level of abnormal disk and the disk fragment. With imaging, this is classified as extrusion. However, it was classified intraoperatively as extrusion and sequestration.

Overall, our accuracy measurements indicate that imaging is better at determining when the annulus is ruptured than it is at determining that the annulus is intact. Analysis of the discrepancies between imaging and surgery (Table 1) showed the following 3 observations: 1) When the height of the abnormal disk material on the sagittal plane is very close to the maximum height of the normal disk, it becomes difficult to differentiate protrusion from extrusion on imaging (Fig 2). 2) When the disk signal is low on all MR images, differentiation of an abnormal disk from the outer annulus becomes challenging and a small extruded disk cannot be differentiated from a contained (protruded) disk (Fig 3). 3) In differentiating extrusion from sequestration, imaging clearly shows a thin connecting tissue, which, by definition, is considered extrusion (Fig 1). The thin sliver of tissue that defines a herniation as an extrusion can be clearly shown on imaging yet not observed in surgery, because the thin sliver of tissue may be very difficult to discern. All discrepancy cases are shown in On-line Appendix 3.

We believe that the effect of variation in imaging parameters on discrepant classifications was minimal in this study. Of the 16 cases with discrepancies, though the magnet strength is not noted



**FIG 2.** Case 4. Sagittal T2-weighted image shows the height of the abnormal disk herniation (bold arrows) compared with the disk height (thin arrows). By imaging, this is classified as extrusion. However, it was classified intraoperatively as protrusion.

in all, most appear to have been performed with at least 1.5T. The TR ranged from 3000 to 5760 ms (a single case at 1000 ms); TE, from 90 to 148 ms; section thickness, between 4 and 5 mm (a single case at 3.6 mm); the FOV, from 24 to 30 cm; and the acquisition matrix, between 256 and 512 (a single case,  $240 \times 175$ ). The overall range of imaging parameters of all these cases is within the standard for achieving high-quality images, which are also in agreement visually as determined by 2 experienced neuroradiologists. With most of the imaging findings being obvious, ultimately we neither think that these discrepancies are due to slight differences in imaging parameters nor believe that variation in imaging parameters compromises the purpose of this study.

Given that at surgery a defect in the posterior annulus differentiates extrusion from protrusion while in imaging it is the relative size of the herniated disk versus the height of base of the annulus that differentiates the 2 categories, it is somewhat surprising that surgery and radiology agree so often. While the approach is different on imaging versus during surgery, the goal in both cases is to determine whether the annulus is ruptured. Because this can be directly observed intraoperatively, surgical standing has been used as the criterion standard, but advancements in imaging can perhaps better define the outer annulus and an-





**FIG 3.** Case 9. Sagittal T2-weighted image shows decreased signal of the disk material, which makes differentiation of the disk from the annular margin difficult.

nular defects. This change would allow a truer imaging prediction of surgical findings. Imaging needs to be better defined so that it can more accurately and more directly determine the presence or lack of an annular defect instead of relying on morphology to provide clues as to the presence or lack of said defect. As imaging improves and this definition becomes possible, it could enable better correlation of symptoms of a disk herniation to the type of herniation and prediction of surgical-versus-nonoperative outcomes for any given patient on the basis of their MR imaging.

This study has a number of limitations. The time between imaging and surgery was 81 days on the average, so progression or resorption of the disk herniation could have occurred from the time of imaging to surgery. One example of this is case 15 (Fig 4), which, on imaging, was classified as an extrusion but was intraoperatively classified as a protrusion. In addition, the number of cases of sequestration was small, thereby limiting conclusions from analysis of this category. In addition, despite the reasonable (70%) agreement between the imaging diagnosis and surgical findings, the probability of the calculated agreement occurring by chance was high owing to the distribution of classifications, which were mostly extrusions. This is reflected in the low  $\kappa$  of 0.19. Despite this high probability, we think that the study is representative of actual practice, and increasing the sample size would not have significantly altered the pathology distribution. Additionally, agreement on a surgical classification is generally moderate at best.<sup>12</sup> For disk classification, the agreement on the difference between protrusion and extrusion/sequestration would likely be good, but the agreement on the difference between extrusion and sequestration would likely be moderate at best. However, due to the pathology distribution and our resulting focus on the distinction between protrusion and sequestration/extrusion rather than



**FIG 4.** Case 15. Sagittal T2-weighted image shows disruption of outer annulus (arrow) even on retrospective interpretation; thus, the imaging classification in this case is extrusion. However, it was classified intraoperatively as a protrusion.

the distinction between sequestration and extrusion, the effect of this limitation is unlikely to be substantial.

## CONCLUSIONS

In this study, which was intended to validate the multisociety combined task force definitions of abnormal disk morphology by using MR imaging with a surgical criterion standard, there was 70% agreement between the imaging diagnosis and surgical findings. Common trends for the discrepancy are described. Future effort may yield better agreement between surgeons and radiologists as to how they describe disk herniation and abnormalities.

Disclosures: Charles H. Cho—UNRELATED: Board Membership: North American Spine Society (no money). Mitchel B. Harris—UNRELATED: Board Membership: North American Spine Society (no money). Christopher M. Bono—UNRELATED: Board Membership: North American Spine Society (nonmonetary, volunteer); Consultancy: United Healthcare; Expert Testimony: CRICO; Royalties: Wolters Kluwer, for edited books; Other: Journal of the American Academy of Orthopaedic Surgeons, Comments: stipend for being Deputy Editor.

## REFERENCES

- Modic MT, Ross JS. Lumbar degenerative disk disease. *Radiology* 2007;245:43–61
- Fardon DF, Williams AL, Dohring EJ, et al. Lumbar disc nomenclature: version 2.0, recommendations of the combined task forces of the North American Spine Society, the American Society of Spine Radiology, and the American Society of Neuroradiology. *Spine (Phila Pa 1976)* 2014;39:E1448–65
- Brant-Zawadzki MN, Jensen MC, Obuchowski N, et al. Interobserver and intraobserver variability in interpretation of lumbar disc abnormalities: a comparison of two nomenclatures. *Spine (Phila Pa 1976)* 1995;20:1257–63; discussion 1264
- Lurie JD, Tosteson AN, Tosteson TD, et al. Reliability of magnetic

- resonance imaging readings for lumbar disc herniation in the Spine Patient Outcomes Research Trial. *Spine (Phila Pa 1976)* 2008;33:991–98
5. Arana E, Kovacs FM, Royuela A, et al. **Influence of nomenclature in the interpretation of lumbar disk countour on MR imaging: a comparison of the agreement using the combined task force and the Nordic nomenclatures.** *AJNR Am J Neuroradiol* 2011;32: 1143–48
  6. Zhuge W, Ben-Galim P, Hipp JA. **Efficacy of MRI for assessment of spinal trauma: correlation with intra-operative findings.** *J Spinal Disord Tech* 2013 Nov 8. [Epub ahead of print]
  7. Hedberg MC, Drayer BP, Flom RA, et al. **Gradient echo (GRASS) MR imaging in cervical radiculopathy.** *AJR Am J Roentgenol* 1988; 150:683–89
  8. Humzah MD, Soames RW. **Human intervertebral disc: structure and function.** *Anat Rec* 1988;220:337–56
  9. Rauschnig W. **Computed tomography and cryomicrotomy of lumbar spine specimens: a new technique for multiplanar anatomic correlation.** *Spine (Phila Pa 1976)* 1983;8:170–80
  10. Tajima N, Kawano K. **Cryomicrotomy of the lumbar spine.** *Spine (Phila Pa 1976)* 1986;11:376–79
  11. Fredrickson BE, Edwards WT, Rauschnig W, et al. **Vertebral burst fractures: an experimental, morphologic, and radiographic study.** *Spine (Phila Pa 1976)* 1992;17:1012–21
  12. Brumback RJ, Jones AL. **Interobserver agreement in the classification of open fractures of the tibia: the results of a survey of two hundred and forty-five orthopaedic surgeons.** *J Bone Joint Surg Am* 1994;76:1162–66