INTER-OBSERVER AGREEMENT IN MRI ASSESSMENT OF LUMBER INTERVERTEBRAL DISCS AND NERVE ROOTS USING PFIRRMAN CLASSIFICATION

Shahla Mohammed Saeed [a], Abeer Kadum Abass [a],
and Mariwan Mahmood Rasool [b]

Submitted: 15/3/2019; Accepted: 29/7/2019; Published: 21/9/2019

ABSTRACT

Background

Low back pain is a common medical condition and is a major health and economic problem. Intervertebral disc degeneration grading is crucial in the evaluation of many degenerative spine conditions including low back pain.MRI is considered the best imaging technique to evaluate intervertebral disc degenerative changes. Back pain is the most common indication for imaging of the spine. A morphologic Pfirrman grading system is commonly used to classify spinal degenerative changes (intervertebral discs & nerve roots).

Objectives

The aim of this study was to evaluate the degree of independent interobserver agreement using Pfirrman classification in the assessment of disc degenerations and nerve root compromise.

Patients and Methods

This study was conducted in Sulaimani, Kurdistan Region, Iraq at the diagnostic imaging center in Sulaimania during the period of 21st April 2015 to 21st October 2015 for 50 patients with low back pain. The same images were interpreted by two radiologists independently for grades of disc degeneration and nerve root compromise using Pfirrmann classification system, up to the authors' knowledge this is the first study in the Kurdistan Region that depends on interobserver agreement regarding MRI interpretations concerning degenerative lumbar spine changes.

Results

The interobserver agreement in our study for Pfirrman grades of disc degeneration and nerve root compromise was moderate with Kappa values of (K= 0.431) and (K=0.417) respectively.

Conclusion

Pfirrman classification system is reliable and provides a standardized assessment of lumbar intervertebral disc and nerve roots.

Keywords: *Disc degeneration, Lumber MRI, Back pain.*

---

[a] Department of Surgery, College of Medicine, University of Sulaimani, Kurdistan Region, Iraq.
Correspondence: shahla.rasul@univsul.edu.iq
[b] Qaladiza General Teaching Hospital.

https://doi.org/10.17656/jsmc.10210

INTRODUCTION

Low back pain (LBP) is a common medical condition, and is a major health and economic problem in the industrialized world. Back pain is the most common indication for imaging of the spine. Most common causes of back pain are those relating to degenerative changes in the intervertebral discs and facet joints [1]. Low back pain is defined as pain and discomfort, localized below the costal margin and above the inferior gluteal folds, with or without referred leg pain. In these guidelines, chronic low back pain is defined as low back pain persisting for at least 12 weeks, unless specified otherwise [2].

The major components of the intervertebral disc are the nucleus pulposus, the annulus fibrosus and the cartilaginous endplates. In adults, the nucleus pulposus is hyperintense on T2-weighted scans and has distinct boundaries with the annulus fibrosus. The outer ring is made of type 1 collagen and is hypointense on MRI; the inner ring is made of type 2 collagen with more proteoglycan and fluid, and is hyperintense on MRI, indistinguishable from the nucleus pulposus. In youth, the water content of the nucleus pulposus is as high as 80%, although this decreases with age. On MRI the relative signal intensity of the disc varies with water content [3]. In adults a septum-like structure can be seen to transect the middle of the nucleus pulposus in the horizontal plane. On T2-weighted images of the normal lumbar discs there is often a linear area of reduced signal traversing the center of the nucleus from anterior to posterior on sagittal views -the so-called intra nuclear cleft. It is not the case in pediatric age group [3, 4].

During degeneration there are a variety of morphological changes that occur, including loss of both height and hydration of the disc [5]. It had been shown that a drop in water content of nucleus pulposus and matrix turnover are related to the age and grade of degeneration [6]. Magnetic resonance imaging (MRI) is the most important method for clinical assessment of intervertebral disc pathology [7]. The signal characteristics of the disc in T2-weighted MRIs reflect changes caused by aging or degeneration [8, 9, 10].

Pfirrmann et al. (2001) classification system is the dominant MRI classification system used for MRI lumbar spine degenerative changes assessment. The Pfirrmann grading system, as a noninvasive, simple, and convenient MR imaging method, can provide a morphologic and semi quantitative evaluation of intervertebral disc degeneration in vivo [8]. This system has been cited more than twice as many times as any other system. Additionally, only Pfirrmann classification has been cited by NIH (National Institutes of Health) funded studies since 2008 and appears to have been widely accepted for MRI [11].

The aim of the study is to evaluate the degree of independent inter observer agreement using pfirrman classification in assessment of disc degenerations and nerve root compromise.

PATIENTS AND METHODS

This prospective study was conducted in Sulaimani, Kurdistan Region, Iraq in the period between 21st April 2015 to 21st October 2015 at the Diagnostic Imaging Center of Sulaimania Teaching Hospital , up to the authors knowledge this is the first time to conduct such study in our Kurdistan region.

MRI examinations of lumbar spine of 50 patients, 17 males (34%**)** and 33 females(66%) containing a total of 250 intervertebral discs, were included in this prospective study. Patients were referred to MRI unit at the diagnostic imaging center because of chronic low back pain. The age of the patients was ranging from 24-72 years; the mean age was 45.1 years (+/- 13.52 SD).

All images were acquired with 1.5-T (Tesla) MRI system (SIEMENS MAGNETOM Symphony Version) (Syngo MRI 2004A) using (CP Spine Array Coil)**.** All patients were examined in supine position, the imaging protocols included in our study are shown in table1.

Two specialist radiologists interpreted the images, both with a long experience, independently (i.e.) both observers were unaware of each other's findings, this is the first study done in our region that depend on inter observatory agreement regarding MRI interpretations concerning degenerative spine changes.

Both of them depended on Pfirrmann classification regarding intervertebral disc findings and nerve roots compression. Pfirrmann grading for interverbral discs is performed on T2-weighted mid sagittal images, as shown in Table 2 and Figure 1. The Pfirrman grading system used for lumber nerve roots compression assessment consists of four grade categories, which are summarized in table 3 and figures 2, 3, 4 & 5.

The study variables that are used in our study are shown in Table 4, all the variables were recoded separately for all lumbar segments (from L1-L2 to L5-S1).

Those patients that have MRI findings of acute spinal infection, recent trauma, tumors, spinal dysraphism,

and previous spinal surgery were excluded from this study to retain some homogeneity of the study.

The findings of both observers were recorded on Microsoft® Excel 2010 program. Interobserver agreement between the two radiologists was calculated by Cohen's Kappa Coefficient [14], using IBM® SPSS® (Statistical Package for the Social Sciences) Statistics 21.The kappa is a popular index of agreement for categorical rating. It ranges from -1 to +1, an agreement is all the best when the kappa is closer to +1 [17]. Interpretation of kappa was done as proposed by Cohen according to the following:

≤ 0 Less than chance agreement
0.01-0.20 Slight agreement
0.21- 0.40 Fair agreement
0.41-0.60 Moderate agreement
0.61-0.80 Good agreement
0.81-0.99 Almost perfect agreement. [15]

Table 1. Sequences for 50 MR Imaging Examinations.

Pulse Sequence	Repetition Time (msec)	Echo Time (msec)	Field of View (mm)	Section Thickness (mm)
Localizer	24	6	400	10
Sagittal T2-weighted	4000	116	300	4
Sagittal T1-weighted	542	10	300	4
Axial T2-weighted*	4000	116	188	4
Axial T1- weighted*	470	14	188	4
Sagittal myelography	8000	1200	280	5
Coronal myelography	6000	281	270	5

*** Parallel to disc plane.**

Table 2. Pfirrmann Classification of disc degeneration [12].

Grade	Structure	Distinction of nucleus and anulus	Signal intensity	Height of intervertebral disc
I	Homogeneous,bright white	Clear	Hyper intense, isointense to cerebrospinal fluid	Normal
II	Inhomogeneous with or without horizontal bands	Clear	Hyper intense, isointense to CSF	Normal
III	Inhomogeneous, gray	Unclear	Intermediate	Normal to slightly decreased
IV	Inhomogeneous, gray to black	Lost	Intermediate to hypo intense	Normal to moderately decreased
V	Inhomogeneous, black	Lost	Hypo intense	Collapsed disc space

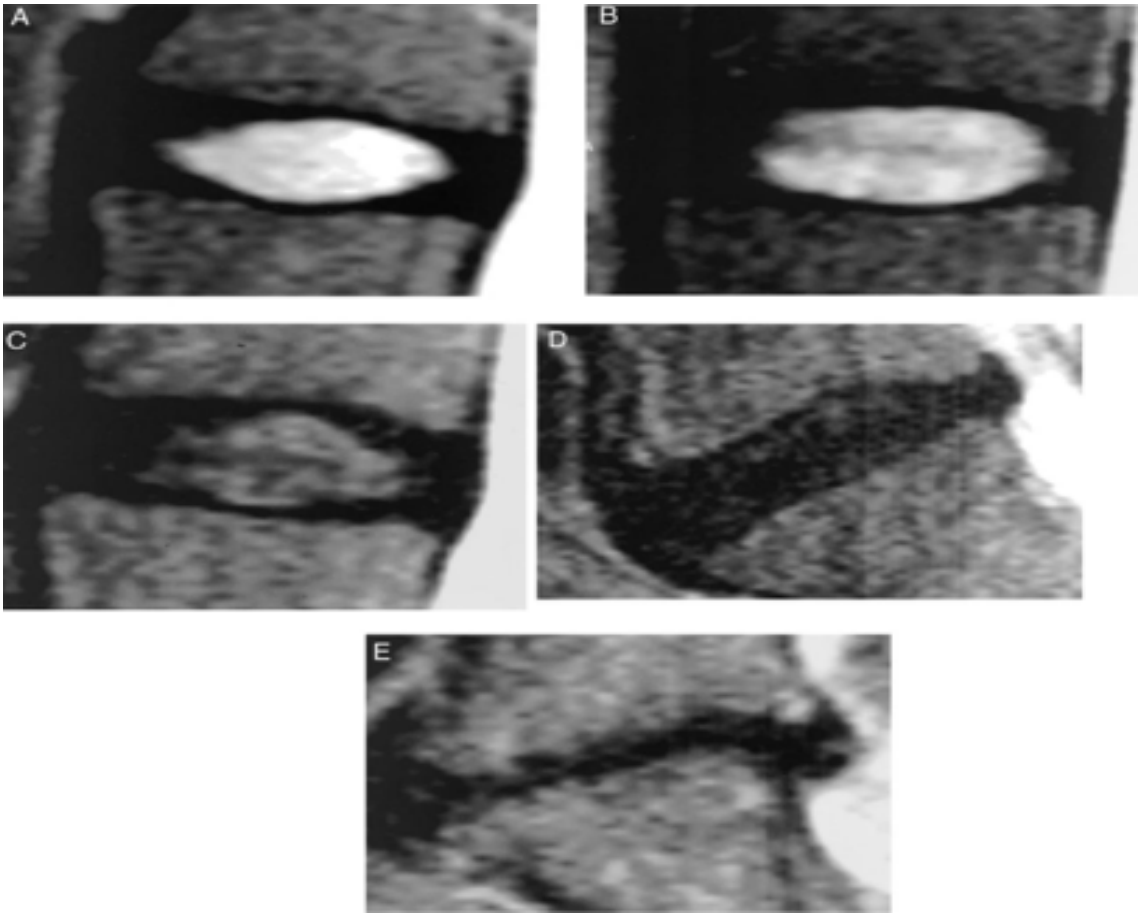

Figure 1. Pfirrmann Grading system for the assessment of lumbar disc degeneration. A- grade I, B- grade II, C- grade III, D- grade IV, E- grade V [12].

Table 3. Pfirrmann grading system of intraspinal extradural lumbar nerve root compression [13].

Grade	Compromise
Grade 0 normal	There is no evident contact of disc material with the nerve root, and the epidural fat layer between the two is preserved. The nerve root has a normal position (Fig. 2).
Grade 1 (contact)	There is visible contact of disc material with the nerve root, and the normal epidural fat layer between the two is not evident. The nerve root has a normal position, and there is no dorsal deviation (Fig. 3).
Grade 2 (deviation)	The nerve root is displaced dorsally by disc material (Fig. 3).
Grade 3 (compression)	The nerve root is compressed between disc material and the wall of the spinal canal; it may appear flattened or be indistinguishable from disc material (Fig. 4).

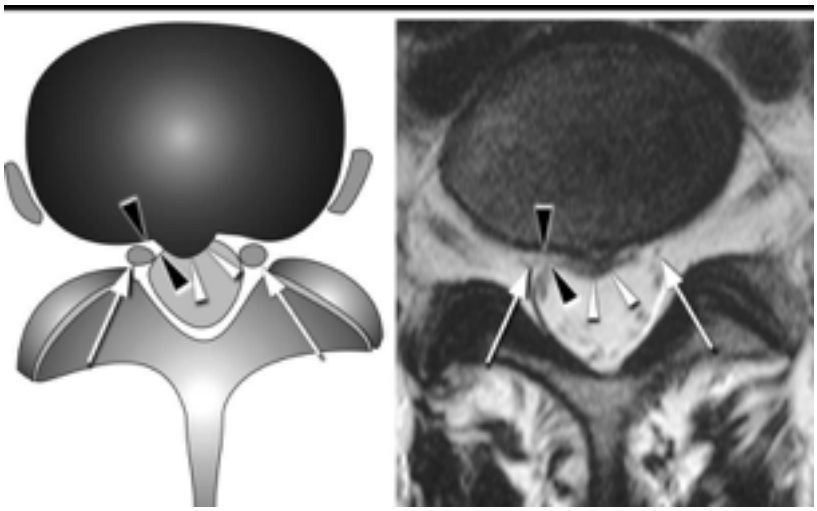

Figure 2. Grade 0, diagram (left) and transverse T2-weighted MR image (right) shows no compromise of the nerve root. A normal epidural fat layer (black arrowheads) is visible between the nerve root (arrows) and the disc material (white arrowheads) [13].

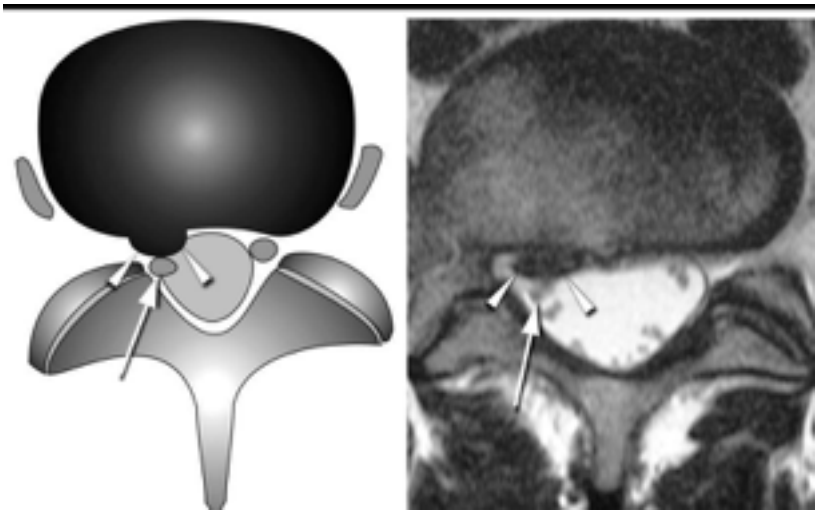

Figure 3. Grade 1, diagram (left) and transverse T2-weighted image (right) shows contact of disc material (arrowheads) with the right nerve root (arrow). No epidural fat layer is visible between the nerve root and the disc material. The nerve root is in the normal position and is not dorsally deviated [13].

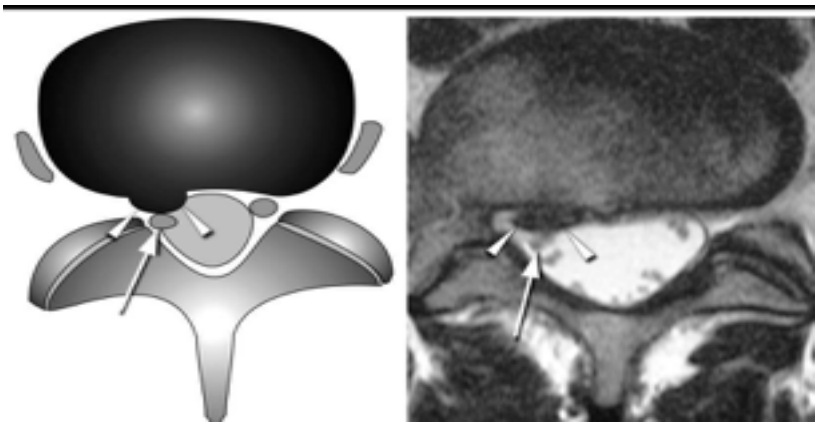

Figure 4. Grade 2, diagram (left) and transverse T2-weighted image (right) shows dorsal deviation of the right nerve root (arrow), caused by contact with disc material (arrowheads)[13].

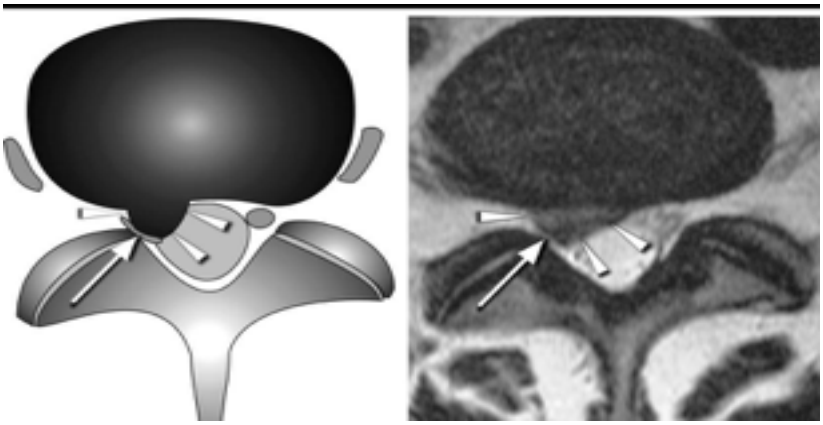

Figure 5. Grade 3, diagram (left) and transverse T2-weighted image (right) shows compression of the right nerve root (arrow) between disc material (arrowheads) and the wall of the spinal canal. The nerve root appears flattened and is indistinguishable from disc material [13].

Table 4. Study variables

Main Variable	Type	Category
Pfirrmann Classification	Signs of disc degeneration (Fig. 1)	grade I
		grade II
		grade III
		grade IV
		grade V
	Intraspinal extradural lumbar nerve root Nerve root compromise (figures 2, 3,4 & 5)	Grade 0 (normal)
		Grade 1 (Contact)
		Grade 2 (Deviation)
		Grade 3 (Compression)

RESULTS

The age of the patients was ranging from 24-72 years, the mean age was 45.1years (+/- 13.52 SD). The age incidence in our study is shown in table (5), while sex distribution of the patients in our study consists of 33 females (66%) and 17 males (34%). It is shown graphically in figure (6).

Table (6) shows the findings of (Pfirrmann' grades for disc degeneration and nerve root compression) reported in 50 MRIs by the 2 radiologists independently .

Pfirrmann grades I and II intervertebral discs are normal while types II -V are degenerated [19].

Normal Intervertebral discs (grade I and II) were most frequently seen at L1-2 level (31.8 %) with grade II being the majority (84 %), Figure (7).

The incidence of disc degeneration (grade III-V) was most frequent at L4-5 level (23.2 %) and least frequent at L1-2 level (15.5 %), Figure (8).

Incidence of various Grades of IVD degeneration at each lumbar level is shown in figure (9).

Among all 3 abnormal grades (3-5); Grade 3 was the most common type (48.6 %), Figure (10).

The incidence of nerve root compression (grade 1-3) was most frequent at L4-5 level (34.85 %) and is shown graphically Figure (11).

Among all grades of nerve compromise (1-3) grade 1 (contact) was most frequent (54.7 %), Figure (12).

Interobserver agreement: The interobserver agreement for disc degeneration (Pfirrmann grades) was moderate (k = 0.431) (P value < 0.0001), Table (7).

Agreement between different grades of degeneration is shown in the Table (8).

The interobserver agreement for nerve root compromise was moderate (k = 0.417) (P value < 0.0001), Table (9).

Agreement between different grades of nerve root is summarized in Table (10).

Table 5 . The age distribution of the cases in our study

Age	Frequency	Percentages
21 - 30 Years	8	16 %
31 - 40 Years	12	24 %
41 - 50 Years	13	26 %
51 - 60 Years	8	16 %
71 - 80 Years	2	4 %

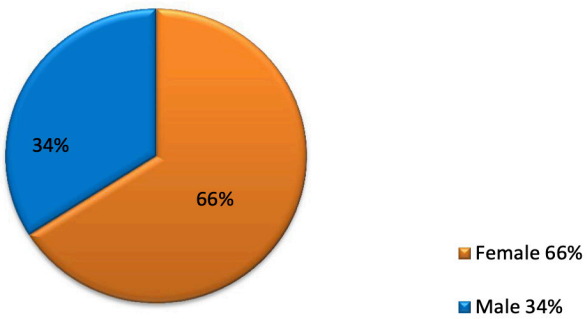

Figure 6. Gender of patients in our study.

Table 6. Findings of disc degeneration & nerve roots compression according to Pfirrmann classification was reported by the two radiologist.

MRI Variables		Number of finding (per 100 reports*)				
Levels		**L1-2**	**L2-3**	**L3-4**	**L4-5**	**L5-S1**
Pfirrmann's score (Disc degeneration)	I	7	3	3	5	4
	II	37	26	19	11	23
	III	28	41	44	27	36
	IV	25	28	27	54	30
	V	3	2	7	3	7
Nerve root compromise	Grade 0 (normal)	89	69	40	16	45
	Grade 1 (contact)	9	23	33	34	33
	Grade 2 (deviation)	2	8	19	26	9
	Grade 3 (compression)	0	0	8	24	13

*** Number of reports was 100 (50 images interpreted by two radiologists).**

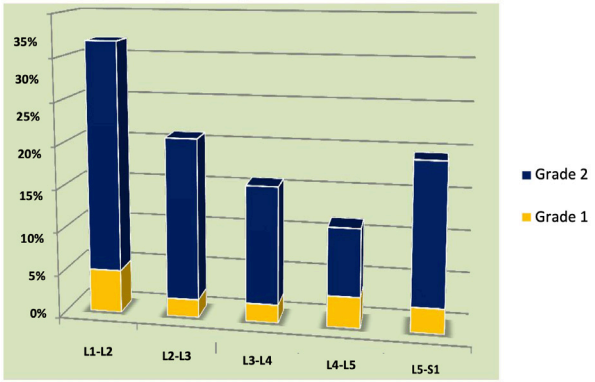

Figure 7. Frequency of normal IVD s (grade I and II) at each level.

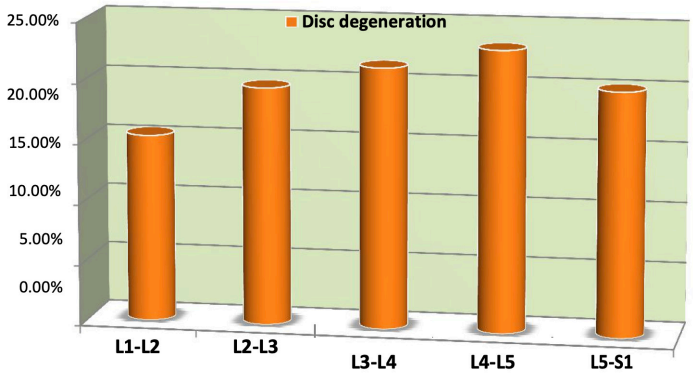

Figure 8. Incidence of disc degeneration at each lumbar level.

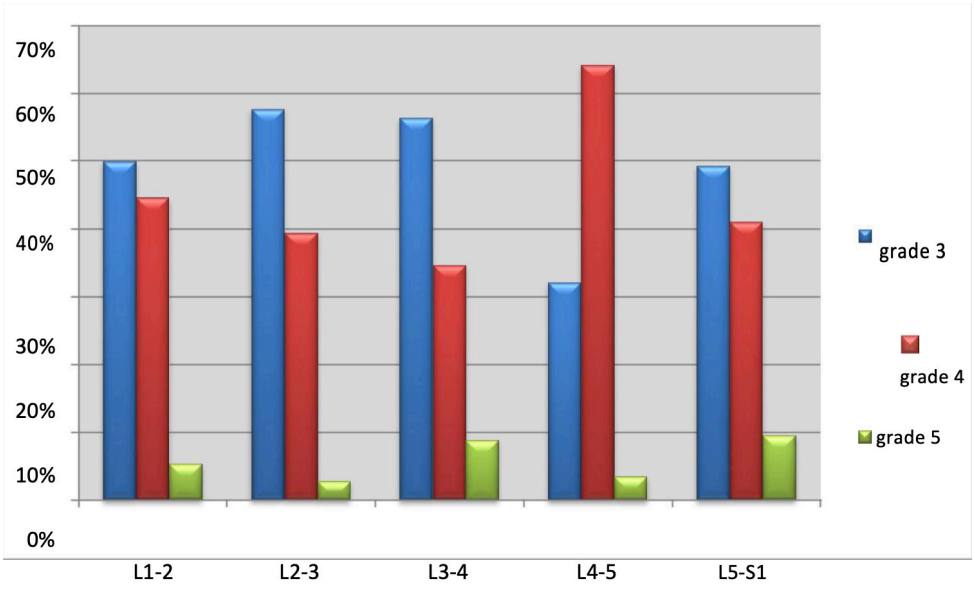

Figure 9. Incidence of various Grades of IVD degeneration at each lumbar level.

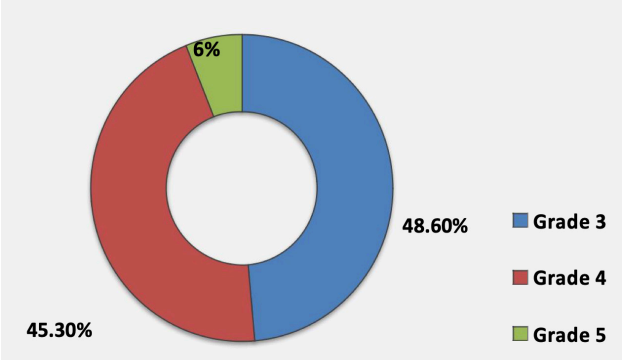

Figure 10. Frequency of various grades of disc degeneration.

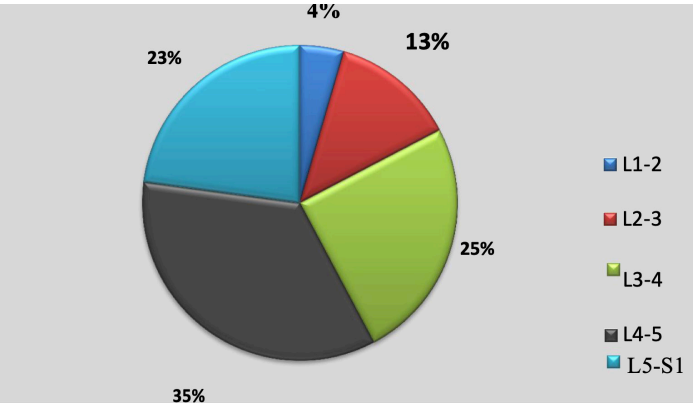

Figure 11. Incidence of nerve root compromise at each level.

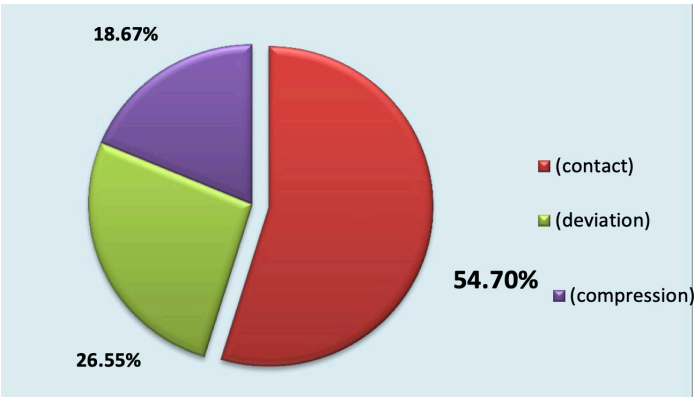

Figure 12. Frequency of nerve root compromise according to type.

Table 7. Interobserver agreement for grades of disc degeneration.

Disc degeneration		Reader A						Kappa	P value
		I	II	III	IV	V	Total		
Reader B	I	0	1	0	0	0	1	0.431	< 0.0001
	II	20	24	19	1	0	64		
	III	1	27	48	27	2	105		
	IV	0	0	4	62	5	71		
	V	0	0	0	3	6	9		
	Total	21	52	71	93	13	250		

Table 8. Influence of Category (Grades I to IV) on Interobserver reliability for disc degeneration

Grades	Interobserver agreement	
	Kappa value	P value
Agreement in one grade difference level		
Grades I and II	0.044	0.37
Grades II and III	0.191	0.037
Grades III and IV	0.568	< 0.0001
Grades IV and V	0.54	< 0.0001
Agreement in two grades difference level		
Grades III and V	0.837	< 0.0001
Grades II and IV	0.972	< 0.0001
Agreement in three grades difference level		
Grades II and V	1.00	< 0.0001

Table 9. Interobserver agreement for nerve root compromise.

Nerve Root		Reader					Kappa	P value
		Grade 0	Grade I	Grade II	Grade III	Total		
	Grade 0	94	55	5	2	156		
	Grade I	8	19	18	3	48		
Reader B	Grade II	1	8	10	4	23	0.417	< 0.0001
	Grade III	0	2	8	13	23		
	Total	103	84	41	22	250		

Table 10. Influence of Category (Grades 0 to III) on Interobserver reliability for nerve root compromise.

Grades	Interobserver agreement	
	Kappa value	P value
Agreement in one grade difference level		
Grades 0 and I	0.195	< 0.0001
Grades I and II	0.249	0.03
Grades II and III	0.318	0.053
Agreement in two grades difference level		
Grades I and III	0.732	< 0.001

DISCUSSION

Magnetic resonance imaging (MRI) is the most important method for the clinical assessment of intervertebral disc pathology and degenerative spinal disease. The signal characteristics of the disc in T2-weighted MRIs reflect changes caused by aging or degeneration [12, 16]. The current study, algorithm & grading system are based on MRI signal intensity , disc structure, disc height & distinguishing disc nucleus from its annulus. Disc height is crucial for discrimination between grade IV & V discs while not in grades III & IV. The changes in the vertebral body marrow adjacent to the intervertebral disc are not included in this study. The signal loss of the disc on T2-weighted MRIs correlates with the progressive degenerative changes of the intervertebral disc [13].

This inter observer agreement was done for the first time in our region regarding MRI interpretation concerning degenerative lumber spine changes at Diagnostic Imaging Center in Sulaimani for 50 patients with low back pain. The same images were interpreted independently by two radiologists for grades of disc degeneration and nerve root compromise.

The interobserver agreement in our study for grades of disc degeneration was moderate (k= 0.431) this is in concordance with a study done by (Arana et al. 2010) [17] in which agreement was moderate (k= 0.491). A study done by (F. M. Kovacs et al. 2009) [18] showed fair agreement (k = 0.219), while our study showed moderate agreement. This is possibly due to:

1- The MRI machine: they used an open 0.2T (Tesla) system [18] , while our MRI was closed 1.5T system.

2- Number of the radiologists was two and were

working in the same center in our study, while they were seven radiologists from two different geographic settings in Spain [18].

A slight agreement was observed between Grades I and II, and between Grades II and III, this can be explained by the main discriminating features between these grades (homogeneous versus inhomogeneous bright nucleus for Grades I and II, as well as disc height which is normal versus normal-slightly decreased for grades II and III respectively) which are subject to a larger scope of interpretation than for the other grades, this grade variation corresponds to the initial study done by (Pfirrmann et al. 2001) [12].

However strong agreement was seen between 2 grades (III and V) (k= 0.837, p value < 0.0001) and absolute agreement seen between 3 grades difference, (II and V =1, p value < 0.0001), there was poor agreement in lower grades (between I-II and II-III) which may affect the overall agreement. Incidence of disc degeneration (grade 3-5) in our study was most frequent at L4-5 level and least frequent at L1-2 level, this is going with three other studies done by (F. M. Kovacs et al. 2009) [18], (Arana et al. 2010) [17] and (S Verma et al.2010) [19]. These findings support the fact that mechanical characteristics of the discs are greater in those that are adjacent to fused lumbar vertebrae, favoring degeneration, with high disc involvement observed more commonly at lower lumbar level. It was evaluated and confirmed that with aging the loss of proteoglycans from the lumbosacral disc exceeds that from the lower lumbar discs because of its proximity to a rigid segment, that is, the sacrum [20].

Agreement for nerve root compromise in our study was moderate (k= 0.417, p <0.0001)this is in concordance with a study done by (Lurie et al. 2008) [21] which showed moderate agreement (k = 0.47). The interobserver agreement between grades of nerve root compromise showed that slight-fair agreement was seen in one grade while there was good agreement between two grades (table 10), this may be due to narrow discriminating feature between the two (adjacent) grades causing higher subjective interpretation.

The analysis of agreement in our study shows that the system we tested is reliable in discriminating between higher grades of nerve root compromise (II and III) but slightly less reliable in discriminating between grades (0 and 1). This grade variation is corresponding to the initial study done by (Pfirrmann et al. 2004) [13] which stated that agreement was slightly less frequent between the two lowest grades (normal and contact) than between the other two grade pairs [13]. The incidence of nerve root compromise in our study (grade 1-3) was most frequent at L4-5 level (34.85 %), the same result was found in a study done by (S Verma et al. 2010) [19].

From this study we recommend using Pfirrmann classification system in our lumbar MRI assessment and reporting. Pfirrmann classification system opens communications tunnels between radiologist & clinicians. We recommend another study being more comprehensive using beside Pfirrmann grading system the disc contour, evaluating vertebral endplate signal changes and spinal canal diameter.

From our study we conclude that moderate interobserver agreement was seen between the two radiologists in assessing lumbar IVD degeneration and nerve root compromise using Pfirrmann classification system. Pfirrmann classification system is reliable and provides standardized assessment of lumbar IVD and nerve roots.

REFERENCES

1. Osborn, Ross, Salzman, Crim, Borg, Moore, Shah Jhaveri, Hamilton , Blaser, Katzman. Expert Differential Diagnosis brain and Spine. First edition. Amirsys, 2009. Spine II. Page: 52.

2. O. Airaksinen, J. I. Brox, C. Cedraschi et al, European guidelines for The management of chronic nonspecific low back pain, chapter 4 , Eur Spine J (2006) 15 (Suppl. 2): S 208.

3. Stephanie Ryan, Michelle McNicholas, Stephen Eustace. Anatomy for Diagnostic Imaging, second edition 2004, Page: 100-104.

4. David Sutton. Textbook Of Radiology And Imaging. Seventh edition 2002, neuroradiology of the spine. Page: 1643. Battie MC, Videman T. Lumbar disc degeneration: epidemiology and genetics. J Bone Joint Surg Am 2006; 882: 3-9.

5. Antoniou J, Steffen T, Nelson F, Winterbottom N et al. The Human Lumbar Intervertebral Disc: Evidence for Changes in the Biosynthesis and Denaturation of the Extracellular Matrix with Growth, Maturation, Ageing and Degeneration. J. Clin.Invest.1996; 98:996–1003.

7. Yu L-P, Qian W-W, Yin G-Y, Ren Y-X, Hu Z-Y (2012) MRI Assessment of Lumbar Intervertebral Disc Degeneration with Lumbar Degenerative Disease Using the Pfirrmann Grading Systems. PLoS ONE

7(12): e48074, page: 1.

8.Modic MT, Masaryk TJ, Ross JS, et al. Imaging of degenerative disk disease. Radiology 1988; 168:177–86.

9. Pearce RH, Thompson JP, Bebault GM, et al. Magnetic resonance Imaging reflects the chemical changes of aging degeneration in the human intervertebral disk. J Rheumatol Suppl 1991; 27:42–3.

10. Sether LA, Yu S, Haughton VM, et al. Intervertebral disk: normal Age-related changes in MR signal intensity. Radiology 1990; 177:385–8.

11. Timothy L. Tan, BS1, Sean L. Borkowski, MS2, Sophia Nicole Sangiorgio, et al. Quantification and Classification of Lumbar Disc Degeneration. http://www.ors.org/Transactions/60/0747.pdf

12. Christian W. A. Pfirrmann, Alexander Metzdorf, Marco Zanetti et al. Magnetic Resonance Classification of Lumbar Intervertebral Disc Degeneration. Spine 2001; 26:1874-1878.

13. Christian W. A. Pfirrmann, Claudio Dora, Marius R. Schmid et al. MR image- based grading of lumbar nerve root compromise due to disk Herniation: reliability study with surgical correlation. Radiology 2004; 230: 583-8.

14.Antony J. Viera, Joanne M. Garret . Understanding Interobserver Agreeement : The Kappa Statistics. FamMed 2005; 37:360-3.

15.Landis JR, Koch GG.The measurmnt of observer agreement for categorical data.Biometrics 1977;33:159-174.

16. Carrino JA, Lurie JD, Tosteson AN, Tosteson TD. Lumbar spine: reliability of MR imaging findings. Radiology 2009; 250:161-70.

17.Estanislao Arana, Ana Royuela , Francisco M. Kovacs et al. Agreement in the Interpretation of 1.5-T MR Images by Using the Nordic Modic Consensus Group Classification Form. Radiology 2010; 254:809-817.

18. Kovacs FM, Royuela A, Jensen TS et al .Agreement in the interpretation of magnetic resonance images of the lumbar spine. Acta Radiol 2009; 50:497–506 .

19. S Verma, P Gupta, A Munshi at al. A Retrospective Analysis Of Magnetic Resonance Imaging Findings In 20 -40 Year Old Patients With Low Back Pain. Experience At A Semi Urban Tertiary Healthcare Centre In Northern India. The Internet Journal of Spine Surgery. 2010 Volume 6. Number 1.

20.Shafaq Saleem, Hafiz Muhammed Aslam, Muhammad Aim Khan Rehmani et al. Lumbar Disc Degenerative Disease: Disc Degeneration Symptoms and Magnetic Resonance Image Findings. Asian Spine J 2013;7(4):322-334

21.Jon D. Lurie, Anna N. A. Tosteson, Tor D. Tosteson et al. Reliability of Magnetic Resonance Imaging Readings for Lumbar Disc Herniation in the Spine Patient Outcomes Research Trial (SPORT). Spine (Phila Pa 1976) 2008; 33: 991–9