Quantifying Vertebral Endplate Degeneration Using the Concavity Index

Menekse Salar Barim[1(✉)], Richard F. Sesek[2], M. Fehmi Capanoglu[2], Wei Sun[3], Sean Gallagher[2], Mark C. Schall Jr.[2], and Gerard A. Davis[2]

[1] Oak Ridge Institute for Science and Education (ORISE) Research Fellow, Cincinnati, OH 45202, USA
mzs0053@auburn.edu
[2] Auburn University, Auburn, AL 36830, USA
[3] Data Scientist, Colaberry Inc., St. Louis, MO, USA

Abstract. A novel morphometric measurement of endplate degradation was compared with qualitative ratings of intervertebral disc degeneration (Pfirrmann Grading) in a double-blinded study to investigate a new, quantitative method for relating disc morphology and bony changes using MR imaging techniques known as the "Concavity Index" (CI). By adding a quantitative measure of vertebral endplate degeneration, the CI could provide further insight into structural changes related to disc breakdown and subsequent low back pain. The continuous nature of the CI may also allow medical professionals to more closely monitor a patient's low back health. T2-weighted MRI scans of the sagittal profile of the lumbar endplates (L2-S1) were collected from 50 subjects (25 females and 25 males) whose ages ranged from 20–40 years. Three trained examiners independently measured the height and the concavity levels of each lumbar vertebrae (L2-S1) as well as assessed the health of the intervertebral discs using Pfirrmann's lumbar disc degeneration grading method. Concavity Indices (CIs) were computed by dividing measured concavity level by disc height (CL/DH). A larger CI was hypothesized to be indicative of spinal degradation and subsequent low back pain. Intra- and inter-rater reliabilities were assessed for both the CI measurements and Pfirrmann's lumbar disc degeneration grades. The categorical intra-observer agreement for Pfirrmann ratings ranged from 26 to 63%. However, the CI, which is a continuous measure, varied by only 2% (average absolute error) among raters. Endplate concavity is indicative of fracturing and damage and is hypothesized to lead to subsequent disc degeneration due to impediment of nutrient flow to the discs themselves. The CI shows promise as a means for potentially quantifying low back health and identifying risk for future low back pain prior to significant disc degeneration.

Keywords: Vertebrae degeneration · MRI · Concavity index · Pfirrmann grading

S. Bagnara et al. (Eds.): IEA 2018, AISC 820, pp. 734–741, 2019.
https://doi.org/10.1007/978-3-319-96083-8_88

1 Introduction

Low back pain (LBP) represents one of the most costly and prevalent musculoskeletal disorders (MSDs). MSDs are the leading cause of disability in the United States and represent 48% of all self-reported chronic medical conditions [1]. LBP is a major health issue affecting millions of people worldwide [2–4].

Despite advances in imaging technology, the etiology of the underlying pain is frequently illusive. Morphological changes related to normal disc aging often appear on MR imaging without any corresponding symptoms. Despite an incomplete understanding of the relationship between physical changes and pain, these MRI-detectible morphological changes show predictive promise and warrant further discussion [5]. Interest in biomechanical models of the spine, particularly detailed knowledge regarding spinal morphometry and the relationships between vertebral segments and corresponding intervertebral discs has been increasing. Several quantitative studies have investigated the external geometry of the vertebrae and adjacent intervertebral discs for different regions of the human spine. In biomechanical models, assumptions have been made assuming the spine acts as a single straight line without considering the volume or the curvature of the spine [6] or incorporating average measurements into models without considering the impact of personal characteristics or differences between individuals [7]. Lakshmanan discovered that the majority of lumbar endplates were concave, while the majority of sacral endplates were flat [8]. In addition to this study, Larsen proposed that concavity was nearly always the result of physical loading [9].

In vertebral motion segment testing, the endplate is typically the first structure to become damaged, and is clearly the weakest link when the spine is loaded in compression [10]. Additionally, it has been shown that compressive loading fractures the endplate before damaging the disc. Increases in concavity related to subsidence can result in vertebral endplate scarring that may impede the flow of nutrients to the intervertebral disc (IVD). It is proposed that increased concavity is related to increased endplate damage and therefore subsequent disc degeneration [10]. In other words, increases in concavity proceed and are related to disc degeneration [11].

The relationship between vertebral disc degeneration and low back pain is well established [12, 13]. This study proposes a novel approach for quantifying vertebral concavity and relating this to disc degeneration. Unlike measures of disc health, which are subjective and discrete, this approach uses a mathematical measure, which quantifies endplate health continuously.

2 Methodology

2.1 Subject Data

MRI scans were collected from fifty subjects (25 male and 25 female) with ages ranging from 20 to 40 (mean 31.1 $\pm$ 5.4 years). All Subjects were scanned on a whole body 3T Magnetic Resonance Imaging machine (Siemens Verio open-bore). All imaging was performed in the supine position. Informed consent was obtained from all subjects. A T2-weighted image, which provides a comprehensive perception of disc structure and good tissue differentiation, was used for the morphological evaluation of intervertebral discs. MRI was performed on the intervertebral discs from L2-S1. MRI data were obtained using a dedicated abdominal coil. The protocol included the following sequences: Axial Continuous T2-weighted, Sagittal Continuous T2-weighted, and Axial Multi group T2-weighted images with the following parameters; T2-weighted spin-echo (TR – 3440 ms; TE – 41 ms). All MR images were obtained at a 3-mm slice thickness with 385 FoV read & 100% FoV phase.

2.2 Image Assessment

Three, Level-3 certified MRI observers, blinded to each other's measurements, graded lumbar IVDs and adjacent vertebral bodies using both the Pfirrmann Grading system and the CI in a randomized sequence. The Pfirrmann Grading System is widely used and related to the height and health of the IVD. Pfirrmann categorizes IVDs into five grades [14, 15]. Grades I and II have normal disc height and have a healthy structure when compared with other levels. Grades III and IV demonstrate some height changes (becoming narrower) relative to other discs and the disc's structure also begins to change. With Grade V, the distinction between disc nucleus and annulus is completely lost and the disc space has collapsed completely. The grading scale and progression from a healthy disc (Grade I) to a severely compromised disc (Grade V) is illustrated in Fig. 1 [16].

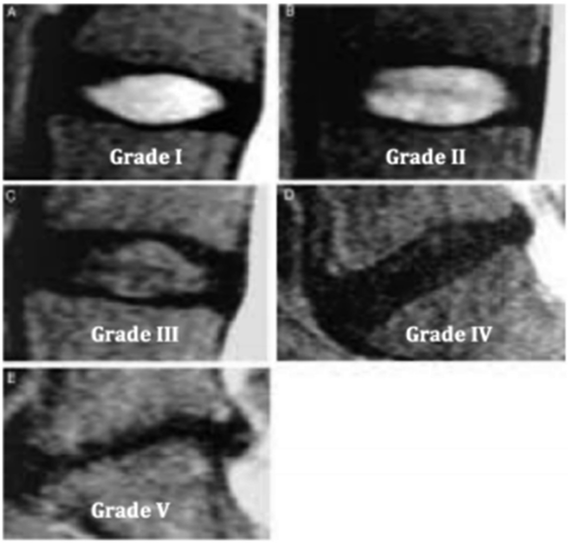

Fig. 1. Pfirrmann grading scores [16]

T2-weighted sagittal MRI scans were used for both Pfirrmann grading and the CI evaluation of lumbar endplates. To calculate the CI, each examiner measured the height and the concavity levels of the lumbar discs (L2–S1). CIs were measured as follows: the superior aspect of vertebral body lengths, which are the distances in the sagittal plane between anterior and posterior borders of vertebral body, were traced first. Then, the perpendicular distance between this line and the vertebral body was measured. This was defined as the concavity level. The concavity level was then divided by the corresponding disc height (CL/DH) (shown in Fig. 2). All measurements were performed using sagittal lumbar spine T2-weighted images and the Osirix software system (OsiriX v8.0.1, 2016).

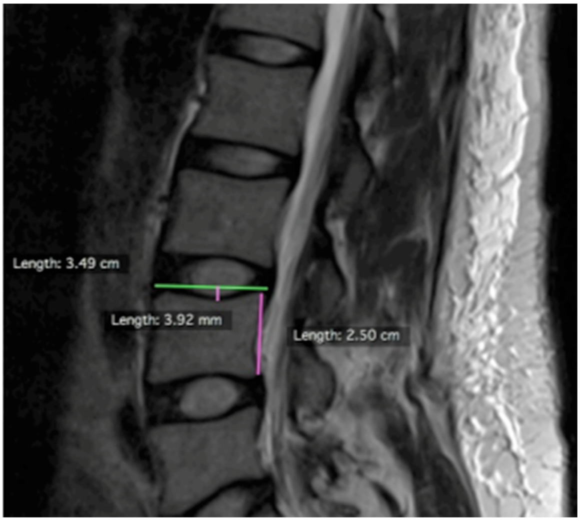

Fig. 2. Measurement of concavity index

3 Results

The absolute agreement between observers was relatively low for Pfirrmann Grading. The highest absolute agreement between any two observers was 63% and the lowest was 26%. Table 1 shows the Cohen's Kappa analysis between observers where "−1" and "+1" represent differences of 1 category.

Table 1. Probability of absolute and near (±1) category inter-observer agreement

Agreement	−1 Category	Absolute agreement	+1 Category	Kappa
Obs 1 vs Obs 2	0.28	0.38	0.32	0.18
Obs 2 vs Obs 3	0.16	0.63	0.17	0.61
Obs 1 vs Obs 3	0.34	0.26	0.36	0.18

High CI agreement (≥ 0.98) between observers suggests that CI measurements are consistent and potentially reliable when compared to Pfirrmann Grading. Table 2 shows the correlation coefficients among CI observers. Each observer's ratings compared against one another and a high level of agreement among observers (r^2 ranging from .963 to .983) was found. The relationship between CI and Pfirrmann is modest with variation in CI at each Pfirrmann level (shown in Fig. 3).

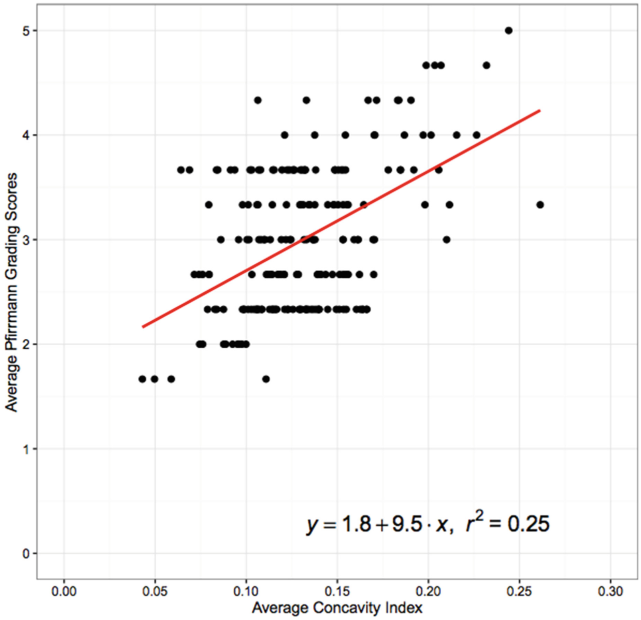

Fig. 3. Relationship between Pfirrmann grading and concavity index

A possible explanation for this modest relationship is that the subject pool was a relatively healthy population with few Grade-V discs. Another possible explanation for lack of fit may be a function of observer experience with Pfirrmann grading.

Table 2. Pearson Correlation Coefficient between observers for concavity index

	Observer 1	Observer 2	Observer 3
Observer 1	1.0	0.99	0.99
Observer 2	0.99	1.0	0.98
Observer 3	0.99	0.98	1.0

There appears to be a clear observer effect with respect to Pfirrmann gradings. However, no such observer effect was detected when ICCs were computed for CIs. In fact, ICC model results were similar for one-way and two-way random effects analyses with both yielding ICC values of 0.985 and standard errors of 0.0026 and 0.0027, respectively. The results indicate that ICC is very high, while agreement among Pfirrmann ratings was relatively poor.

4 Discussion

Most inter-rater disagreements for the Pfirrmann scores were within one category and occurred when classifying grade II and III discs. The difference between grades II and III is heavily dependent open disc height which can be difficult to reliable assess visually. This is further complicated since normal disc height is not uniform across all levels in all subjects and disc height often decreased at the L5-S1 compared with other levels, even when disc health appeared to be otherwise healthy (e.g., good color and uniformity). Since the study population was asymptomatic, very few Grade-V discs were observed. Surprisingly, there were also very few Grade-I discs. A broader range of subject ages and symptom statuses would likely provide a wider distribution of Pfirrmann grades and CIs. With a wider distribution of data, relationships may become more clear. Pfirrmann grading relies solely on visual appearance of T2-weighted images. It is possible, however, to more quantitatively determine water and proteoglycan content of discs using MRI signals other than T2-weighted images. This may be used to enhance the Pfirrmann grading system and make the process more objective. The novice Pfirrmann graders in this study might benefit from such an enhancement.

Also, this study did not consider subject symptoms or low back pain. A prospective study including subjects with and without LBP could address this limitation. All of the subjects in this study were young (20 to 40 years of age) and were relatively healthy college students free of LBP. A diverse sampling of subjects from a greater age range and with varying occupational risk factors could address this limitation. Accurate knowledge of normal and degenerative lumbar intervertebral discs is important for medical professionals. Using the CI, medical professionals can potentially make more accurate and early diagnostic interpretations and, subsequently, more precise surgical interventions regarding lumbar vertebrae and intervertebral discs. The CI demonstrated very high agreement despite the lack of medical experience of the research team. Adding objective elements to the Pfirrmann grading system (such as water content evaluation using other MRI signals) could be beneficial for both inexperienced and experienced observers alike.

An ideal classification system for disc degeneration should be simple, easy to apply, discriminatory, and reproducible with good intra-rater reliability [21]. Results of this anatomic study of morphometric measurement for the lumbar vertebrae suggest that the CI method described within has promise for objectively quantifying low back health and possibly predicting future low back pain. The CI allows for relative comparisons because it is a continuous measure rather than an ordinal scale. On the contrary, and consistent with other studies [17], agreement between observers for Pfirrmann grading was relatively low. The Pfirrmann scoring system is simple and easy to use, but its subjective nature lacks the ability to subtly discriminate degradation. The CI, on the other hand, has demonstrated strong intra-rater reliability while being a more objective approach to assessing vertebral health. Together, CI and Pfirrmann paint a more complete picture of intervertebral motion segment health.

5 Conclusion

A novel approach for quantifying vertebral degeneration has been proposed and there appears to be a positive linear relationship between the CI and Pfirrmann grading. The Pfirrmann grading system is widely used and accepted. The CI may provide a complimentary measure capable of predicting disc degeneration and that could be used in conjunction with Pfirrmann grading to provide a more complete assessment of the health of a given spinal motion segment. The CI is easy to apply; requiring limited previous knowledge of MRI scans or low back geometry.

The CI in conjunction with the established Pfirrmann ratings can provide a more complete picture of low back health and could potentially provide a more comprehensive assessment of spinal segment health.

Acknowledgement. This publication was partially supported by Grant # 2T420H008436 from NIOSH. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention. Use of trade names is for identification only and does not constitute endorsement by the Public Health Service or by the U.S. Department of Health and Human Service.

References

1. BMUS, The Burden of Musculoskeletal Diseases in the United States (2011) United States bone and joint initiative: the burden of musculoskeletal diseases in the United States, prevalence, societal and economic cost, 2nd edn. American Academy of Orthopaedic Surgeons, Rosemont
2. Pope MH, Goh KL, Magnusson ML (2002) Spine ergonomics. Annu Rev Biomed Eng 4:49–68
3. Brooks PM (2006) The burden of musculoskeletal disease – a global perspective. Clin Rheumatol 25(6):778–781 ISSN 0770-3198
4. Woolf AD, Pfleger B (2003) Burden of major musculoskeletal conditions. Bull World Health Organ 81(9):646–656 ISSN 0042-9686
5. Ract I, Meadeb JM, Mercy G, Cueff F, Husson JL, Guillin RA (2015) Review of the value of MRI signs in low back pain. Diagn Interv Imaging 96(3):239–249
6. Merryweather AS, Loertscher MC, Bloswick DS (2009) A revised back compressive force estimation model for ergonomic evaluation of lifting tasks. Work 34(3):263–272
7. Chaffin DB (1969) A computerized biomechanical model-development of and use in studying gross body actions. J Biomech 2(4):429–441
8. Lakshmanan P, Purushothaman B, Dvorak V, Schratt W, Thambiraj S, Boszczyk BM (2012) Sagittal endplate morphology of the lower lumbar spine. Eur Spine J 21:16–164
9. Larsen JL (1985) The posterior surface of the lumbar vertebral bodies, part I. Spine 10 (1):50–58
10. Adams MA (2004) Biomechanics of low back pain. Acupunct Med 22(4):178–188
11. Wang Y, Videman T, Battie MC (2012) Lumbar vertebral endplate lesions: associations with disc degeneration and back history. Spine 1(31):1490–1496
12. Kumar MN, Baklanov A, Chopin D (2001) Correlation between sagittal plane changes and adjacent segment degeneration following lumbar spine fusion. Eur Spine J 10:314–319

13. Kepler CK, Ponnappan RK, Tannoury CA, Risbud MV, Anderson DG (2013) The molecular basis of intervertebral disc degeneration. Spine J 13(3):318–330
14. Rajasekaran S, Babu JN, Arun R, Armstrong BR, Shetty AP, Murugan S (2004) ISSLS prize winner: a study of diffusion in human lumbar discs—a serial magnetic resonance imaging study documenting the influence of the endplate on diffusion in normal and degenerate discs. Spine (PhilaPa 1976) 29(23):2654–2667
15. Marinelli NL, Haughton VM, Munoz A, Anderson PA (2009) T2 relaxation times of intervertebral disc tissue correlated with water content and proteoglycan content. Spine (Phila Pa 1976) 34(5):520–524
16. Pfirrmann CW, Metzdorf A, Zanetti M, Hodler J, Boos N (2001) Magnetic resonance classification of lumbar intervertebral disc degeneration. Spine 26:1873–1878
17. Griffith JF, Wang YX, Antonio GE, Choi KC, Yu A, Ahuja AT et al (2007) Modified Pfirrmann grading system for lumbar intervertebral disc degeneration. Spine 32:708–712