The lumbar spine anatomy, degeneration, and low back pain

Intervertebral discs (IVD) have a complex structure, whereby a central nucleus pulposus (NP) is surrounded by layers of annulus fibrosus (AF). In young adults with healthy IVDs, the NP has high water and glycosaminoglycan contents, which provide resistance to compression. AF naturally has low water content, but it is high in collagen content from concentric lamellar sheets of collagen fibers. Between an IVD and a vertebral body (Figure 1B), or the disco-vertebral junction (DVJ), there exists a ~1 mm thin layer of connective tissue known as cartilaginous endplate (CEP). The normal CEP in adults consists of ~0.1 mm thick calcified and thicker uncalcified cartilage layers, whose roles include attachment of the IVD to the vertebral body, and facilitation of transport of solutes into and out of the IVD via an adjacent bed of capillaries within the bony vertebral endplate (1). The vertebral body, consisting of cortical and trabecular bone along with bone marrow, is another integral part of the lumbar spine.

Low back pain afflicts a large number of people (2) and may involve degeneration or injury in a number of components of the lumbar spine, along with back muscle injury. These may include IVD degeneration (3), fracture of the vertebral body due to compression (4), fracture at the DVJ developing into Schmorl's node (5), as well as bone marrow Modic changes (6,7). Cartilaginous and bony endplates near the DVJ are also likely to be pain sources, given rich nerve endings in the region (8), and increased
density of nerve fibers in the DVJ observed in painful patients (9).

There is a need to evaluate multiple components of the lumbar spine non-invasively. This article covers a number of techniques available to evaluate the lumbar spine, including conventional and novel morphologic and quantitative sequences, as well as considerations for analyzing and post-processing MR images for quantitative evaluation.

**Conventional morphologic MR imaging of the lumbar disc**

The anatomy of the lumbar spine can be partially evaluated by conventional MR techniques. Using spin echo sequences with different weighting, a number of components of the spine can be observed directly, unlike images from the plain film or computed tomography. In clinical settings, T2- and T1-weighted spin echo images taken in sagittal (Figure 2A,B) and axial (Figure 2B,C) planes may be used to evaluate spinal conditions including disc degeneration, disc herniation, and abnormalities of the vertebral bones, bone marrow and the spinal cord. T2-weighted images are sensitive to disc hydration (10), and shows NPs of relatively normal discs with high signal intensity (Figure 2A, L2/L3 and L3/L4, arrowheads), and NPs of degenerated discs with low signal intensity (Figure 2A, L1/L2 and L5/S1, arrows). The degenerated disc at L5/S1 also has a posterior protrusion with a high signal intensity zone (Figure 2A,B, curved arrow). On the axial images, the shape of the protrusion (Figure 2C,D, curved arrows) is seen as broad-based, covering 25 to 50% of disc circumference.

Morphologic grading systems have been devised to
determine severity and morphology of IVD degeneration visible on conventional MR images. The Pfirrmann grading looks at disc structure, signal intensity and disc height (11,12) in sagittal T2-weighted spin echo images, to grade IVDs from 1 (normal) to 5 (complete collapse). These grading systems have been adapted to evaluate efficacy of disc therapeutics in animals (13,14). Additionally, other features such as annular tears and fissures, vertebral body marrow changes, shape of disc herniation in the sagittal and axial planes can be evaluated per established classification (15) and nomenclature (16) schemes with high accuracy (17,18).

Conventional quantitative MRI of the lumbar discs

In an effort to supplement the conventional methods, and to provide more objective biomarkers for spine health, several quantitative MR techniques have been implemented. Quantitative MR measures of the disc have been shown to correlate with biochemical content (19), biomechanical function (20), and even discogenic pain (21), which may supplement conventional MR evaluation. Potential applications include the early detection of disc degeneration and evaluation of the efficacy of disc treatment.

T1 relaxation constant is the rate of the regrowth of longitudinal magnetization, and techniques such as inversion recovery (22), saturation recovery (23), and variable flip angle (24) method are available. It has been reported that T1 decreases with disc degeneration (10,25), disc herniation (26), and water loss (10,25,27). T1 measurement has also been demonstrated useful when used in conjunction with the delayed gadolinium-enhanced MRI of cartilage (dGEMRIC) technique (28), which uses negatively charged contrast agents that distribute inversely proportional to GAGs and decreases T1 value. Using this technique Vaga et al. (29) reported that ΔT1 (T1 value without contrast minus T1 value with contrast) values correlated inversely with GAG content in discectomy tissues; ΔT1 was higher where there was less GAG.

T2 relaxation is another fundamental MR behavior, and it refers to the decay of transverse magnetization. T2 value is often determined by obtaining multiple images at constant repetition time (TR) while varying echo times (TE), and fitting the signal intensity to an exponential signal decay model (Figure 3). Most scanners nowadays offer product sequences such as multi echo spin echo, which obtains multiple T2-weighted images in one scan (30, 31), and create color maps of T2 values of the disc (Figure 3, insert). While as few as two images with different T2-weighting can be used to created color maps, additional images improve the accuracy in the presence of random noise. T2 values also correlate with water and proteoglycan contents of the disc such that T2 decreases with water loss (10,25,27,32) and proteoglycan loss (27,32). T2 values are also found to be lower in herniated discs from symptomatic subjects (33), suggesting correlation with pain, although the causality is not clear. T2 mapping may be useful when following the efficacy of biologic treatments (34), where subtle biochemical or cellular changes are expected, rather than marked morphologic and signal changes.

More recently, T1rho techniques have been used to evaluate slowly moving macromolecules such as proteoglycans in the NP by measuring transverse relaxation in the presence of a spin-locking pulse (35). In most implementations, reviewed in (36), T1rho preparation pulses with varying spin lock times are applied, followed by image acquisition with a set TR and TE. The resulting images are T1rho-weighted, to which an exponential fitting is performed to obtain T1rho maps. It has been reported that T1rho decreases with water and proteoglycan loss (19,37) and it has also been correlated with discogenic pain (21). Both the T2 (Figure 4A,B) and T1rho (Figure 4C,D) values are sensitive to disc degeneration, and correlate strongly with each other when measured on the same sample (38,39). However, there are subtle differences in spatial distribution of T2 and T1rho values (Figure 4, arrows), even when considering that T1rho values are higher than T2 values in
general. In addition, it has been reported that T1rho values have greater dynamic range and are more sensitive to clinical indices (38) than T2 values. However, the published data is still limited, and additional work is needed to establish which quantitative techniques are best suited to evaluate disc degeneration. Additionally, T1rho sequence are not yet widely available, limiting the use to translational research.

**MR imaging of the CEP and other short T2 tissues with UTE techniques**

While the conventional MR techniques are useful for evaluation of IVD and other soft tissues visible on conventional MR images, several other components of the lumbar spine appear dark and cannot be examined directly. These include the CEP (Figure 1B), longitudinal ligaments (Figure 5A, curved arrow) and ligamentum flavum. Additionally, due to the lack of contrast between the CEP and bony vertebral endplate (Figure 5A, arrows), sclerosis and other subtle bony endplate changes cannot be discerned on conventional images. The low signal intensity from these tissues is due in part to their intrinsically short T2 characteristics. Their T2 values can range from less than 1 ms in the bone (40) to about 2 to 4 ms in the ligaments (41) and the CEP (42). Given relatively long TE values in conventional spine imaging such as spin echo techniques.
(typically 10 to over 100 ms), MR signals from these short T2 tissues decay too rapidly to be captured. While it is possible to capture signal from uncalcified CEP using gradient echo techniques employing shorter TE of 2 to 4 ms (43), the image contrast for the CEP may be suboptimal, and it may not capture signal from calcified layers of the CEP with even shorter T2 values, likely less than 1 ms.

MR sequences have been developed (44-51) in order to image tissues with very short T2 by utilizing the ultrashort TE (UTE) in the order of several microseconds. In certain UTE techniques, the minimum TE can be reduced to as short as 0.008 ms (40). When imaged with a two-dimensional UTE technique (40), a lumbar disc segment can be seen with high signal intensity throughout (Figure 5B), since the MR signal from both long and short T2 tissues are captured at the same time. To modulate image contrast, specifically to suppress long T2 signal while preserving short T2 signals of interest, a simple technique of digital image subtraction (52,53) can be used. Here, two images at different echo times are acquired; the first image is obtained at the minimum UTE, and the second image is obtained at longer TEs, typically greater than a few ms. The first image contains MR signal from both short and long T2 tissues (Figure 5B), while the second image contains the signal largely from the long T2 tissues (Figure 5C) since the short T2 signals have decayed markedly. Therefore, subtracting the second image from the first image can produce an image that accentuates only the short T2 tissues (Figures 5D). Using different TEs (by varying the second TE usually), short T2 contrast can be optimized (54). Other techniques to suppress long T2 signal includes inversion nulling of water (55,56), nulling of water and fat using dual adiabatic inversion recovery (57), and long T2 water saturation (46,58). In addition, short T2 tissues often have short T1 values. NP has a long T1 value over 1000 ms, while T1 value of the CEP is about half of that (43). Therefore, the signal intensity from short T2 tissues can be accentuated by changing TR and flip angle utilizing such T1 differences (43,59). While promising, these techniques have not yet seen a wide-spread use clinically, due to the lack of commercial availability on all platforms, as well as additional cost for the sequence.

Quantitative analysis: MR relaxation fitting

As described earlier, quantitative MR techniques can provide objective and sensitive biomarkers for biochemical contents of tissues and disc health. Quantitative MR relaxation constants can be measured by acquiring series of images with different weighting, and by fitting the signal intensity to an appropriate relaxation model. For example, in order to estimate T2 values for a specific voxel of tissue, multiple T2-weighted images are acquired at different TEs, and then decaying MR signal for each voxel is fitted to a mono-exponential relaxation model, with or without a noise constant, to determine the T2 values and create color maps (Figure 6A). A noise constant is an important factor for accurate estimation of T2 values, especially when the T2 value is small and signal-to-noise ratio (SNR) is low. SNR is defined by the ratio of signal strength (or intensity) to noise strength (or noise standard deviation, σ). When the SNR is low, the noise introduced in the acquired MR image is persistent regardless of scanning parameters (e.g., at long TE), which makes fitting inaccurate. Overestimation of T2 is often observed in this case. To correct for noise, a noise constant may be included in the fitting algorithm (60,61). When the noise strength (σ) can be reliably estimated in the magnitude MR images, these noise-corrected methods result in accurate T2 estimation even when SNR is as low as 20 (61). If SNR is over 100, the noise statistics approximates to a zero-mean Gaussian distribution and an uncorrected fitting without a noise constant can still result in an accurate estimation of T2 values.

It has been evident from nuclear magnetic resonance (NMR) spectroscopic studies that certain biological tissues exhibit a mixed relaxation behavior from multiple T2 components (62). With sufficient SNR (for both long and short T2 components), and number of images, MR imaging could also be used to determine multiple T2 or T2* values at a voxel. In musculoskeletal systems, the advent of UTE techniques has enabled acquisition of short T2* signal and facilitated multicomponent analysis of bony (63), ligamentous (63), and cartilaginous (63-65) tissues. An example of multi-component T2* mapping of a lumbar segment is shown in Figure 6, illustrating large proportion of long T2* component in the NP (Figure 6D, square) and a small proportion of short T2* component in the CEP (Figure 6D, arrow). While promising, multi-component analyses of MR images are difficult since non-unique solutions may exist for a given MR data, especially when SNR is low and substantial noise is present in the images. Many limitations of multi-component analyses have been described (66,67).

Because multicomponent fitting is sensitive to SNR, certain types of constraints are generally used in the fitting procedure. For example, the total number of components...
Figure 6: Comparison of a single component vs. bicomponent T2* analysis of a lumbar intervertebral disc. (A) Single component and (B,C,D) bi-component analysis of UTE T2* data. In the NP (square), (A) single T2* and (B) long T2* values are similar, and (C) there is little variation in short T2* values or (D) their fraction. In the CEP (arrows), however, both (B) long and (C) short T2* values vary in distribution and (D) fraction.

could be limited to 2, representing a short T2* and a longer T2* bi-components in a single voxel. Since the MR images are generally produced as a sum of squares of the real and imaginary signal measurements, the noise has a non-zero value which may act as a long relaxation component in the signal and affect the fitting accuracy (68). Therefore, the noise is also usually included in the fitting model as a constant. The noise constant can be estimated in a separate procedure (63,69,70), or estimated as a parameter during the fitting process (71-73). Care must be taken when the magnitude MR images are obtained with multichannel coils and parallel imaging techniques since these conditions change the noise statistics (74).

In addition to fitting data to models that assume two or more discrete T2 components, models have been developed that assume a distributed T2 components or a continuous T2 spectrum (75). In this approach, a distribution of T2 components (i.e., T2 spectrum) are included in the decay signal model, which is fitted using non-linear algorithms such as a non-negative least squares (NNLS) algorithm (76). Since there are too many unknown parameters (spectrum magnitudes) to be estimated and the problem is ill-posed (77), a constraint is usually incorporated to smooth the magnitudes of the neighboring T2 values (i.e., regularize) (75). This approach does not require a prior information on the number of components, and provides a distribution of T2 components in a tissue. Figure 7 shows an example of multicomponent fitting performed with a bicomponent model (Figure 7A) and a distributed model (Figure 7B). Regularized NNLS algorithms provide resilient fitting results, but biases can occur depending on the SNR and noise types (68). Several variations of the NNLS technique have also been described (78-81). These fitting approaches may be useful when the number of components cannot be determined a priori.

Quantitative analysis: MR image post-processing

After acquisition of MR images, quantitative analyses often require segmentation of the images (82) to define appropriate regions of interest. For the lumbar spine, automated segmentation of the vertebral body and the IVD have been challenging due their irregular shape and heterogeneous signal intensity, especially for degenerated discs (83). Nonetheless, a number of segmentation techniques based on post-processing has been proposed, including thresholding (84,85), edge-detection (86), graph-cut based methods...
Edge detection is a simple but an effective technique that detects an edge based on spatial gradient of signal intensities (91). When applied to a proton-density weighted (TR = 2,000 ms, TE = 10 ms) MR images (Figure 8A), edge detection could effectively segment the vertebral body (Figure 8B) and the IVD (Figure 8C). In this example, the canny edge detection method (91) was used. For segmentation of vertebral body, voxel intensities were evaluated radially from the center of the ROI outward to detect rapidly changing intensity due to the cortical shell. The outermost boundary edges, disconnected focally, are then connected through erosion and dilation process (86) to complete the segmentation. Difficulties arise when the image contains irregular and disconnected bony boundaries, or when bone marrow changes result in irregular internal signal intensities. The segmentation of the IVD in this example relied upon successful segmentation of the vertebral body for the superior and inferior boundaries, and additional edge detection for the anterior and posterior boundaries. Distortion of geometry and signal intensity due to conditions such as disc herniation may degrade the segmentation performance.

The graph-cut method, a relatively recent development that gained popularity, performs segmentation by making “cuts” that have associated energies that can be minimized. As an example, a variation of graph-cut method (92) was applied to a spine MR data. This method finds an energy cost for making cuts between a set of nodes that are on the radial lines extending from the center of the vertebral body outward, to efficiently search a square-shaped structure such as the vertebral body. The algorithm has also been extended to 3-D (93), which uses a cuboid search region and the rays that extend radially in a sphere. Graph-cut methods are being advanced by including spatial information (94) to improve robustness (overcome noise) and accuracy. Limitations of this technique includes variable performance that depends on the position of seed point, as well as lower spatial resolution of the segmentation when the boundary of the segmented region is far from the seed point. Additionally, the technique is not well-suited for segmenting thin objects, or multiple regions of interest simultaneously (95).

Yet another approach is atlas-based segmentation, involving the creation of an atlas (or a template) of a region of interest based on a training data, and applying the atlas to the target data. This approach may be useful for analyzing the NP and AF of the intervertebral disc separately, especially in degenerated discs that lack defined NP (Figure 2A, arrow). In one implementation (89), an atlas of NP and AF was created from MR images of grade 1 discs (that had distinct NPs) and registered (96) against the AF boundary of the target images, for the purposes of determining average T2 values of the NP and AF on color maps. In another implementation (90), a probabilistic atlas was used to overcome the partial volume effect and the overlapping problems of gray-level values. However, the performance of this method depends on the amount and quality of images of discs which form the probabilistic atlas (i.e., training data) and showed reduced accuracy for degenerated cases.

Figure 8 Segmentation of vertebral body and intervertebral disc on MR data. (A) Proton density weighted image was analyzed using a Canny edge detection technique to perform (B) vertebral body segmentation, and subsequently, (C) intervertebral disc segmentation.
Conclusions

In conclusion, quantitative MR sequences and UTE sequences offer additional ways for evaluation of health and injury of the lumbar spine to supplement conventional clinical MR imaging. Sensitivity of quantitative MR measures to degeneration of IVD have been well-established in literature, facilitated by techniques to segment regions of interest and to determine accurate MR measures from the regions using different relaxation models with consideration for noise. UTE sequences offer a unique contrast mechanism useful for previously unevaluated discovevertebral junction of the spine, creating new opportunities to diagnose and characterize the tissue. Combined with continuing basic research on pathophysiology of spine diseases, along with advancements in techniques to accurately localize and quantitatively evaluate different regions of the lumbar spine, a wider-spread in quantitative and novel MR techniques may be realized.

Acknowledgements

Funding: This article was made possible in part by grants from the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (2016R1A2B4015016) in support of Dr. Hwang, the Clinical Science Research & Development of the VA Office of Research and Development (Award Number 5101CX000625; Project ID: 1161961) in support of Dr. Christine B. Chung, and the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health in support of Dr. Won C. Bae (Grant Number R01 AR066622).

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

Disclaimer: The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of the funding agencies.

References


46. Techaowiboonwong A, Song HK, Wehrli FW. In vivo MRI of submillisecond T2(2) species with two-dimensional and three-dimensional radial sequences and applications to the measurement of cortical bone water. NMR Biomed 2008;21:59-70.


