

UCSF

UC San Francisco Previously Published Works

Title

Measuring and reporting of vertebral endplate bone marrow lesions as seen on MRI (Modic changes): recommendations from the ISSLS Degenerative Spinal Phenotypes Group

Permalink

<https://escholarship.org/uc/item/9dw5894q>

Journal

European Spine Journal, 28(10)

ISSN

0940-6719

Authors

Fields, Aaron J
Battié, Michele C
Herzog, Richard J
et al.

Publication Date

2019-10-01

DOI

10.1007/s00586-019-06119-6

Peer reviewed



Published in final edited form as:

Eur Spine J. 2019 October ; 28(10): 2266–2274. doi:10.1007/s00586-019-06119-6.

Measuring and reporting of vertebral endplate bone marrow lesions as seen on MRI (Modic changes); recommendations from the ISSLS Degenerative Spinal Phenotypes Group

Aaron J. Fields¹, Michele C. Battié², Richard J. Herzog³, Jeffrey G. Jarvik⁴, Roland Krug⁵, Thomas M. Link⁵, Jeffrey C. Lotz¹, Conor W. O'Neill¹, Aseem Sharma⁶ ISSLS Degenerative Spinal Phenotypes Group

¹Department of Orthopaedic Surgery, University of California, San Francisco, CA USA

²Faculty of Health Sciences & Western's Bone and Joint Institute, University of Western Ontario, London, Ontario Canada

³Department of Radiology, Hospital for Special Surgery, New York, NY USA

⁴Departments of Radiology, Neurosurgery and Health Services, and the Comparative Effectiveness, Cost and Outcomes Research Center, University of Washington, Seattle, WA USA

⁵Department of Radiology and Biomedical Imaging, University of California, San Francisco, CA USA

⁶Mallinckrodt Institute of Radiology, Washington University School of Medicine in St. Louis, MO USA

Abstract

Purpose: The positive association between low back pain and MRI evidence of vertebral endplate bone marrow lesions, often called Modic changes (MC), offers the exciting prospect of diagnosing a specific phenotype of chronic low back pain (LBP). However, imprecision in the reporting of MC has introduced substantial challenges, as variations in both imaging equipment and scanning parameters can impact conspicuity of MC. This review discusses key methodological factors that impact MC classification and recommends guidelines for more consistent MC reporting that will allow for better integration of research into this LBP phenotype.

Methods: Non-systematic literature review.

Results: The high diagnostic specificity of MC classification for a painful level contributes to the significant association observed between MC and LBP, whereas low and variable sensitivity underlies the between- and within-study variability in observed associations. Poor sensitivity may be owing to the presence of other pain generators, to the limited MRI resolution, and to the imperfect reliability of MC classification, which affects diagnostic sensitivity, and thus, influences the association between MC and LBP. Importantly, magnetic field strength and pulse sequence

parameters also impact detection of MC. Advances in pulse sequences may improve reliability and prove valuable for quantifying lesion severity.

Conclusions: Comparison of MC data between studies can be problematic. Various methodological factors impact detection and classification of MC, and the lack of reporting guidelines hinders interpretation and comparison of findings. Thus, it is critical to adopt imaging and reporting standards that codify acceptable methodological criteria.

Keywords

low back pain; Modic changes; bone marrow lesion; endplate damage; magnetic resonance imaging

What are Modic changes and why are they important?

Modic *et al.* [1] and de Roos *et al.* [2] separately observed that the bone marrow near damaged endplates in LBP patients had a distinct appearance when viewed with MRI. Compared to the adjacent normal bone marrow, the endplate lesions with active inflammation and fibrovascular replacement of the hematopoietic marrow appeared hypointense on T1-weighted images and hyperintense on T2-weighted images (type 1 changes; MC1); the endplate lesions with fatty replacement of the marrow appeared hyperintense on T1-weighted images and isointense to slightly hyperintense on conventional T2-weighted images (type 2 changes; MC2). A third type of endplate lesion, characterized by sclerotic subchondral bone, appeared hypointense on both T1- and T2-weighted images (type 3 changes; MC3) [3]. The etiology, risk factors, and management of Modic changes are the focus of a large body of work that has been reviewed elsewhere [4–8]. Briefly, it appears Modic changes may be either transient or permanent stages of a chronic pathologic process [8–11]; they associate with age and disc degeneration [12,5]; and they predominate in the lower lumbar spine (L4–S1)[12]. Perhaps most importantly, systematic reviews suggest a positive association between LBP and Modic changes [5,6].

Yet, the clinical relevance of Modic changes varies widely between studies [4,5,13], and controversies remain. Such inconsistencies may be due in part to misclassification and to imprecision in reporting of Modic changes. While some investigators follow standardized imaging and grading procedures, many utilize equipment, imaging sequences or reporting methods that can inadvertently misrepresent Modic changes. Also, the frequent omission of basic methodological details makes study comparisons difficult. As a result, discrepancies between study findings may reflect differences in imaging procedures, particularly between older and newer imaging technologies. This issue is becoming more apparent and more important as clinical interest in Modic changes grows. Thus, we recommend that investigators reporting Modic change data follow simple reporting guidelines (Box 1). These guidelines are not meant to constrain investigators to one particular methodology; instead, they are meant to ensure that the methodological details that impact the appearance and detection of Modic changes are clearly reported.

Identifying Modic changes: the high cost of misclassification

Misclassification prevents a clear understanding of the clinical significance of Modic changes and is a potentially important source of discrepancy between studies. The conventional MRI assessment of Modic changes amounts to a binary classification test. Owing to the difficulty of measuring test performance against histopathologic findings as ground truth, prior studies compared Modic classification to clinical pain diagnoses from provocative discography. Although the usefulness and safety of provocative discography are controversial [14,15], the accuracy of provocative discography can be quite high (specificity of 0.94 and false-positive rate of 6%) if performed using a low-pressure technique [16]. A review of the performance of Modic classification shows that high specificity may underlie the significant associations between Modic changes and LBP. Specifically, among six studies that reported the diagnostic performance of Modic classification for identifying a painful disc concordant with a positive discogram, the specificity was over 95% [17–21] in all but one [22] (Table 1). In other words, among patients with chronic LBP receiving discography, observation of a Modic change has a low false positive rate at a given spinal level. The specificity may be higher for MC1 than MC2 or MC3, although it is difficult to compare lesion types since some lesions show mixed elements of both MC1 and MC2 or MC 2 and MC 3 [23].

Conversely, the low and variable sensitivity of Modic classification for detecting discography-concordant pain may contribute to the weak associations with symptoms and to inter-study variability. For example, among the same six studies, the sensitivity of Modic classification for identifying a painful disc was less than 50%; moreover, it was highly variable between studies, ranging from 14% to 48% (Table 1). Thus, the absence of a Modic change is not sufficient for ruling out pain at a given spinal level.

The low and variable sensitivity can potentially reflect the existence of other pain generators. It may also be that poor sensitivity of Modic changes is related to limited spatial resolution, signal-to-noise, or lack of fat-saturation in the MRI exam. For example, one study reported that only 11% of fibrovascular marrow lesions and 62% of fatty marrow lesions identified on histologic sections were visible with a conventional imaging protocol [24]. As we discuss in the next section, this finding underscores the importance of reporting technical imaging parameters when interpreting the reported associations between Modic changes and LBP (Box 1).

Finally, the imperfect accuracy, or *reliability*, of Modic classification directly impacts diagnostic sensitivity and thereby contributes to inter-study variability. Jones *et al.* [25] reported good inter-rater reliability ($\kappa = 0.85$) for 5 raters, with repeatability for the individual raters ranging from $\kappa = 0.71$ –1.00. Other studies reported modest to good agreement between raters ($\kappa = 0.64$ –0.85 [26,12,9,27,28]; Table 2). The impact of rater agreement and its variation between studies highlights the need to measure and report inter-rater values (Box 1).

To emphasize the impact of imperfect reliability, consider the hypothetical effects of false negative classifications on the relationship between Modic changes and LBP (Figure 1). In

three scenarios with different sample sizes, rater disagreements resulting in false negative classifications were used to calculate inter-rater reliabilities and odds ratios (ORs). For a given sample size, ORs were highly sensitive to inter-rater reliability, especially for inter-rater agreements below $\kappa = 0.80$. Moreover, for the average inter-rater agreement of the six studies reported in Table 2 ($\kappa = 0.788$), associations between Modic changes and LBP ranged from a mean OR = 3.83 (CI 0.43–34.85, *not significant*) to a mean OR = 9.99 (CI 2.14–52.16, $p < 0.001$). Together these data suggest that imperfect reliability of Modic classification could be source of discrepancies between studies. Reliability is expected to improve by implementation of standardized imaging protocols and adoption of quantitative classification schemes that are based on a continuous rather than categorical measurement scale, thereby increasing sensitivity, and bringing much-needed clarity to the complex relationship between endplate bone marrow lesions and LBP.

The importance of MRI field strength, pulse sequence, and fat suppression

The ability for MRI to interrogate the endplate bone marrow is limited by instrument precision. MRI precision is influenced by field strength, acquisition matrix, and pulse sequence parameters because these factors directly affect image resolution, signal-to-noise, and contrast-to-noise ratios. In addition, field strength and fat saturation (for T2) also alter the T1 and T2 relaxation times and tissue visualization. All of these effects impact the appearance of water and fat signals, which are used to distinguish MC1 from MC2. For example, when older magnets with field strengths less than 1.0T are used to image the vertebral bone marrow, they tend to greatly mitigate chemical shift, susceptibility artifacts, and flow artifacts compared to 1.5T magnets. This improves the clarity of MC1, but makes it more difficult to distinguish MC2. Conversely, the 1.5T magnets make it more difficult to identify MC1 because marrow inhomogeneities are more pronounced compared to magnets with field strengths less than 1.0T [29]. Newer magnets with field strengths of 3.0T or higher have better conventional fat saturation than 1.5T magnets [30]. This reduces the tendency to overlook the marrow edema that occurs with MC1, although it is unclear whether this results in a clinically important difference. Variable field strengths also cause discrepancies in the observed frequency of different types of Modic changes. For example, whereas MC1 were 3–4 times more prevalent on 0.3T scanners, MC2 were identified twice as often with 1.5T scanners [29]. Generally speaking, the higher spatial resolution and greater dynamic range and signal- and contrast-to-noise ratios of 1.5T and 3.0T magnets, as compared to weaker magnets, make it easier to identify Modic changes. Thus, magnetic field strength should be reported in studies of Modic changes (Box 1).

Pulse sequence parameters also play a key role because they can dramatically alter the marrow signal. On T1-weighted images, fat is bright and fluid is dark. This makes fat conspicuous and helps radiologists identify boundaries between bone marrow lesions and normal marrow, which is useful for studying the bone marrow. Also, the short echo time of T1-weighted images provides a high intrinsic signal-to-noise ratio, which enhances anatomic detail. On T2-weighted images, fluid appears bright and fat is variable, depending on whether the images are acquired as conventional spin echo (moderate-to-high fat signal) vs. fast spin echo (high fat signal). It is important to recognize that the original studies of Modic *et al.* [1] and de Roos *et al.* [2] did not use fat-saturated T2-weighted sequences; hence, MC1

and MC2 both were defined as appearing hyperintense on T2-weighted MRI. However, many investigators now use fat-saturated T2-weighted sequences, which cause MC1 to appear hyperintense and MC2 to appear hypointense (Figure 2). Spectral fat suppression has the added benefit of increasing the dynamic range, and if combined with modest echo times of conventional T2-weighted sequences (60–80 ms), it can do so while preserving anatomic detail. The efficiency of fat suppression for Modic classification depends on the field homogeneity, which is better in the center of the magnet bore and improves with higher-order shimming. Spectral fat suppression can only be effectively used at field strengths greater than 1.0T. A number of fat suppression sequences are often used to supplement T1- and T2-weighted sequences, including chemical shift imaging sequences (such as Dixon) and Short T1 inversion recovery (STIR) sequences. STIR sequences can be used at low field strengths, and they provide more uniform fat saturation than fast spin echo sequences. Depending on the type of spectral fat suppression, STIR sequences can be less susceptible to magnetic field inhomogeneities. Authors should report the type of fat suppression used (Box 1).

Advances in the classification and detection of endplate bone marrow lesions

Recent advances in image analysis techniques may improve the reliability of Modic classification while at the same time provide quantitative methodologies that are needed to evaluate treatments. These advanced techniques are exploratory, with small studies performed mostly at single imaging centers. One strategy involves contouring the shape of the Modic change, which allows classification to be based on continuous measurements of size rather than on “present or absent” categorizations (Figure 3). For example, Wang *et al.* [28] reported outstanding intra- and inter-rater reliability values for three quantitative indices measured from manual contours of mid-sagittal slices: affected/unaffected vertebral area ratio (intra-rater: 0.96; inter-rater: 0.81), cerebrospinal fluid-adjusted mean signal intensity of the Modic change (intra-rater: 0.99; inter-rater: 0.92), and total signal intensity of the Modic change (intra-rater: 0.96; inter-rater: 0.92). Semi-automated contouring approaches [31] improve measurement efficiency and may further enhance the reliability of indices based on Modic change size and intensity.

A second strategy for quantitative and objective classification of endplate bone marrow lesions is based on the assessment of bone marrow lesion composition rather than on lesion size or structure. Measuring bone marrow composition is especially advantageous for evaluating lesion progression and for monitoring the response to treatments, where biochemical changes in the marrow compartment are likely to precede any visible structural changes. MR imaging based on chemical shift encoding-based water-fat imaging enables the spatially resolved assessment of bone marrow fat at trabecular sites with heterogeneous red marrow distribution [32]. For example, multi-echo gradient echo acquisitions with echo time steps that result in a water-fat phase difference different from 0 and 2π give robust water-fat separation, and accurate fat quantification is possible with techniques that incorporate a precalibrated multi-peak fat spectrum in the signal model [33]. These methods may be useful for quantifying the extent of marrow edema and fatty replacement that coincide with

MC1 and MC2 (Figure 3). Compared to single voxel spectroscopy, water-fat MRI sequences also facilitate measuring the spatial heterogeneity in marrow content, which is useful for classifying lesions that exhibit both edematous and fatty changes. Currently, classification of these “mixed-type” Modic changes [34,35] with T1 and T2-weighted sequences alone is highly subjective.

A third strategy involves diffusion-weighted imaging, which may help differentiate patients that have endplate bone marrow lesions with degenerative *vs.* infectious etiologies. Using diffusion-weighted MRI, Patel *et al.* [36] found that patients with well-margined, linear regions of high signal intensity situated within adjacent vertebrae at the interface between normal and abnormal bone marrow were predominantly infection-free; conversely, absence of this “claw sign” was associated with discitis/osteomyelitis. Those authors hypothesized that a gradual, progressive degenerative process results in a well-defined border between the normal and affected bone marrow, although this remains unconfirmed. Apparent diffusion coefficient (ADC) maps remove the T2 shine through of diffusion-weighted imaging and thereby provide quantifiable signal that is directly proportional to the diffusivity of water inside the tissue. ADC values can accurately distinguish between infectious spondylitis and MC1 [37,38]. One limitation with this approach is the sensitivity of image quality and ADC values to chosen strength and timing of the gradients (*b*-values) [38].

A complementary strategy for identifying spinal levels with endplate bone marrow lesions involves using pulse sequences that enhance visualization of endplate damage. Damage to the cartilage endplate and subchondral bone is believed to be an important factor in the etiology of endplate bone marrow lesions because damage promotes cross-talk between inflammatory factors expressed in the disc and the quiescent bone marrow [8,39]. However, conventional T2-weighted sequences used in the spine (echo time = 60–80 ms) are unable to show the cartilage endplate, because the cartilage has short T2 values, and thus, its signal is not captured in sequences with long echo times. Newer sequences such as ultra-short time-to-echo [40–42] (UTE) can overcome this limitation. For example, Law *et al.* [42] first demonstrated the feasibility of assessing cartilage endplate integrity using UTE MRI, which simultaneously improves visualization of endplate morphology (Figure 3). In addition to assessing the morphology of the cartilage endplate, UTE also has the capability of assessing its biochemical composition [41], which could potentially be used to assess early degeneration. In the future, quantitative measurements of cartilage endplate degeneration and damage may provide a more comprehensive evaluation of endplate bone marrow lesions [43].

Summary and recommendations

Comparison of Modic change data between studies can be problematic. Depending on various technical factors such as imaging sequence and magnetic field strength, MC1 can be detected at greatly varying frequencies and MC2 can appear hyper- rather than hypointense. These variations may result in inconsistencies in the apparent relationship between Modic changes and LBP. Problems with comparability can be even greater in longitudinal studies where the comparison is between images acquired at baseline with older MR units that have lower signal-to-noise ratios and images acquired at follow-up with newer MR units with

improved signal-to-noise ratios and fat-saturated sequences. Even when identical sequences are used with the same imager at all time points, the subjective and categorical nature of Modic classification limits inter-rater and intra-rater reliability. All of these factors affect the reported associations between Modic changes and LBP and may underlie the wide variability between studies.

Overall, as research and clinical interest in Modic changes increases, and as stronger magnets and newer sequences gradually replace older ones, it will be critical to appreciate how technological advancements can influence reported clinical associations and comparability of study results. It will also be necessary to ensure the methodological details that accompany Modic change data are sufficiently documented to understand which technologies and techniques were used for image acquisition and analysis. This technical information is essential for consistent interpretation of Modic change data. Therefore, it is critical to adopt imaging and reporting standards that codify acceptable methodological information that is necessary to accompany Modic change data (Box 1).

Finally, while qualitative Modic classification with T1 and T2 sequences is currently the norm, novel quantitative methods show potential for assessing the severity of changes in marrow composition and for characterizing endplate structure and function in a more objective and operator- and scanner-independent manner. These developments may form the basis for more accurate future classification systems.

Acknowledgements:

The Degenerative Spinal Phenotypes Group of the International Society for the Study of the Lumbar Spine was co-chaired by Dino Samartzis, Joseph Assheuer, Bradley Weiner and Michele Battié, with the goal of investigating and promoting common concepts, nomenclature, definitions, and core measures for degenerative spinal phenotypes to advance knowledge in the field. The imaging data accompanying this commentary were acquired with support from the National Institutes of Health, grants R01AR070198 (AJF), R01AR063705 (JCL), and P30AR066262 (JCL). The authors acknowledge Dave Glidden, Ph.D. for providing biostatistics consultation.

References

1. Modic MT, Steinberg PM, Ross JS, Masaryk TJ, Carter JR (1988) Degenerative disk disease: assessment of changes in vertebral body marrow with MR imaging. *Radiology* 166 (1 Pt 1):193–199 [PubMed: 3336678]
2. de Roos A, Kressel H, Spritzer C, Dalinka M (1987) MR imaging of marrow changes adjacent to end plates in degenerative lumbar disk disease. *AJR Am J Roentgenol* 149 (3):531–534. doi:10.2214/ajr.149.3.531 [PubMed: 3497539]
3. Modic MT, Masaryk TJ, Ross JS, Carter JR (1988) Imaging of degenerative disk disease. *Radiology* 168 (1):177–186. doi:10.1148/radiology.168.1.3289089 [PubMed: 3289089]
4. Jensen RK, Leboeuf-Yde C (2011) Is the presence of modic changes associated with the outcomes of different treatments? A systematic critical review. *BMC Musculoskelet Disord* 12:183. doi:10.1186/1471-2474-12-183 [PubMed: 21831312]
5. Jensen TS, Karppinen J, Sorensen JS, Niinimäki J, Leboeuf-Yde C (2008) Vertebral endplate signal changes (Modic change): a systematic literature review of prevalence and association with non-specific low back pain. *Eur Spine J* 17 (11):1407–1422. doi:10.1007/s00586-008-0770-2 [PubMed: 18787845]
6. Zhang YH, Zhao CQ, Jiang LS, Chen XD, Dai LY (2008) Modic changes: a systematic review of the literature. *Eur Spine J* 17 (10):1289–1299. doi:10.1007/s00586-008-0758-y [PubMed: 18751740]

7. Rahme R, Moussa R (2008) The modic vertebral endplate and marrow changes: pathologic significance and relation to low back pain and segmental instability of the lumbar spine. *AJNR Am J Neuroradiol* 29 (5):838–842. doi:ajnr.A0925 [pii] 10.3174/ajnr.A0925 [PubMed: 18272564]
8. Dudli S, Fields AJ, Samartzis D, Karppinen J, Lotz JC (2016) Pathobiology of Modic changes. *Eur Spine J* 25 (11):3723–3734. doi:10.1007/s00586-016-4459-7 [PubMed: 26914098]
9. Kuisma M, Karppinen J, Niinimäki J, Kurunlahti M, Haapea M, Vanharanta H, Tervonen O (2006) A three-year follow-up of lumbar spine endplate (Modic) changes. *Spine (Phila Pa 1976)* 31 (15):1714–1718. doi:10.1097/01.brs.0000224167.18483.14 [PubMed: 16816768]
10. Hutton MJ, Bayer JH, Powell JM (2011) Modic vertebral body changes: the natural history as assessed by consecutive magnetic resonance imaging. *Spine (Phila Pa 1976)* 36 (26):2304–2307. doi:10.1097/BRS.0b013e31821604b6 [PubMed: 21358572]
11. Farshad-Amacker NA, Hughes A, Herzog RJ, Seifert B, Farshad M (2017) The intervertebral disc, the endplates and the vertebral bone marrow as a unit in the process of degeneration. *Eur Radiol* 27 (6):2507–2520. doi:10.1007/s00330-016-4584-z [PubMed: 27709276]
12. Karchevsky M, Schweitzer ME, Carrino JA, Zoga A, Montgomery D, Parker L (2005) Reactive endplate marrow changes: a systematic morphologic and epidemiologic evaluation. *Skeletal Radiol* 34 (3):125–129. doi:10.1007/s00256-004-0886-3 [PubMed: 15647940]
13. Laustsen AF, Bech-Azeddine R (2016) Do Modic changes have an impact on clinical outcome in lumbar spine surgery? A systematic literature review. *Eur Spine J* 25 (11):3735–3745. doi:10.1007/s00586-016-4609-y [PubMed: 27177469]
14. Carragee EJ, Don AS, Hurwitz EL, Cuellar JM, Carrino JA, Herzog R (2009) 2009 ISSLS Prize Winner: Does discography cause accelerated progression of degeneration changes in the lumbar disc: a ten-year matched cohort study. *Spine (Phila Pa 1976)* 34 (21):2338–2345. doi:10.1097/BRS.0b013e3181ab5432 [PubMed: 19755936]
15. Carragee EJ, Lincoln T, Parmar VS, Alamin T (2006) A gold standard evaluation of the “discogenic pain” diagnosis as determined by provocative discography. *Spine (Phila Pa 1976)* 31 (18):2115–2123. doi:10.1097/01.brs.0000231436.30262.dd 00007632-200608150-00018 [pii] [PubMed: 16915099]
16. Wolfer LR, Derby R, Lee JE, Lee SH (2008) Systematic review of lumbar provocation discography in asymptomatic subjects with a meta-analysis of false-positive rates. *Pain Physician* 11 (4):513–538 [PubMed: 18690280]
17. Braithwaite I, White J, Saifuddin A, Renton P, Taylor BA (1998) Vertebral end-plate (Modic) changes on lumbar spine MRI: correlation with pain reproduction at lumbar discography. *Eur Spine J* 7 (5):363–368 [PubMed: 9840468]
18. Ito M, Incorvaia KM, Yu SF, Fredrickson BE, Yuan HA, Rosenbaum AE (1998) Predictive signs of discogenic lumbar pain on magnetic resonance imaging with discography correlation. *Spine* 23 (11):1252–1258; discussion 1259–1260 [PubMed: 9636979]
19. O’Neill C, Kurgansky M, Kaiser J, Lau W (2008) Accuracy of MRI for diagnosis of discogenic pain. *Pain Physician* 11 (3):311–326 [PubMed: 18523502]
20. Thompson KJ, Dagher AP, Eckel TS, Clark M, Reinig JW (2009) Modic changes on MR images as studied with provocative diskography: clinical relevance—a retrospective study of 2457 disks. *Radiology* 250 (3):849–855. doi:250/3/849 [pii] 10.1148/radiol.2503080474 [PubMed: 19244050]
21. Weishaupt D, Zanetti M, Hodler J, Min K, Fuchs B, Pfirrmann CW, Boos N (2001) Painful lumbar disk derangement: relevance of endplate abnormalities at MR imaging. *Radiology* 218 (2):420–427 [PubMed: 11161156]
22. Kokkonen SM, Kurunlahti M, Tervonen O, Ilkko E, Vanharanta H (2002) Endplate degeneration observed on magnetic resonance imaging of the lumbar spine: correlation with pain provocation and disc changes observed on computed tomography diskography. *Spine* 27 (20):2274–2278. doi:10.1097/01.BRS.0000029248.13169.51 [PubMed: 12394906]
23. Wang Y, Videman T, Battie MC (2012) Lumbar vertebral endplate lesions: prevalence, classification, and association with age. *Spine (Phila Pa 1976)* 37 (17):1432–1439. doi:10.1097/BRS.0b013e31824dd20a [PubMed: 22333959]

24. Fields AJ, Liebenberg EC, Lotz JC (2014) Innervation of pathologies in the lumbar vertebral end plate and intervertebral disc. *Spine J* 14 (3):513–521. doi:S1529-9430(13)01197-2 [pii] 10.1016/j.spinee.2013.06.075 [PubMed: 24139753]
25. Jones A, Clarke A, Freeman BJ, Lam KS, Grevitt MP (2005) The Modic classification: inter- and intraobserver error in clinical practice. *Spine (Phila Pa 1976)* 30 (16):1867–1869 [PubMed: 16103857]
26. Chung CB, Vande Berg BC, Tavernier T, Cotten A, Laredo JD, Vallee C, Malghem J (2004) End plate marrow changes in the asymptomatic lumbosacral spine: frequency, distribution and correlation with age and degenerative changes. *Skeletal Radiol* 33 (7):399–404. doi:10.1007/s00256-004-0780-z [PubMed: 15138721]
27. Peterson CK, Gatterman B, Carter JC, Humphreys BK, Weibel A (2007) Inter- and intraexaminer reliability in identifying and classifying degenerative marrow (Modic) changes on lumbar spine magnetic resonance scans. *J Manipulative Physiol Ther* 30 (2):85–90. doi:10.1016/j.jmpt.2006.12.001 [PubMed: 17320728]
28. Wang Y, Videman T, Niemelainen R, Battie MC (2011) Quantitative measures of modic changes in lumbar spine magnetic resonance imaging: intra- and inter-rater reliability. *Spine (Phila Pa 1976)* 36 (15):1236–1243. doi:10.1097/BRS.0b013e3181ecf283 [PubMed: 21217445]
29. Bendix T, Sorensen JS, Henriksson GA, Bolstad JE, Narvestad EK, Jensen TS (2012) Lumbar modic changes—a comparison between findings at low- and high-field magnetic resonance imaging. *Spine (Phila Pa 1976)* 37 (20):1756–1762. doi:10.1097/BRS.0b013e318257ffce [PubMed: 22498993]
30. Zhao J, Krug R, Xu D, Lu Y, Link TM (2009) MRI of the spine: image quality and normal-neoplastic bone marrow contrast at 3 T versus 1.5 T. *AJR Am J Roentgenol* 192 (4):873–880. doi:10.2214/AJR.08.1750 [PubMed: 19304689]
31. Kazakia GJ, Fields AJ, Liebenberg EC, Gary KN, Carballido-Gamio J, Han M, Krug R, Lotz JC (2016) Magnetic resonance imaging accurately measures vertebral marrow lesion size. Paper presented at the 43rd International Society for the Study of the Lumbar Spine Annual Meeting,
32. Karampinos DC, Melkus G, Baum T, Bauer JS, Rummeny EJ, Krug R (2014) Bone marrow fat quantification in the presence of trabecular bone: initial comparison between water-fat imaging and single-voxel MRS. *Magn Reson Med* 71 (3):1158–1165. doi:10.1002/mrm.24775 [PubMed: 23657998]
33. Karampinos DC, Ruschke S, Dieckmeyer M, Diefenbach M, Franz D, Gersing AS, Krug R, Baum T (2018) Quantitative MRI and spectroscopy of bone marrow. *J Magn Reson Imaging* 47 (2):332–353. doi:10.1002/jmri.25769 [PubMed: 28570033]
34. Fayad F, Lefevre-Colau MM, Drape JL, Feydy A, Chemla N, Quintero N, Rannou F, Poiradeau S, Fermanian J, Revel M (2009) Reliability of a modified Modic classification of bone marrow changes in lumbar spine MRI. *Joint Bone Spine* 76 (3):286–289. doi:10.1016/j.jbspin.2008.09.012 [PubMed: 19119042]
35. Kuisma M, Karppinen J, Haapea M, Lammentausta E, Niinimäki J, Tervonen O (2009) Modic changes in vertebral endplates: a comparison of MR imaging and multislice CT. *Skeletal Radiol* 38 (2):141–147. doi:10.1007/s00256-008-0590-9 [PubMed: 18807028]
36. Patel KB, Poplawski MM, Pawha PS, Naidich TP, Tanenbaum LN (2014) Diffusion-weighted MRI “claw sign” improves differentiation of infectious from degenerative modic type 1 signal changes of the spine. *AJNR Am J Neuroradiol* 35 (8):1647–1652. doi:10.3174/ajnr.A3948 [PubMed: 24742801]
37. Oztekin O, Calli C, Kitis O, Adibelli ZH, Eren CS, Apaydin M, Zileli M, Yurtseven T (2010) Reliability of diffusion weighted MR imaging in differentiating degenerative and infectious end plate changes. *Radiol Oncol* 44 (2):97–102. doi:10.2478/v10019-010-0006-z [PubMed: 22933898]
38. Daghighi MH, Poureisa M, Safarpour M, Behzadmehr R, Fouladi DF, Meshkini A, Varshochi M, Kiani Nazarlou A (2016) Diffusion-weighted magnetic resonance imaging in differentiating acute infectious spondylitis from degenerative Modic type 1 change; the role of b-value, apparent diffusion coefficient, claw sign and amorphous increased signal. *Br J Radiol* 89 (1066):20150152. doi:10.1259/bjr.20150152 [PubMed: 27452260]

39. Dudli S, Sing DC, Hu SS, Berven SH, Burch S, Deviren V, Cheng I, Tay BK, Alamin TF, Ith MA, Pietras EM, Lotz JC (2017) ISSLS PRIZE IN BASIC SCIENCE 2017: Intervertebral disc/bone marrow cross-talk with Modic changes. *Eur Spine J*. doi:10.1007/s00586-017-4955-4
40. Bae WC, Statum S, Zhang Z, Yamaguchi T, Wolfson T, Gamst AC, Du J, Bydder GM, Masuda K, Chung CB (2013) Morphology of the cartilaginous endplates in human intervertebral disks with ultrashort echo time MR imaging. *Radiology* 266 (2):564–574. doi:10.1148/radiol.12121181 [pii] 10.1148/radiol.12121181 [PubMed: 23192776]
41. Fields AJ, Han M, Krug R, Lotz JC (2015) Cartilaginous end plates: quantitative MR imaging with very short echo times-orientation dependence and correlation with biochemical composition. *Radiology* 274 (2):482–489. doi:10.1148/radiol.14141082 [PubMed: 25302832]
42. Law T, Anthony MP, Chan Q, Samartzis D, Kim M, Cheung KM, Khong PL (2013) Ultrashort time-to-echo MRI of the cartilaginous endplate: Technique and association with intervertebral disc degeneration. *J Med Imaging Radiat Oncol* 57 (4):427–434. doi:10.1111/1754-9485.12041 [PubMed: 23870338]
43. Bailey JF, Fields AJ, Ballatori A, Cohen D, Jain D, Coughlin D, O’Neill C, McCormick Z, Han M, Krug R, Demir-Deviren S, Lotz JC (2019) The Relationship Between Endplate Pathology and Patient-Reported Symptoms for Chronic Low Back Pain Depends on Lumbar Paraspinal Muscle Quality. *Spine (Phila Pa 1976)*. doi:10.1097/BRS.0000000000003035

Box 1.**SUGGESTED GUIDELINES FOR PRESENTATION OF MODIC CHANGE DATA****MR unit:**

Identify the MRI instrument and magnetic field strength, plus any surface coils or specialized tables used to collect and amplify signal. Include model number and manufacturer. For longitudinal or multi-center studies, use scanners with the same field strength.

Sequences:

Specify the T1- and T2-weighted sequence parameters, including the type of spin echo (*i.e.* fast *vs.* conventional), repetition time/echo time, field of view, matrix size, slice thickness/spacing, number of echoes, and the type of fat suppression that was applied (T2 only). If fat suppression was used, the Modic classification for type 2 changes should be defined as having hyperintense signal on T1-weighted images and hypointense on fat-saturated T2-weighted images. If additional sequences, *e.g.* STIR/Dixon, will be used for classifying Modic changes, these sequences should be reported in addition to the T1- and T2-weighted sequences.

Image evaluation:

Describe which image slices were rated and which levels were evaluated.

Rater agreement:

Report the inter-rater and intra-rater kappa statistics for categorical Modic classification. Also, report the inter-rater and intra-rater intraclass correlation coefficients (ICC) if using quantitative measurements, *e.g.* lesion size, cerebrospinal fluid-normalized intensity, *etc.*

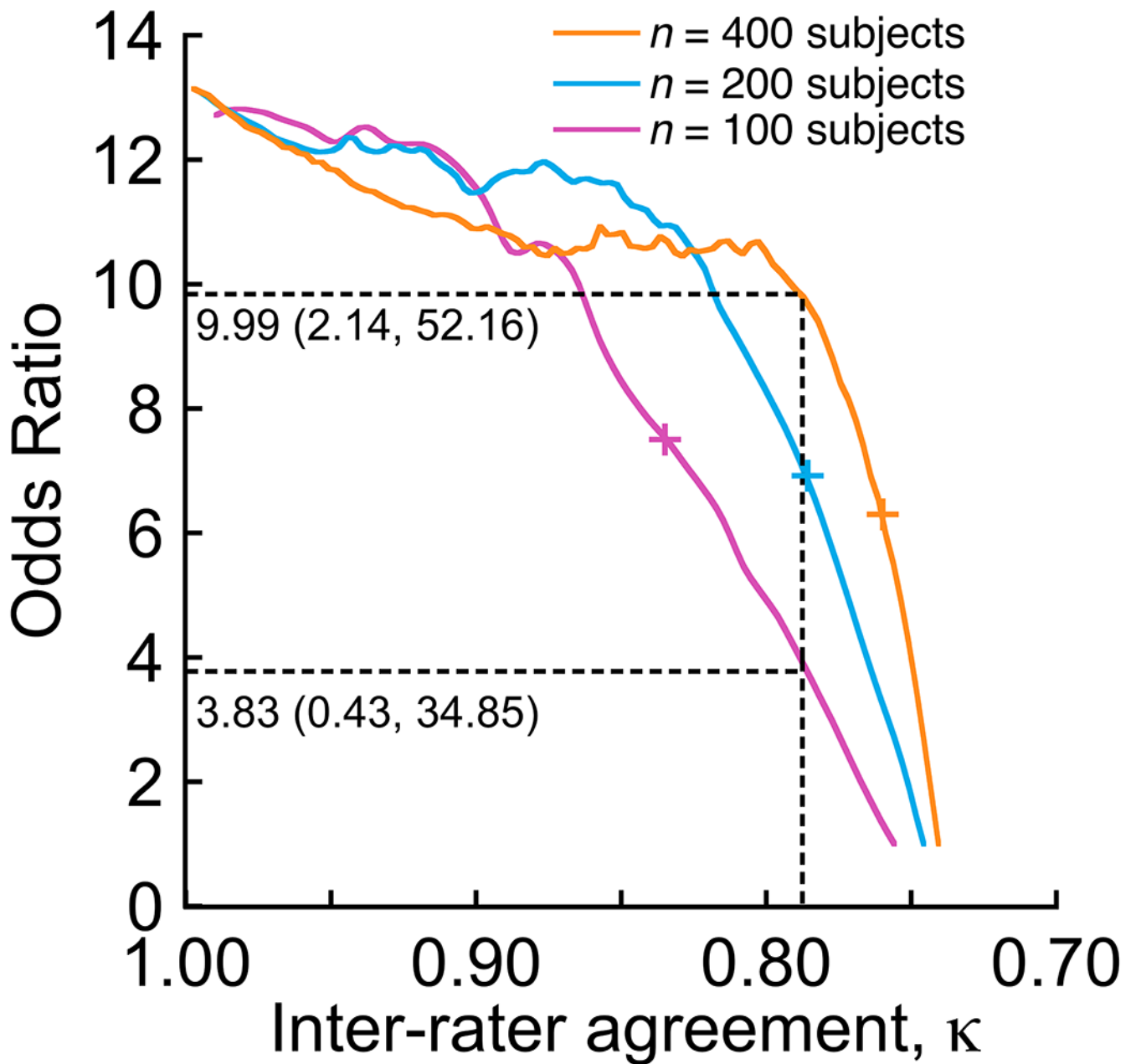


Figure 1.

Hypothetical scenarios show how the reliability of Modic classification, or inter-rater agreement, impacts the relationship between Modic changes and LBP. In each scenario, the actual prevalence of Modic changes in LBP patients (46% [5]) and asymptomatic controls (6% [5]) was simulated, and increasing numbers of false negatives were randomly assigned to the groups (group size = $n/2$). After each false negative assignment, the inter-rater agreement (κ) and odds ratio was calculated. This simulation was repeated 100 times in order to consider cases where the false negative assignments randomly affected members of each group. The mean odds ratio was calculated for 100 repeat simulations. Plus signs (+) for each scenario indicate the cutoff value of κ that produced a significant odds ratio for

each scenario: 100 subjects, $\kappa = 0.837$, OR = 7.69 (CI 1.09–56.81); 200 subjects, $\kappa = 0.786$, OR = 7.04 (CI 1.03–51.03); 400 subjects, $\kappa = 0.761$, OR = 7.27 (CI 1.04–53.53). For $\kappa = 0.788$ (vertical dashed line), which was the mean reliability of Modic classification from the six studies summarized in Table 2, the association between Modic changes and LBP ranged from a mean OR = 3.83 (CI 0.43–34.85, *not significant*) to a mean OR = 9.99 (CI 2.14–52.16, $p < 0.001$).

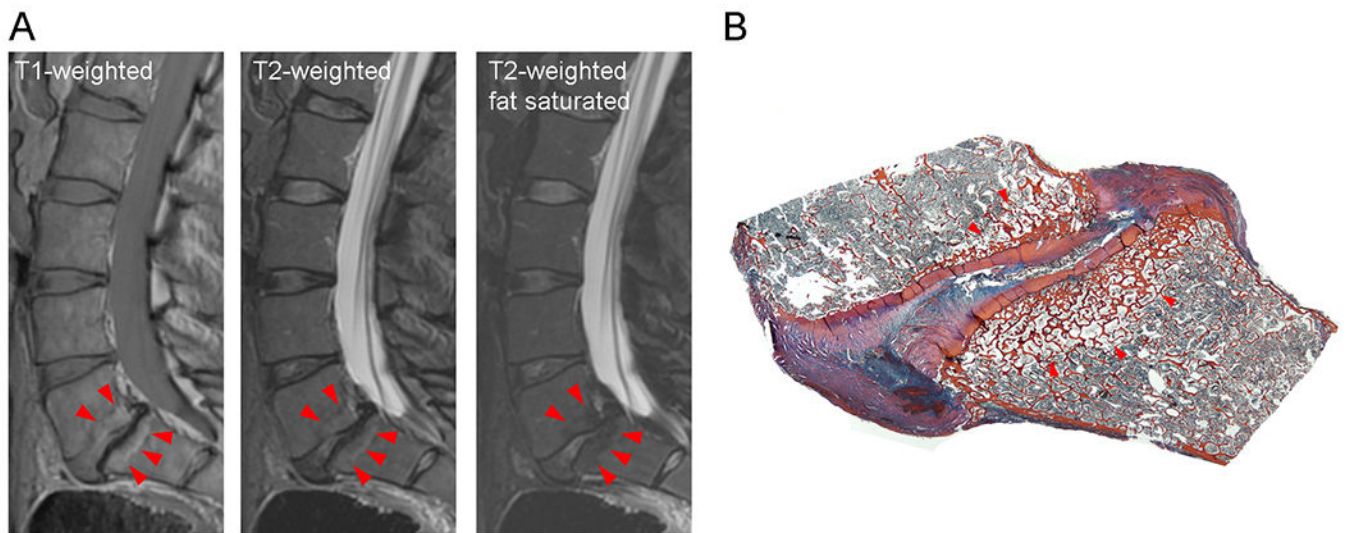


Figure 2.

(A) Type 2 Modic changes seen on sagittal T1- and T2-weighted images of a fresh cadaveric lumbar spine. Use of fat suppression reverses the hyperintense signal on the T2-weighted images. Images were acquired at 3.0T with a fast spin-echo sagittal T2 sequence (repetition time msec/echo time msec 4282/85; 27-cm field of view; 3-mm slice thickness) and a sagittal T1 sequence (556/14, 27-cm field of view, 3-mm slice thickness). (B) Matching sagittal histologic section of L5-S1 level indicating endplate bone marrow lesion with fatty replacement of the hematopoietic elements. Heidenhain tri-chrome stain.

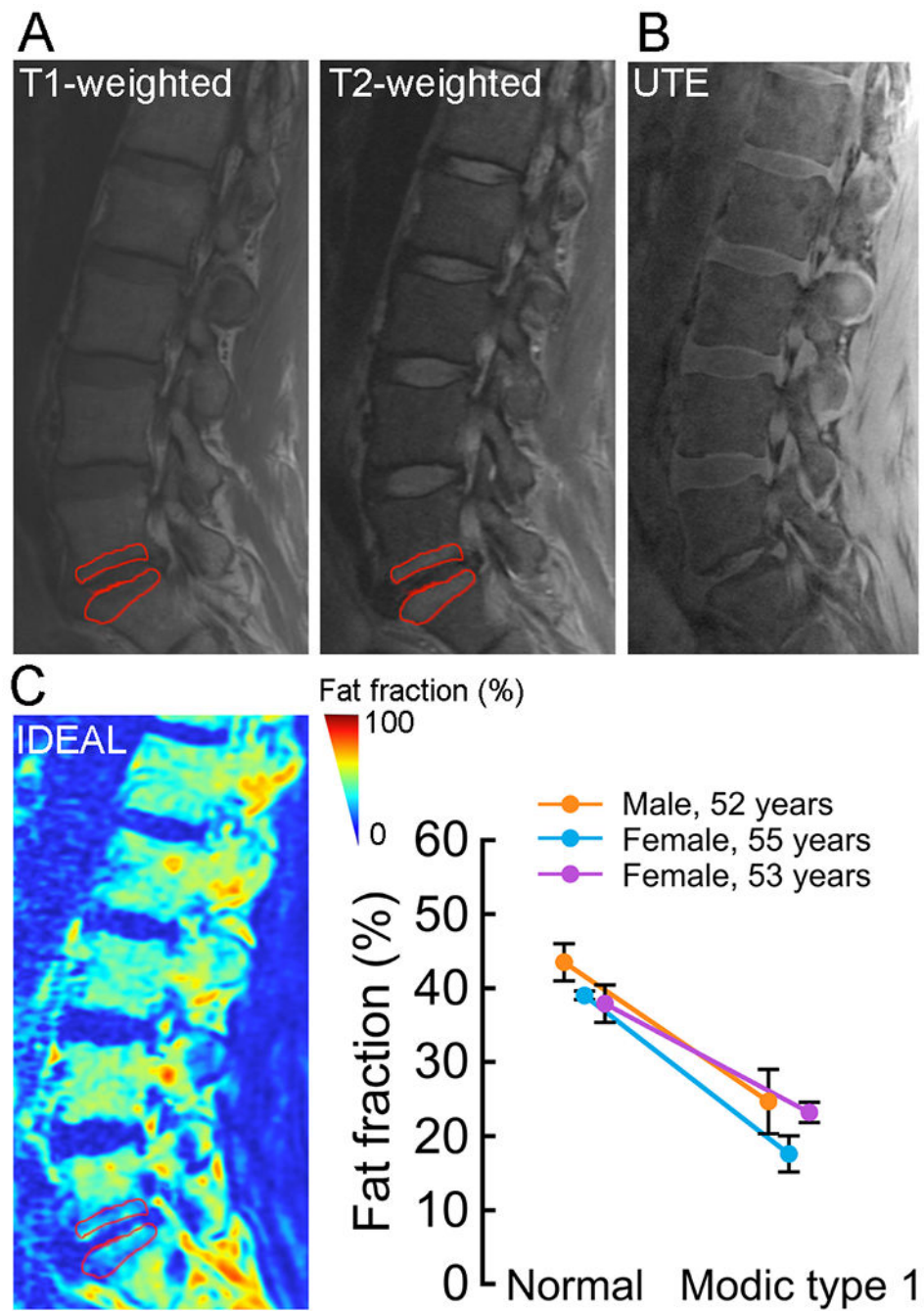


Figure 3.

(A) Type 1 Modic change contoured on sagittal T1- and fat saturated T2-weighted images from a 52-year-old male subject presenting with chronic LBP. Contouring enables measurement of the size and relative intensity of the Modic change. Images were acquired with a fast spin-echo sagittal T2 sequence (repetition time msec/echo time msec 4933/66; 26-cm field of view; 4-mm slice thickness) and a sagittal T1 sequence (694/15, 26-cm field of view, 4-mm slice thickness). (B) Matching sagittal UTE image showing disruptions in the continuity of the cartilage endplate adjacent to the Modic change. In addition to providing an

assessment of cartilage endplate integrity, UTE shows good marrow contrast. Images were acquired with a UTE sequence that combines a nonselective hard pulse and 3D radial acquisition (10/0.236, 19-degree flip angle, 15-cm field of view, 1.5-mm slice thickness). (C) Matching sagittal fat fraction map from water-fat MRI showing reduced fat fraction in the region corresponding to the contour in (A). Compared to site-matched normal regions in the adjacent, non-affected vertebrae, regions with type 1 Modic changes had lower fat fraction, suggesting that water-fat MRI is sensitive to the fibrovascular replacement of the normal bone marrow that occurs with this type of endplate bone marrow lesion. Data are shown for three subjects presenting with chronic LBP, and fat fractions are shown as mean \pm SD for three ROIs (3-mm diameter) per region. Images were acquired with a 3D spoiled gradient recalled sequence (10/1.3, 3-degree flip angle, 22-cm field of view, 4-mm slice thickness) with six echoes and iterative decomposition of water and fat with echo asymmetry and least squares estimation (IDEAL). All images were acquired at 3.0T.

Table 1.

Summary of diagnostic performance of Modic classification for discography-concordant pain

Author	Year	Levels	Age (years)	Field	Specificity (%)	Sensitivity (%)
Braithwaite ¹⁷	1998	290	42	0.5T, 1.5T	97	23
Ito ¹⁸	1998	101	37	1.5T	95	22
Kokkonen ²²	2002	103	40	<i>NR</i>	64	41
O'Neill ¹⁹	2008	460	43	<i>NR</i>	98	14
Thompson ²⁰	2009	2457	43	<i>NR</i>	95	26
Weishaup ²¹	2001	116	42	1.0T	96	48

NR – not reported

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Summary of the inter- and intra-rater repeatability of Modic classification

Author	Year	Subjects	Field strength	κ , inter-rater	κ , intra-rater	Raters
Jones ²⁵	2005	50	<i>NR</i>	0.85	0.71-1.00	5
Chung ²⁶	2004	59	0.5T-1.5T	0.83	<i>NR</i>	2
Karchevsky ¹²	2005	100	1.5T	0.81	<i>NR</i>	3
Kuisma ⁹	2006	60	1.5T	0.64	0.90	2
Peterson ²⁷	2007	51	0.6T	0.81	0.70-0.86	2
Wang ²⁸	2011	83	1.5T	0.79	0.88	2

NR – not reported

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript