Interventional Pain Medicine 1 (2022) 100092

Contents lists available at ScienceDirect

Interventional Pain Medicine

journal homepage: www.journals.elsevier.com/interventional-pain-medicine

# The interrater reliability of the novel Udby classification of Modic Changes: A first estimate

David Sherwood [a,*], R. Sterling Haring [b], Benjamin Gill [c], Scott Miller [b], Adam Epps [b], Oksana Zhivotenko [b], Samir Khan [b], Theodora L. Swenson [b], James Gardner [b], Christian Roehmer [b], Dann Martin [b], David J. Kennedy [b], Byron Schneider [b], Michael Modic [b], Peter Udby [d,e]

[a] University Health, Lakewood Medical Center, Department of Orthopedics, 7900 Lee's Summit Rd, Kansas City, MO, 64139, USA
[b] Vanderbilt University Medical Center, 2201 Children's Way, Suite 1318, Nashville, TN, 37212, USA
[c] University of Missouri, Columbia, 1 Hospital Drive DC046.00, Columbia, MO, 65212, USA
[d] Spine Unit, Department of Orthopedic Surgery, Zealand University Hospital, 4600, Koege, Denmark
[e] Spine Surgery and Research, Spine Center of Southern Denmark, Lillebaelt Hospital, 5500, Middelfart, Denmark

ARTICLE INFO

*Keywords:*
Low back pain
Basivertebral nerve ablation
Modic changes
Lumbar spine
Kappa
Udby

ABSTRACT

*Background:* Modic change grading is heterogeneous, inconsistent, and lacks a single nomenclature across the published literature. A new method of Modic change classification has been established by Dr. Peter Udby which hopes to unify how Modic changes are classified while also adding grading of the cranial/caudal extent of the Modic change across the vertebral body from the respective endplate involved to best capture the clinically relevant information of Modic changes.

*Methods:* Twenty magnetic resonance images of potential basivertebral nerve ablation candidates were independently reviewed by two board-certified and fellowship trained neuroradiologist and two board-certified and fellowship-trained interventional spine physiatrists for the presence and characterization of Modic changes using the newly described Udby classification. 100% agreement of all four reviewers of Modic change presence, Type, and Udby classification was required to be classified as agreement. There were 480 total data points each with 10 unique choices to compare across the four independent reviewers.

*Results:* The kappa value of their agreement was 0.5899 (95% CI 0.4860–0.6939).

*Conclusion:* This study, requiring unanimous agreement between 4 physicians in application of the Udby classification, demonstrated an interrater reliability score of 0.5899 (95% CI 0.4860–0.6939). While this figure provides a first estimate, larger scale research is necessary before definitive claims regarding the interrater agreement validity of the Udby characterization system may be made.

## 1. Introduction

A newly proposed characterization of Modic changes was published by lead author Peter Udby [1]. Beyond defining the presence and type of Modic changes described by de Roos and Modic, the new system classifies the physical extent of vertebral involvement (A = <25%, B = 25–50%, C = >50%) via sagittal evaluation on magnetic resonance imaging (MRI) [1–3]. All sagittal slices at each vertebral body are reviewed. The sagittal slice with the Modic change of the largest cephalad extent if measuring from the inferior endplate or caudal extent if measuring from the superior endplate is then graded. The preferred nomenclature for this system for a Type 1 Modic change at L5 with >50% vertical involvement on sagittal imaging would be: L5 MC 1 Grade C. Fig. 1 provides a visualization of the Udby method of characterizing Modic changes.

Modic changes at a given level are correlated with reproduction of a patient's low back pain by provocation discography [4]. Moreover, it has been reported that the vertical extent of the intracorporal Modic change may also have clinical implication [5,6]. The new classification system hopes to offer a more granular characterization of Modic changes to thus further delineate clinical implications for researchers and physicians.

The interrater reliability (Kappa) of Modic changes has been previously established typically ranging from moderate (0.6–0.79) to strong

* Corresponding author. University Health, Lakewood Medical Center, Department of Orthopedics, 7900 Lee's Summit Rd, Kansas City, MO, 64139, USA.
*E-mail address:* dhs988@gmail.com (D. Sherwood).

https://doi.org/10.1016/j.inpm.2022.100092
Received 25 February 2022; Received in revised form 12 April 2022; Accepted 14 April 2022

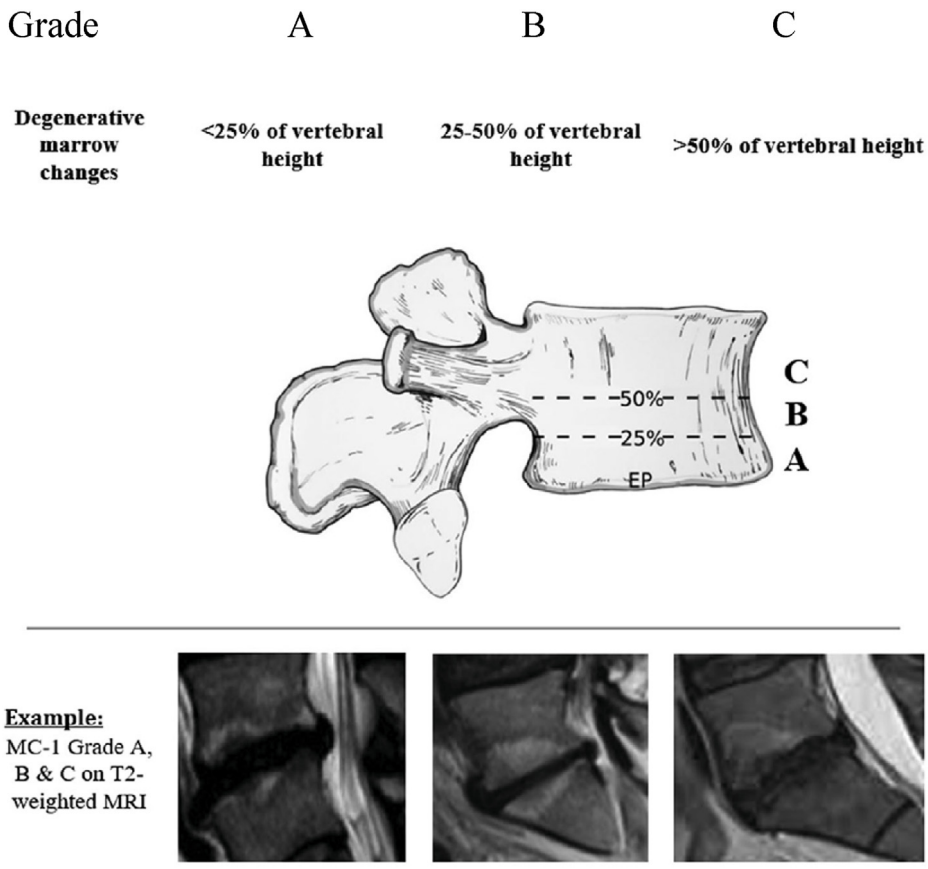


**Fig. 1.** *Udby Classification Example*[1]
An illustrative version of the sagittal grading scale. Measurement begins from the endplate (EP) and is determined by the sagittal slice with the most caudal or cephalad extent of involvement. <25% = A, 25–50% = B, >50% sagittal involvement = C. The imaging studies should be interpreted from as Left: Grade A <25%, Middle: Grade B 25–50%, Right: Grade C >50%.

(0.8–0.9) [5,7–13]. However, there is heterogeneity in these results subject to differing definitions of Modic changes, grading criteria, cohorts evaluated, pedigree of reviewer, and field strength of the MRI magnet [1,4,14]. Such heterogeneity prompts the need for creation of a unified classification system. A unified classification system would address issues related to nomenclature and allow for greater interstudy comparisons. For a new classification system to serve a purpose in an academic and clinical setting, the grading score needs to be unambiguous and valid. The aim of this study was to evaluate the interrater reliability of the Udby classification of Modic changes in a clinically relevant cohort.

## 2. Methods

This is a retrospective study of a single patient cohort from an academic medical center. Institutional Review Board (#210039) approval was obtained. Patient charts from four fellowship-trained spine physiatrists at a large academic spine center for records dated between January 1, 2019, and January 1, 2020 were identified using the diagnostic codes found in Table 1 from the International Classification of Diseases, Tenth Revision (ICD-10) [15]. The selection for the diagnostic codes was a collaborative effort between the primary investigator (PI) and the VUMC Research Derivative to maximize the capture rate of patients with isolated axial low back pain without radicular symptoms. Multiple ICD-10 combinations were queried until the largest sample size was identified for the focus of the study.

**Table 1**
Combination of diagnostic ICD-10 codes used to obtain initial cohort.

- ICD-10
  - Low back pain (M54.5)
  - Spondylopathy in diseases classified elsewhere, lumbosacral region (M49.87)
  - Spondylopathy in diseases classified elsewhere, lumbar region (M49.86)
  - Other Spondylosis (M47.8)
  - Spondylosis, unspecified (M47.9)
  - Spondylosis without myelopathy or radiculopathy (M47.816)

Table 1 includes the complete combination of search terms which were submitted to the VUMC Research Derivatives Database for chart retrieval.

Patient records with the relevant diagnoses were then cross-referenced with Current Procedural Terminology (CPT®) codes for MRI of the lumbar spine, including CPT 72148 and 72158. Patient records were excluded if no lumbar MRI code was ordered between January 1, 2019, and January 1, 2020.

The initial cohort was obtained via Vanderbilt University Medical Center's (VUMC) Research Derivatives database (RDD). The RDD is a database of clinical and related data derived from the VUMC clinical systems and restructured for research. Data is repurposed from VUMC's enterprise data warehouse, which includes data from StarPanel, Vanderbilt Perioperative Information Management System, and Operating Room Management Information System, EPIC, Medipac, and Horizon Expert Orders. The medical record number and other person identifiers are preserved within the RDD. Data is subject to atlerations and updates made by medical billers after initial coding is submitted, such changes to coding are updated monthly. Data types include reimbursement codes, clinical notes and documentation, nursing records, medication data, laboratory data, encounter and visit data, among others. Output may include structured data points, such as ICD-9 or -10 codes and encounter dates, semi-structured data such as laboratory tests and results, or unstructured data such as physician progress reports. The database is maintained by the Office of Research Informatics [16].

This study was nested within another study that evaluated the prevalence of basivertebral nerve ablation (BVNA) candidates, which is particularly relevant to Modic changes. A team of eight physicians reviewed the charts for exclusion criteria used in the studies that have

**Table 2**
*Exclusion criteria used to establish study population, as used in efficacy studies of BVNA.*

- Exclusion criteria:
  - Radicular pain (any pain or neurological deficit that traveled along a dermatomal distribution into the lower extremity at or below the medial thigh)
  - Previous lumbar spine surgery
  - Symptomatic spinal stenosis, defined as the presence of neurogenic claudication
  - Diagnosed with osteoporosis
  - Disc extrusion or protrusion >5 mm on MRI
  - Spondylolisthesis >2 mm at any level
  - Involved in litigation related to back pain or injury
  - BMI >40
  - Facet arthrosis or edema correlated with axial low back pain via a positive radiofrequency ablation
  - Receiving disability benefits
  - Currently taking extended-release narcotics
  - Chronic low back pain for less than six months
- Additional exclusion criteria
  - Not seen by a physician within the spine center
  - No MRI of the lumbar spine performed from January 1, 2019 to January 1, 2020 and available within electronic record
  - Duplicate Medical Record Number

demonstrated potential benefit of BVNA [17–19]. There were additional exclusionary criteria applied in order to reduce redundancy of the medical records included and exclude patients who were not seen by the spine-trained physiatrists. Each patient chart that met exclusion criteria was verified a second time by the study PI prior to removal from the cohort. The final cohort before imaging review included those charts of the studied physicians which met the appropriate diagnosis codes, did not have any of the exclusion criteria listed in Table 2, had a lumbar MRI ordered during the study time frame, and had an MRI within our electronic medical record for review (Fig. 2).

A team of imaging reviewers was established with two board-certified and fellowship trained neuroradiologists and two board-certified and fellowship-trained interventional spine physiatrists. Collectively these reviewers, who were all familiar with identifying Modic changes, met with the investigator for whom Modic changes are eponymously named, Dr. Michael Modic, for standardization of categorization. Furthermore, reviewers were provided with a document of visual examples and written characterization instructions as seen in Table 3 [7].

Each of the four reviewers then independently analyzed each lumbar MRI for the presence of Modic changes at vertebral levels L3-S1. Each reviewer noted the presence or absence of Modic changes at each endplate. Modic changes were characterized as either Type 1, Type 2, or Type 3. Each reviewer then provided a grade for the size of the respective Modic change using the Udby classification system [1]. No other radiologic features were collected. As each reviewer provided independent determinations, only the PI and statistical investigator had access to and compared conclusions between reviewers.

### 2.1. Udby Classification [1]

The respective endplate with a suspected Modic change is identified. The Modic change should be measured on the sagittal T1 and T2/STIR sequence in accordance with the classification found in Table 3. If there are believed to be mixed Modic Changes present, the grading is deferred to the most clinically significant characterization. For example, if mixed Type 1 and 2 changes, then the grade would be Type 1. If mixed, Type 2 and Type 3, then the grade would be Type 2. If mixed Type 1 and Type 3, then the grade would be Type 1.

All sagittal slices at each vertebral body are reviewed. The sagittal slice with the Modic change of the largest cephalad extent if measuring from the inferior endplate or caudal extent if measuring from the superior endplate is then given a grade. If <25% involvement from the respective endplate upon which the Modic change begins, then it is given a "Grade A." If 25–50% involvement from the respective endplate upon which the Modic change begins, then it is given a "Grade B." If >50% involvement from the respective endplate upon which the Modic change begins, then it is given a "Grade C." The preferred nomenclature for this system for a Type 1 Modic change at L5 with >50% vertical involvement on sagittal imaging would be: L5 MC 1 Grade C.

### 2.2. MRI parameters

MRI of the lumbar spine was performed at one institution using identical protocols and the same 1.5-T scanner with the same software version (Magnetom Avanto B19; Siemens). This MRI included sagittal T1- and T2-weighted fast spin echo ('T1/T2'). The integrated spine array coil was used, but no surface coils. Echo time (ms)/repetition time (ms) was 11/575 for T1, 87/3700 for T2. Echo train length was 5 for T1, 17 for T2. Matrix was 384 × 269 for T1/T2. Slice thickness/spacing was 4 mm/0.4 mm and field of view was 300 mm × 300 mm for all three sequences.

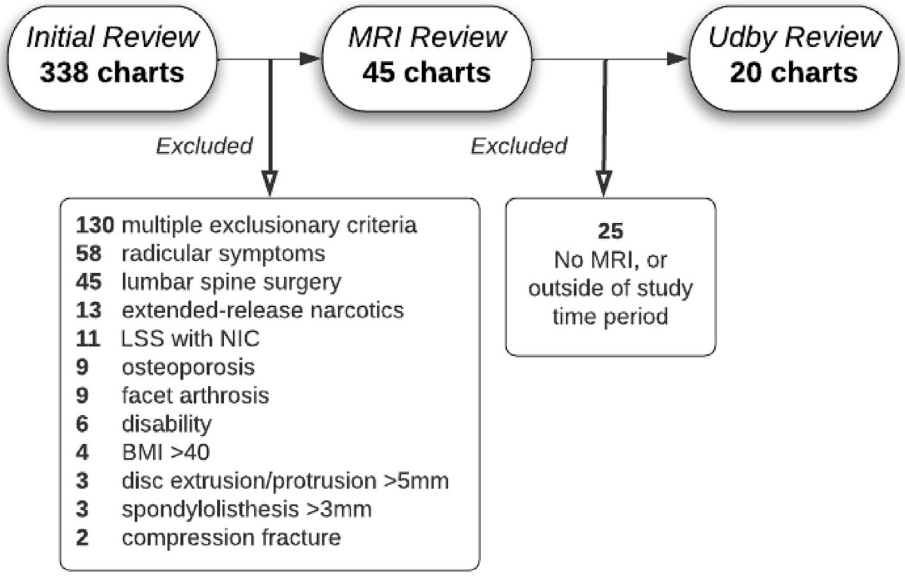


**Fig. 2.** Eligibility flow diagram.

**Table 3**
*MC Definitions* [4].

| Grade | T1 MRI Signal | T2/STIR MRI Signal |
|---|---|---|
| 1 | Hypointense | Hyperintense |
| 2 | Hyperintense | Hyperintense |
| 3 | Hypointense | Hypointense |

Grade 1 Modic Changes: Hypointense on T1, and Hyperintense on T2/STIR MRI.
Grade 2 Modic Changes: Hyperintense on T1, and Hyperintense on T2/STIR MRI.
Grade 3 Modic Changes: Hypointense on T1, and Hypointense on T2/STIR MRI.

**Table 4**
Interrater agreement of Udby Classification.

| | Kappa value | 95% confidence interval |
|---|---|---|
| All 4 reviewers | 0.5899 | 0.4860-0.6939 |
| Radiologists only | 0.6275 | 0.4966-0.7584 |
| Physiatrists only | 0.4684 | 0.3062-0.6307 |

### 2.3. Statistical methodology

When classifying changes to the vertebral endplates, findings were grouped by vertebral level and location on the level of interest (i.e., upper or lower endplate). Modic changes were classified into one of ten categorical combinations: no changes, Modic type 1, 2 or 3 changes; and the percent of the vertebra involved (<25% involvement [a], 25–50% [b], or >50% [c]). Kappa was weighted in an ordinal fashion to reflect the progressive nature of findings (e.g., the difference between a finding of Modic type 2a and Modic type 2c is greater than the difference between type 2a and type 2b). Agreement for the purposes of Kappa scoring was defined as 100% agreement regarding the presence or absence of Modic changes, the type of Modic change, and the Udby classification method across all 4 reviewers.

Interrater reliability across two (intra-specialty) and four (all raters) was calculated using Cohen's weighted kappa via the Stata module Kappaetc, with design-based approach to determine 95% confidence intervals [20,21]. All analyses were conducted using the Stata statistical software package, version 16.1 (StataCorp, College Station, TX), and alpha was set at 0.05.

## 3. Results

20/338 charts met the criteria for review (Fig. 2). There were 10 unique choices at each of 6 endplates evaluated across 20 patients. Thus, there were 480 unique data points to compare across the four independent reviewers. Of these, 269 (56.0%) were read as having no endplate changes, 61 (12.7%) were read as having Modic Type 1 changes, and 150 (31.3%) were read as having Modic Type II changes. There were no endplates read as having Modic Type III changes. The inferior endplate at L5 and the superior S1 endplate were most commonly read as having pathologic changes (46, or 57.5% of endplate-reads), followed by the superior endplate of L5 (41 [51.3%]), the inferior endplate of L4 (35 [43.8%]), the superior endplate of L4 (25 [31.3%]), and the inferior endplate of L3 (18 [22.5%]). At least one reviewer noted the presence of pathology in 8/20 images at the inferior endplate of L3, 12/20 at the superior endplate of L4, 14/20 at the inferior endplate of L4, 17/20 at the superior endplate of L5, 13/20 at the inferior endplate of L5, and 14/20 at the superior endplate of S1.

The Udby classification system produced kappa scores of 0.5899 (95% CI 0.4860–0.6939) across all four reviewers within our cohort (Table 4). Using the McHugh interpretation of kappa scores, our level of agreement is weak to moderate [13]. The two radiologists tended to have a greater level of agreement (K = 0.6275 [95% CI 0.4966–0.7584]) than the two physiatrists (K = 0.4684 [95% CI 0.3062–0.6307]), though the study was not powered sufficiently to identify differences between reviewer specialty.

## 4. Discussion

This manuscript should provide an initial data point from which to further assess the interrater reliability of the Udby classification method as a tool for characterizing Modic changes. The intent of these authors was to inform others regarding the Udby method while providing a first estimate, but not the final word on the validity of the tool.

Other grading methods have been suggested for the description of Modic changes on MRI. Often these methods have introduced many variables (e.g. volume, location, size, endplate involvement, osteophyte formation) which will ultimately increase both inter and intra rater disagreement. None of these grading methods have been implemented broadly across studies. A grading score should be implemented if it is clinically reliable and applicable to many patients with Modic changes. The advantages of the Udby method is that it ensures a consistent relevant grading for all types of Modic changes. This makes interstudy comparison of clinical studies on Modic changes possible.

Future research directives should focus on assessing the measurement tool across larger cohorts. If consistent interrater reliability can be demonstrated, then research should move forward assessing the clinical significance of the different grading scores pre and post procedurally in patients undergoing basivertebral nerve ablation, lumbar medial branch ablation, lumbar disc herniation surgery, and lumbar fusion. If the vertical extent of Modic changes truly do have clinical implications, then the Udby method may be a tool to measure those implications. In theory, the Udby method may then be able to help provide clinicians with information to help guide decision making to improve patient outcomes.

Modic changes are not clinically homogenous. Research on the topic has suggested that Type 1 Modic changes are more correlative with chronic low back pain than Type 2 Modic changes. Other research suggests the size of the lesion in the sagittal plane may present differing severity of symptoms. The type and size of the Modic change may suggest different clinical courses [1,4–6,22–24]. However, a definitive relationship has yet to be established. The Udby method has the potential to be a tool to help researchers assess the validity of that relationship, and clinicians a tool to communicate to patients the severity of their imaging findings. With the advent of basivertebral nerve ablation the more granular characterization of Modic changes via the Udby method may yield improved patient selection. Currently, this is speculative and subject to future research to more definitively measure the value, or lack thereof, of this novel grading criteria.

There are limitations of this study's methodology which should shape the lens through which readers interpret its findings. There was no standardization of the screen or visualization software upon which images were reviewed. The reviewers were not standardized to the use of a measuring tool for the vertical extent of each Modic change which adds a degree of subjectivity to a process meant to provide increased objectivity. A classification of the vertical extent of each Modic change without a standardized tool to measure each may be insufficient to produce high agreement. Thus, more standardization of quantifying the vertical extent of each Modic change may be necessary to produce higher agreement. Additionally, the relatively high frequency of endplates read as having no pathological changes likely introduces some bias toward a greater level of agreement. We believe that the presence of vertebral endplate pathology in the general population is more likely to fit a Poisson distribution (higher frequency of little or no pathology, lower frequency of severe pathology) than a Gaussian distribution, and therefore decided that agreement should be calculated without removing these “no pathology” data points in order to better approximate reality.

Alternatively, our study has many strengths which increase internal and external validity. The cohort reviewed is clinically relevant as they were intentionally selected for review based on potential BVNA qualifications. The imaging review process was completed by four clinicians experienced in identification of Modic changes. The reviewers were not standardized to screen, software, or digital measuring tool which provides generalizability to the study. Moreover, reviewers independently

assessed each lumbar MRI following a review of Modic changes grading criteria led by Dr. Michael Modic.

Interestingly, the imaging reviewers raised a concern regarding how to determine the extent of an individual Modic changes in which overlap of consecutive changes were seen. Questions arose with how to assess whether such a Modic change represented 25% involvement from endplate A and >50% involvement from endplate B, or vice versa. This study failed to explain how to make that distinction to reviewers. While there were no Modic changes which fit this description in our cohort, the concern on how to address such a Modic change moving forward remained.

Overall, we suggest that future studies on Modic changes should include the Udby classification in order to validate the reliability and potential clinical significance or the grading scheme in larger cohorts.

## 5. Conclusion

The Udby system offers a unified method to further characterize Modic changes. This study demonstrated an interrater reliability score of 0.5899 (95% CI 0.4860–0.6939) for the system. While this figure is promising, further research is necessary before definitive claims regarding the interrater agreement validity of the Udby characterization system may be made.

## Conflicts of interest

The authors declare no conflicts of interest.

## Funding

The authors have no sources of funding to declare for this manuscript.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

[1] Udby Peter M, Samartzis Dino, Carreon Leah Y, Anderson Mikkel, Karppinen Jaro, Modic Michael. A definition and clinical grading of Modic changes. J Orthop Res 2021:1–7. https://doi.org/10.1002/jor.25240.

[2] de Roos A, Kressel H, Spritzer C, Dalinka M. MR imaging of marrow changes adjacent to end plates in degenerative lumbar disk disease. AJR Am J Roentgenol 1987;149(3):531–4. https://doi.org/10.2214/ajr.149.3.531.

[3] Modic MT, Masaryk TJ, Ross JS, Carter JR. Imaging of degenerative disk disease. Radiology 1988;168(1):177–86. https://doi.org/10.1148/radiology.168.1.3289089.

[4] Herlin C, Kjaer P, Espeland A, et al. Modic changes—Their associations with low back pain and activity limitation: a systematic literature review and meta-analysis. PLoS One 2018;13(8):e0200677. https://doi.org/10.1371/journal.pone.0200677.

[5] Määttä JH, Karppinen J, Paananen M, et al. Refined phenotyping of modic changes: imaging biomarkers of prolonged severe low back pain and disability. Medicine (Baltim) 2016;95(22):e3495. https://doi.org/10.1097/MD.0000000000003495.

[6] Weishaupt D, Zanetti M, Hodler J, et al. Painful lumbar disk derangement: relevance of endplate Abnormalities at MR imaging. Radiology 2001;218(2):420–7. https://doi.org/10.1148/radiology.218.2.r01fe15420.

[7] Jones A, Clarke A, Freeman BJC, Lam KS, Grevitt MP. The Modic classification: inter- and intraobserver error in clinical practice. Spine 2005;30(16):1867–9. https://doi.org/10.1097/01.brs.0000173898.47585.7d.

[8] Fields AJ, Battié MC, Herzog RJ, et al. Measuring and reporting of vertebral endplate bone marrow lesions as seen on MRI (Modic changes); recommendations from the ISSLS Degenerative Spinal Phenotypes Group. Eur Spine J Off Publ Eur Spine Soc Eur Spinal Deform Soc Eur Sect Cerv Spine Res Soc 2019;28(10):2266–74. https://doi.org/10.1007/s00586-019-06119-6.

[9] Wang Y, Videman T, Niemeläinen R, Battié MC. Quantitative measures of modic changes in lumbar spine magnetic resonance imaging: intra- and inter-rater reliability. Spine 2011;36(15):1236–43. https://doi.org/10.1097/BRS.0b013e3181ecf283.

[10] Peterson CK, Gatterman B, Carter JC, Humphreys BK, Weibel A. Inter- and intraexaminer reliability in identifying and classifying degenerative marrow (Modic) changes on lumbar spine magnetic resonance scans. J Manip Physiol Ther 2007;30(2):85–90. https://doi.org/10.1016/j.jmpt.2006.12.001.

[11] Jensen TS, Karppinen J, Sorensen JS, Niinimäki J, Leboeuf-Yde C. Vertebral endplate signal changes (Modic change): a systematic literature review of prevalence and association with non-specific low back pain. Eur Spine J Off Publ Eur Spine Soc Eur Spinal Deform Soc Eur Sect Cerv Spine Res Soc 2008;17(11):1407–22. https://doi.org/10.1007/s00586-008-0770-2.

[12] Lawan A, Leung A, Battié MC. Vertebral endplate defects: nomenclature, classification and measurement methods: a scoping review. Eur Spine J Off Publ Eur Spine Soc Eur Spinal Deform Soc Eur Sect Cerv Spine Res Soc 2020;29(6):1397–409. https://doi.org/10.1007/s00586-020-06378-8.

[13] McHugh ML. Interrater reliability: the kappa statistic. Biochem Med 2012;22(3):276–82.

[14] Viswanathan VK, Shetty AP, Rajasekaran S. Modic changes - An evidence-based, narrative review on its patho-physiology, clinical significance and role in chronic low back pain. J Clin Orthop Trauma 2020;11(5):761–9. https://doi.org/10.1016/j.jcot.2020.06.025.

[15] ICD-10 resources | CMS. https://www.cms.gov/Medicare/Coding/ICD10/ICD-10Resources. [Accessed 4 August 2021].

[16] National center for advancing translational sciences (NCATS) clinical translational science award (CTSA) program, *Award Number 5UL1TR002243-03*.

[17] Fischgrund JS, Rhyne A, Franke J, et al. Intraosseous basivertebral nerve ablation for the treatment of chronic low back pain: a prospective randomized double-blind sham-controlled multi-center study. Eur Spine J Off Publ Eur Spine Soc Eur Spinal Deform Soc Eur Sect Cerv Spine Res Soc 2018;27(5):1146–56. https://doi.org/10.1007/s00586-018-5496-1.

[18] Khalil JG, Smuck M, Koreckij T, et al. A prospective, randomized, multicenter study of intraosseous basivertebral nerve ablation for the treatment of chronic low back pain. Spine J Off J North Am Spine Soc 2019;19(10):1620–32. https://doi.org/10.1016/j.spinee.2019.05.598.

[19] Relievant Medsystems, Inc. A prospective randomized, double-blind, controlled investigation evaluating the intracept intraosseous nerve ablation system for the reduction of pain in patients with chronic axial low back pain. clinicaltrials.gov. https://clinicaltrials.gov/ct2/show/NCT01446419. [Accessed 13 January 2022].

[20] Klein D. Implementing a general framework for assessing interrater agreement in Stata. STATA J 2018;18(4):871–901.

[21] Feng GC. Factors affecting intercoder reliability: a Monte Carlo experiment. Qual Quantity 2013;47(5):2959–82. https://doi.org/10.1007/s11135-012-9745-9.

[22] Dudli S, Fields AJ, Samartzis D, Karppinen J, Lotz JC. Pathobiology of modic changes. Eur Spine J Off Publ Eur Spine Soc Eur Spinal Deform Soc Eur Sect Cerv Spine Res Soc 2016;25(11):3723–34. https://doi.org/10.1007/s00586-016-4459-7.

[23] Jensen RK, Leboeuf-Yde C, Wedderkopp N, Sorensen JS, Jensen TS, Manniche C. Is the development of Modic changes associated with clinical symptoms? A 14-month cohort study with MRI. Eur Spine J 2012;21(11):2271–9. https://doi.org/10.1007/s00586-012-2309-9.

[24] Thompson KJ, Dagher AP, Eckel TS, Clark M, Reinig JW. Modic changes on MR images as studied with provocative diskography: clinical relevance—A retrospective study of 2457 disks. Radiology 2009;250(3):849–55. https://doi.org/10.1148/radiol.2503080474.