



Differential patient responses to spinal manipulative therapy and their relation to spinal degeneration and post-treatment changes in disc diffusion

Arnold Y. L. Wong^{1,2}  · Eric C. Parent¹ · Sukhvinder S. Dhillon³ · Narasimha Prasad⁴ · Dino Samartzis⁵ · Gregory N. Kawchuk¹

Received: 10 November 2017 / Accepted: 30 November 2018 / Published online: 2 January 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

Purpose Our prior study revealed that people with non-specific low back pain (LBP) who self-reported a > 30% improvement in disability after SMT demonstrated significant post-treatment improvements in spinal stiffness, dynamic muscle thickness and disc diffusion, while those not having self-reported improvement did not have these objective changes. The mechanism underlying this differential post-SMT response remains unknown. This exploratory secondary analysis aimed to determine whether persons with non-specific LBP who respond to spinal SMT have unique lumbar magnetic resonance imaging (MRI) findings compared to SMT non-responders.

Methods Thirty-two participants with non-specific LBP received lumbar MRI before and after SMT on Day 1. Resulting images were assessed for facet degeneration, disc degeneration, Modic changes and apparent diffusion coefficient (ADC). SMT was provided again on Day 4 without imaging. SMT responders were classified as having a $\geq 30\%$ reduction in their modified Oswestry disability index at Day 7. Baseline MRI findings between responders and non-responders were compared. The associations between SMT responder status and the presence/absence of post-SMT increases in ADC values of discs associated with painful/non-painful segments as determined by palpation were calculated. In this secondary analysis, a statistical trend was considered as a *P* value between 0.05 and 0.10.

Results Although there was no significant between-group difference in all spinal degenerative features (e.g. Modic changes), SMT responders tended to have a lower prevalence of severely degenerated facets ($P=0.05$) and higher baseline ADC values at the L4-5 disc when compared to SMT non-responders ($P=0.09$). Post hoc analyses revealed that 180 patients per group should have been recruited to find significant between-group differences in the two features. SMT responders were

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00586-018-5851-2>) contains supplementary material, which is available to authorized users.

✉ Arnold Y. L. Wong
arnold.wong@polyu.edu.hk

- ¹ Department of Physical Therapy, University of Alberta, Edmonton, AB, Canada
- ² Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hung Hom, Hong Kong, SAR, China
- ³ Department of Radiology, University of Alberta, Edmonton, AB, Canada
- ⁴ Department of Mathematical and Statistical Sciences, University of Alberta Hospital, Edmonton, AB, Canada
- ⁵ Department of Orthopaedic Surgery, Rush University Medical Center, Chicago, Illinois, USA

also characterized by significant increases in post-SMT ADC values at discs associated with painful segments identified by palpation ($P < 0.01$).

Conclusions The current secondary analysis suggests that the spines of SMT responders appear to differ from non-responders with respect to degeneration changes in posterior joints and disc diffusion. Although this analysis was preliminary, it provides a new direction to investigate the mechanisms underlying SMT and the existence of discrete forms of treatment-specific LBP.

Graphical abstract These slides can be retrieved under Electronic Supplementary Material.

Key points
[Spinal manipulative therapy, low back pain, apparent diffusion coefficient, facet joint, degeneration]

The objectives of this secondary analysis were to determine whether:

- There were statistically significant differences in spinal degenerative features between low back pain patients with and without clinically significant improvement in disability following spinal manipulative therapy (SMT).
- SMT responder/non-responder status was related to increased diffusion in discs associated with painful segments identified by palpation.

Wong AY, Parent EC, Dhillon SS, Prasad N, Samarziz D, Kawchuk GN (2018) Differential patient responses to spinal manipulative therapy and their relation to spinal degeneration and post-treatment changes in disc diffusion. Eur Spine J.

Interesting characteristics of SMT responders and non-responders as classified by clinically significant improvement in Oswestry disability scores

	Spinal manipulation responders	Spinal manipulation non-responders	Significance ($p < 0.05$)
Severe facet joint degeneration	Less common	More common	No
Diffusion at the L4-5 disc	More	Less	No
Post-SMT diffusion in discs associated with painful segments identified by palpation	Increased	No change	Yes

Wong AY, Parent EC, Dhillon SS, Prasad N, Samarziz D, Kawchuk GN (2018) Differential patient responses to spinal manipulative therapy and their relation to spinal degeneration and post-treatment changes in disc diffusion. Eur Spine J.

Take Home Messages

- Although no statistically significant differences in spinal structural features between SMT responders and non-responders, those responders demonstrate a trend of a higher prevalence of severe facet degeneration and relatively lower baseline disc diffusion at the L4-5 disc.
- Spinal degeneration (especially the discs) may be a potential treatment effect modifier for SMT in patients with non-specific low back pain.
- SMT responders demonstrate post-SMT increases in water diffusion in discs associated with painful segments.

Wong AY, Parent EC, Dhillon SS, Prasad N, Samarziz D, Kawchuk GN (2018) Differential patient responses to spinal manipulative therapy and their relation to spinal degeneration and post-treatment changes in disc diffusion. Eur Spine J.

Keywords Spinal manipulative therapy · Low back pain · Apparent diffusion coefficient · Facet joint · Degeneration

Introduction

Low back pain (LBP) is the most common disabling health condition in the world [1]. Despite its high prevalence, approximately 90% of these cases are diagnosed as non-specific LBP due to unknown aetiology [2]. Accordingly, patients with non-specific LBP are thought to be a heterogeneous collection of various aetiologies [2] as demonstrated by variable responses to any given intervention [3–7].

Notably, spinal manipulative therapy (SMT) is a common treatment for non-specific LBP that has shown a range of therapeutic effects [3–5]. Our recent study confirmed this variation in outcomes. Some participants demonstrated clinically significant improvements in LBP-related disability after 1 week of SMT treatments displayed post-SMT decreases in spinal stiffness, increases in lumbar multifidus contraction and increases in water diffusion at the L4-5 and L5-S1 discs [3]. Conversely, no such changes were observed in SMT non-responders, untreated non-specific LBP controls or untreated asymptomatic controls [3]. Divergent physical and clinical responses in patients with non-specific LBP following an identical SMT application highlight the importance of matching patient subgroups to the right LBP intervention [8]. Importantly, these findings also offer a unique opportunity to identify whether the differential treatment responses arise from patients with distinct underlying lumbar problems—a long-held premise with little empirical evidence.

Since our prior results have shown that only SMT responders demonstrated post-treatment improvement in

water diffusion in multiple lumbar discs [3], it may imply that SMT non-responders may display other spinal degenerative features (e.g. vertebral subchondral bone marrow lesions) that differ from SMT responders [9–13]. Additionally, because post-SMT improvement in disc diffusion may benefit disc health and repair [13, 14], an exploration of the association between post-SMT improvement in disc diffusion at certain segments (especially discs associated with painful segments as determined by palpation) and corresponding clinical improvements would help determine whether post-SMT enhancement of diffusion in certain discs on magnetic resonance imaging (MRI) was related to improved clinical outcomes.

Thus, our primary objective was to perform an exploratory secondary analysis of our previous data to determine whether SMT responders differed in their lumbar degeneration status from SMT non-responders, which might explain differential therapeutic responses of patients with LBP after SMT. An additional objective was to determine whether the presence/absence of increase in disc diffusion of painful segments following the first application of SMT was related to SMT responder/non-responder status after 1 week. It was hypothesized that SMT non-responders displayed significantly more degenerative spinal features than SMT responders. It was also hypothesized that an improvement in disc diffusion of pain segments after the first application of SMT was associated with favourable clinical improvements after 1 week.

Materials and methods

Study design and participants

We performed an exploratory secondary analysis of a prospective non-randomized clinical trial investigating post-SMT differential responses of individuals with LBP [3] (ClinicalTrials.gov identifier: NCT01761838). Participants aged 18–60 years with non-specific LBP were recruited from multidisciplinary settings. All enrolled participants had a minimum LBP intensity of 2 on the 11-point numeric pain rating scale and at least 20% on the modified Oswestry Disability Index (mODI) [3]. Exclusion criteria included but were not limited to: prior lumbosacral surgery, prior SMT within the past 4 weeks, scoliosis, pregnancy, spinal fractures/tumours or any contraindications for MRI [3].

To minimize the risk of unknowingly recruiting an excessive number of SMT responders or non-responders, we attempted to balance these groups by employing a clinical prediction rule for identifying potential SMT responders during enrolment [15]. Individuals with ≤ 2 items in the clinical prediction rule were recruited as potential SMT non-responders, while those with ≥ 4 characteristics were recruited as potential responders (“Appendix 1”) [3, 4]. Individuals with exactly three characteristics were excluded as per prior research [3, 4]. All participants gave written informed consent for the study approved by the institutional research ethics board.

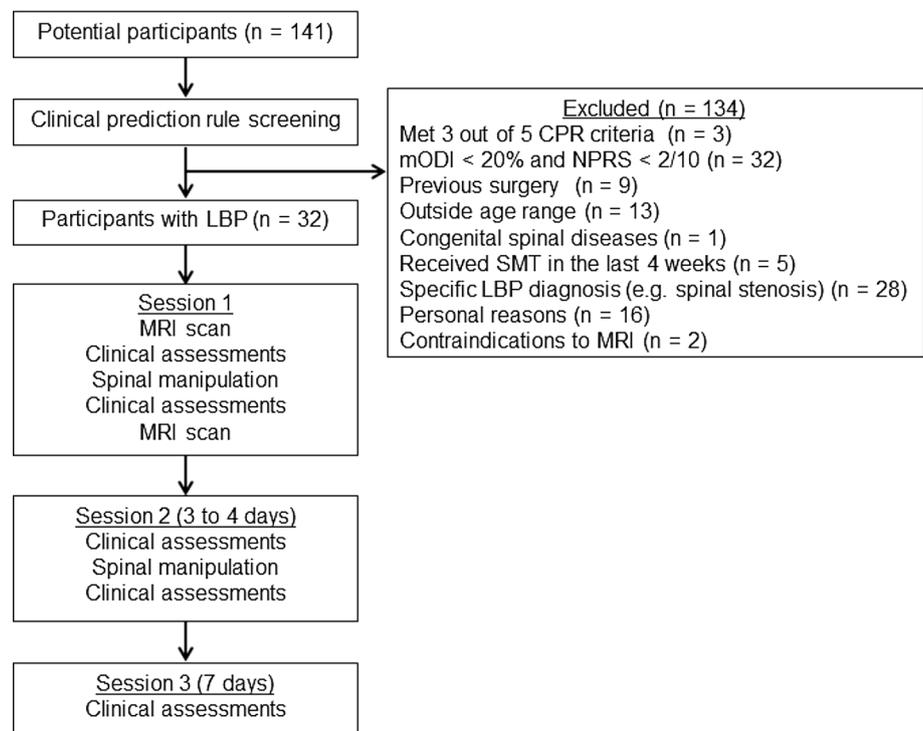
Procedure

Thirty-two participants with LBP attended three morning sessions in 1 week (Fig. 1). On session 1, participants provided demographics and completed the mODI. The lumbar spine was palpated at the spinous processes to identify painful and non-painful segment(s) [16]. The cephalic and caudal intervertebral discs (IVDs) associated with the painful segment were noted. Additionally, participants underwent lumbar MRI (L1–S1), spinal stiffness assessments and multifidus ultrasonography before, and after, the application of SMT in the first session [3]. The mean time between scans was 1 h. On session 2 (3–4 days later), participants repeated these procedures but without MRI. On session 3 (Day 7), only the mODI score, spinal stiffness and multifidus function were completed; no SMT was provided. Participants were then dichotomized as SMT responders and non-responders based on $\geq 30\%$ and $< 30\%$ reduction in baseline mODI scores. This cut-off is recognized as the minimal clinically important improvement in LBP disability [17] although different minimal clinically important improvement values have been suggested by different studies [18–20].

Spinal manipulative therapy

Participants received standardized SMT described elsewhere [3]. Briefly, it involved a long-lever application of a postero-inferior thrust to the pelvis (“Appendix 2”). During sessions

Fig. 1 Flow diagram of the study. CPR, clinical prediction rule for successful spinal manipulative therapy; LBP, low back pain; MRI, magnetic resonance imaging; mODI, modified Oswestry disability index; NPRS, numeric pain rating scale; SMT, spinal manipulative therapy



1 and 2, a maximum of two SMT applications were allowed for each side of the participant.

Imaging

Participants' lumbar spines were scanned by a 1.5 Tesla MRI unit (MAGNETOM Symphony, Siemens Medical Solutions) using multi-element spine coils. The pre-SMT imaging protocol included sagittal T1- and T2-weighted turbo spin echo, axial T2-weighted turbo spin echo, sagittal STIR sequence and sagittal diffusion-weighted imaging (DWI) sequences (Table 1), while post-SMT imaging only involved DWI [3]. Sagittal DWI were acquired using a single shot, dual spin echo, echo-planar imaging acquisition with multi-element spine coils and abdominal coils. The image processing system of the scanner constructed mean apparent diffusion coefficient (ADC) maps.

Grading of MRI findings

A certified radiologist, blinded to the ADC findings, reviewed each participants' imaging and provided scoring for: facet joint degeneration, IVD degeneration and Modic changes (MCs) between the L3 and S1 segments. These segments were chosen as LBP was commonly noted in the lower lumbar spine [21, 22].

The facet joint degeneration at L3–L5 levels was graded by a validated 4-point scale having a high intra-observer reliability (kappa 0.70–0.77) and clinical relevance [23]. Facet degeneration was graded based on joint space, the presence of osteophyte, articular process hypertrophy, subarticular erosion and subchondral cyst [23]. A zero score denoted no degeneration, while 3 represented severe degeneration [24].

Disc degeneration at all lumbar levels was graded by the 5-point Pfirrmann grading system that is based on disc height, homogeneity of the disc structure, differentiation between nucleus and annulus, and T2-weighted signal intensity [25]. This grading system has demonstrated excellent intra-rater reliability (kappa from 0.84 to 0.90) [25]. Higher grades indicate more severe degeneration.

The MCs of vertebrae were graded by a method with substantial to almost perfect intra-observer reliability (kappa from 0.77 to 1.00) [26]. Specifically, type I MC (fibrovascular replacement of subchondral bone) was characterized by decreased T1 and increased T2 signal intensity [27]. Type 2 MC (endplate fissures and fatty replacement of subchondral bone) showed hyperintense T1 and isointense/hyperintense T2 signals [27]. Type 3 MC (sclerosis of subchondral bone) appeared as hypointense T1 and T2 signals [27]. If multiple types of MCs coexisted at the endplate/subchondral bone, it was labelled as mixed type 1/2 or 2/3 MC. The locations and sizes (endplate area, maximum height, volume or anteroposterior diameter) of MCs were recorded [28, 29].

ADC measurements

An examiner blinded to the radiologist's grading used a customized programme to measure mean ADC values (ADC mean) of five lumbar discs on the midsagittal ADC map. Notably, a 40 mm² circular region of interest (ROI) was placed in the central, nuclear portion of a disc on the midsagittal ADC map to minimize the noise signals from adjacent vertebral structures. The ADC mean of a disc was estimated from signal intensity within the ROI. If the diameter of the ROI exceeded the disc height, the disc was excluded from analysis [3]. A high ADC value represents higher IVD diffusion [13].

Table 1 Sequence of lumbar magnetic resonance imaging from L1 to S1

Imaging sequence	Details
A localizer sequence	TR/TE 7.0/2.5 ms, flip angle 20°, two coronal and two sagittal images in orthogonal planes
Sagittal T1-weighted turbo spin echo	TR/TE 450/14 ms, voxel size 0.9×0.5×4.0 mm, 265 mm field of view, 15 slices of 4 mm thickness, distance factor 10%, four acquisitions (4 min 35 s scan time)
Sagittal T2-weighted turbo spin echo	TR/TE 3030/121 ms, voxel size 1.0×0.5×4.0 mm, 265 mm field of view, 15 slices of 4 mm thickness, distance factor 10%, four acquisitions (3 min 17 s scan time)
Axial T2-weighted turbo spin echo	TR/TE 3310/122 ms, voxel size 1.3×0.6×4.0 mm, 280 mm field of view, 41 slices of 4.0 mm thickness, distance factor 10%, two acquisitions, axial images from each lumbar intervertebral level (L1/2 to L5/S1) parallel to the vertebral endplates with at least one image per disc (3 min 52 s scan time)
Sagittal short T1 inversion recovery (STIR) sequence	TR/TE 4000/68 ms, inversion time 160 ms, voxel size 1.2×0.8×4.0 mm, 265 mm field of view, 15 slices of 4.0 mm thickness, distance factor 10%, two acquisitions (4 min 30 s scan time)
DWI sequence	TR/TE 6000/76 ms, <i>b</i> -factor 400 s/mm ² , voxel size 2.3×2.3×4.0 mm, 300 mm field of view, 10 excitations, three-scan trace diffusion mode for ADC maps construction, three orthogonal diffusion directions, 15 slices of 4.0 mm thickness, distance factor 0% (4 min 20 s scan time)

The first scan included all the sequences but the second scan only involved the DWI sequence

ADC apparent diffusion coefficient, DWI diffusion-weighted imaging; mm, millimetres, TE echo time, TR relaxation time

Intra-observer reliability of assessments

Three weeks after initial imaging, the radiologist and the examiner independently repeated the grading of MRI findings and ADC measurements on the scans of randomly selected participants. Both readers were blinded to all measurement results.

Statistical analysis

The exploratory, secondary analyses of underpowered data were conducted using SPSS 22.0 software (IBM, Armonk, NY). As some categories of the MRI findings showed low prevalence of specific grades (< 10%) (e.g. type 3 MC), each variable was dichotomized to calculate the intra-observer reliability using kappa analysis [30]. Specifically, Pfirrmann grading was dichotomized into “no/mild disc degeneration” (grade 1–3) and “severe disc degeneration” (grade 4 or 5) [31]. MC was dichotomized as “no” and “yes” (regardless of types or locations). Facet joint degeneration was dichotomized as “no/mild degeneration” (grade 0 and 1) and “severe degeneration” (grade 2 and 3) [32]. Kappa coefficients were interpreted as follows: poor (≤ 0), slight (0.01–0.20), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80) and excellent (0.81–1.00) [33]. The intra-observer reliability of ADC measurements was reported elsewhere [3] and demonstrated excellent intra-class correlation coefficients ($ICC_{3,1}$) estimates (0.97–0.98) (“Appendix 3”) [3, 34].

Chi-square/Fisher’s exact tests assessed differences in the frequency of dichotomized MRI findings between responders and non-responders. Demographic data and baseline ADC mean of all lumbar IVDs in responders and non-responders were checked for normality using Shapiro–Wilk tests. If continuous data were not normally distributed, Mann–Whitney U tests would replace independent *t* tests for between-group comparisons. The effect sizes of all between-group comparisons were calculated [35]. Additionally, to

determine the baseline ADC mean cut-off of each IVD that could differentiate responders and non-responders [36], area under the curve (AUC) of the receiver operating characteristic curve and Youden index were used [34, 36]. An AUC value of 0.5 indicates that the baseline ADC mean of a disc is no better than chance to distinguish between responders and non-responders, while an AUC of 1.00 represents the baseline ADC mean that could always identify the responders [34]. The cut-off was chosen if Youden index value ($J = \max_c [Se(c) + Sp(c) - 1]$) was maximum, where *Se* and *Sp* were the probability of truly identifying responders and non-responders, respectively, at a given cut-off (*c*) [36].

The differences in post-SMT changes in ADC mean of IVDs associated with painful/non-painful segments between responders and non-responders were examined by separate Chi-square tests. To dichotomize the post-SMT changes in ADC mean into increase or no change/decrease in ADC mean, values of the minimal detectable change at the 95% confidence interval (“Appendix 4”) [37] derived from intra-observer reliability of ADC measurements at the respective disc levels were used as the cut-offs. Cramer’s V tests were used to estimate the strength of the association. Cramer’s V values were interpreted as weak (< 0.10), moderate (0.11–0.33) and strong (> 0.30) [38]. The significance level was set at 0.05 for all tests. For this exploratory secondary analysis, a *P* value between 0.05 and 0.10 was considered as a trend towards significance.

Results

Based on mODI scores, 15 participants were classified as SMT responders and 17 as non-responders. The baseline demographics of these two groups were comparable (Table 2). The intra-observer reliability of the various lumbar degeneration grading was excellent ($\kappa = 0.80$ –0.84) (“Appendix 3”) [33].

Table 2 Baseline demographic characteristics of all participants and in the responder/non-responder subgroups as defined by self-reported improvement in disability on session 3*

	All participants (<i>n</i> = 32)	Responders (<i>n</i> = 15)	Non-responders (<i>n</i> = 17)	<i>P</i> value
Age, years	32.2 (10.2)	32.2 (11.3)	32.2 (9.4)	<i>t</i> = −1.0; <i>P</i> = 0.19
Sex (male), %	34.4	40	29.4	Fisher exact test <i>P</i> = 0.71
Body mass index, kg/m ²	25.0 (4.3)	24.6 (3.0)	25.1 (5.3)	<i>t</i> = −0.4; <i>P</i> = 0.07
Pretreatment mODI score (0–100)	23.9 (5.2)	22.9 (4.7)	24.7 (5.6)	<i>t</i> = −1.0; <i>P</i> = 0.17
Pretreatment current NPRS (0–10)	4.3 (1.6)	4.6 (1.6)	4.1 (1.5)	<i>t</i> = 0.9; <i>P</i> = 0.40
FABQW score (0–24)	12.3 (7.7)	11.4 (7.9)	13.2 (7.6)	<i>t</i> = −0.5; <i>P</i> = 0.88
Prior history of LBP, %	87.5	93.3	82.4	Fisher exact test <i>P</i> = 0.60

Values are mean (standard deviation)

FABQW fear avoidance behaviour questionnaire work subscale; kg/m², kilogram per metre squared, LBP/SMT participants with low back pain receiving spinal manipulative therapy, mODI modified Oswestry Disability Index, NPRS numeric pain rating scale

Responders' and non-responders' MRI findings

Compared to SMT responders, SMT non-responders displayed a trend of higher prevalence of severely degenerated facet joints ($\chi^2 = 3.86, P = 0.05, \text{Cramer's } V = 0.14$) (Fig. 2). From the post hoc analysis using G*Power program [39], 180 participants per group should have been recruited to detect a significant between-group difference in facet joint degeneration (Power = 0.80, $\alpha = 0.05$). There was no statistically significant between-group difference in the prevalence of IVD or MC grading. While most IVDs in both subgroups were not degenerated, more degenerated discs were found at the L4-5 and L5-S1 levels (Fig. 3). Modic changes were observed in 46.7% and 58.8% of responders' and non-responders' endplates, respectively. For responders, the prevalence estimates of types 1, 2 and 3 MCs in endplates were 9.5%, 4.8% and 1.0%, respectively (Fig. 4). Similarly, type 1, type 2, type 3 and mixed type 1/2 MCs were noted in 6.7%, 5.0%, 0.0% and 5.9% of non-responders' endplates, respectively (Fig. 4). The anterior superior endplate showed the highest relative frequency of MCs (33.3% to 34.5%) in both subgroups although it did not significantly differ from other locations. The majority of the observed MCs involved only endplates (88.9% in responders, 68.9% in non-responders).

Baseline ADC mean of responders and non-responders

Five out of 160 IVDs were excluded from the ADC measurements because the disc space was smaller than the ROI diameter. All the excluded discs were from five non-responders at

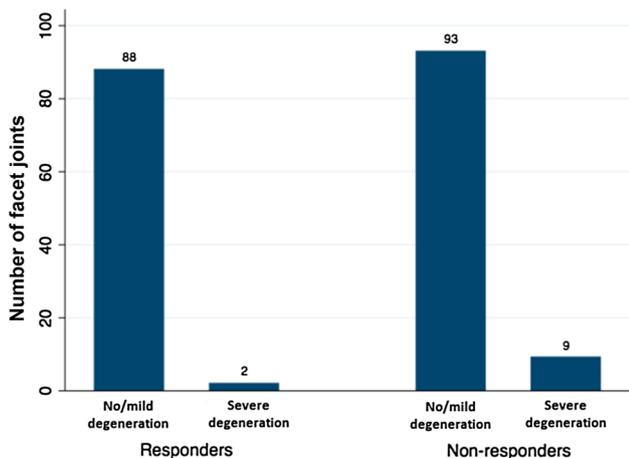


Fig. 2 Different grades of facet joint degeneration in modified Oswestry disability index defined responders (a total of 90 facet joints) and non-responders (a total of 102 facet joints). There was significant difference in the proportion of severely degenerated facet joint between the two subgroups ($\chi^2 = 3.86, \text{Cramer's } V = 0.14, P = 0.05$)

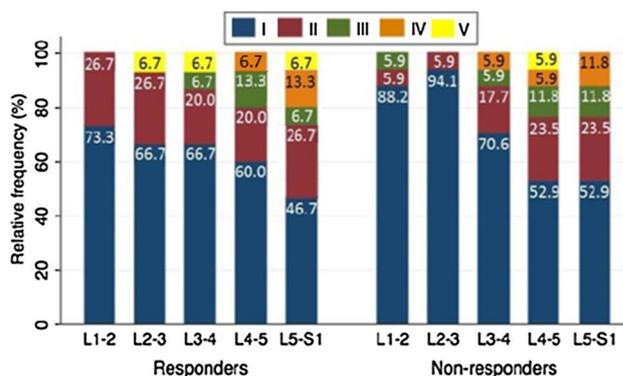


Fig. 3 Different Pfirrmann grades of disc degeneration in modified Oswestry disability index defined responders [$n = 15$] and non-responders [$n = 17$]. Pfirrmann grading was dichotomized into “no/mild disc degeneration” (grade 1 to 3) and “severe disc degeneration” (grade 4 or 5). No statistically significant difference in the prevalence of Pfirrmann grades between the two subgroups. ($\chi^2 = 1.91, \text{Cramer's } V = 0.04, P = 0.17$)

L1-2 ($n = 2$) and L4-5 ($n = 1$) and L5-S1 discs ($n = 2$). There was no significant difference in ADC mean of all intervertebral discs between responders and non-responders. However, compared to responders, non-responders had a trend of lower estimated baseline ADC mean of the L4-5 disc ($U = 78.00, P = 0.09, r = 0.30$) (Table 3). The post hoc analysis revealed that a sample size of 43 per group should have been recruited to detect a significant between-group difference of this magnitude (Power = 0.80, $\alpha = 0.05$). Similarly, while AUCs for baseline ADC mean of discs at each level were not statistically significant, the AUC of the L4-5 disc was 0.70 (95% CI: 0.50–0.89; $P = 0.07$). Specifically, patients with baseline ADC mean higher than $2.07 \times 10^{-3} \text{ mm}^2/\text{s}$ at the L4-5

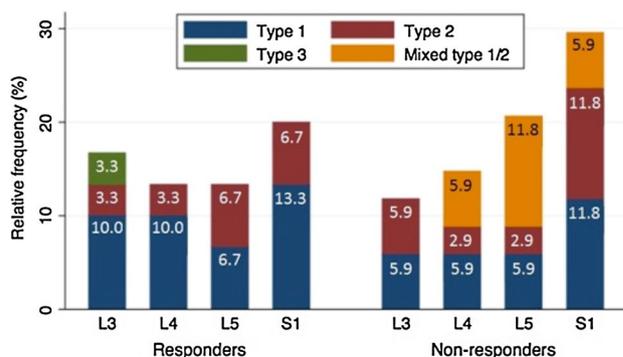


Fig. 4 Different types of Modic changes in responder and non-responder subgroups. Responders (30 endplates/level for L3 to L5 levels, 15 endplates for the S1 level) and non-responders (34 endplates/level for L3 to L5 levels, 17 endplates for the S1 level). Modic changes were dichotomized as “no” and “yes” (regardless of types or locations). No statistically significant difference in the prevalence of Modic changes between these subgroups ($\chi^2 = 1.47, \text{Cramer's } V = 0.04, P = 0.23$)

Table 3 Comparisons of baseline mean apparent diffusion coefficient (ADC) value of each lumbar disc level between spinal manipulative therapy responders and non-responders

Disc levels	Median ADC values of responders (mm ² /s)	Median ADC values of non-responders (mm ² /s)	Mann–Whitney U test	<i>P</i> value	Effect size (<i>r</i>)	Estimated sample size per group*
L1–L2	2.16 × 10 ⁻³	2.26 × 10 ⁻³	- 1.06	0.29	0.19	111
L2–L3	2.01 × 10 ⁻³	2.04 × 10 ⁻³	- 1.04	0.30	0.18	140
L3–L4	2.03 × 10 ⁻³	2.03 × 10 ⁻³	- 0.19	0.85	0.03	4568
L4–L5	2.09 × 10 ⁻³	2.02 × 10 ⁻³	- 1.66	0.09	0.30	43
L5–S1	2.09 × 10 ⁻³	2.08 × 10 ⁻³	- 0.92	0.36	0.17	158

*Sample size calculation based on a significant level of 0.05 and statistical power of 0.80

Table 4 Associations between the post-spinal manipulative therapy (SMT) changes in mean apparent diffusion coefficient (ADC) values of intervertebral discs at painful or non-painful segments as identified by palpation and clinical responses on session 3

Post-SMT ADC change of lumbar discs on session 1	Response on session 3*	
	Responders	Non-responders
Discs associated with the painful segments		
Increase ^a	30 (20)	10 (21)
% within the column	73.2%	23.3%
Decrease/no change	11 (22)	33 (23)
% within the column	26.8%	76.7%
Post-SMT ADC change of lumbar discs on session 1		
Response on session 3 [#]		
	Responders	Non-responders
Discs associated with the non-painful segments		
Increase ^a	9 (6)	4 (7)
% within the column	27.3%	10.5%
Decrease/no change	24 (27)	35 (31)
% within the column	72.7%	89.5%

The number in parentheses represents the expected frequency of a particular category. Some responders/responders might have one or more lumbar segments identified as painful during palpation

*Significant association between discs at painful segments and the corresponding clinical responses: $\chi^2 = 20.96$; Cramer's $V = 0.50$, $P < 0.01$

[#]No significant association between discs at non-painful segments and the corresponding clinical responses: $\chi^2 = 3.31$; Cramer's $V = 0.22$, $P > 0.05$

^aThe post-SMT changes in ADC mean were dichotomized as an increase and no change/decrease categories based on the minimal detectable change values of repeated ADC measurements at the respective disc levels (“Appendix 4”)

disc were more likely to be responders (sensitivity = 0.71; specificity = 0.68).

Comparing post-SMT ADC mean changes in discs at painful/non-painful segments between responders and non-responders

Compared to non-responders, a significantly higher proportion of discs identified as painful in responders demonstrated immediate post-SMT increases in ADC mean (Cramer's $V = 0.50$, $P < 0.01$) (Table 4). Conversely, both responders and non-responders did not show significant differences in the patterns of post-SMT changes in ADC mean of discs associated with non-painful segments ($P > 0.01$) (Table 4).

Discussion

While the current secondary analysis has revealed no statistically significant differences in spinal structural features between SMT responders and non-responders, lumbar degeneration status appears to be related to post-SMT treatment outcomes. Compared to SMT responders, the non-responders demonstrated a trend of a higher prevalence of severe facet degeneration ($P = 0.05$) and relatively lower baseline ADC mean of the L4-5 disc ($P = 0.09$). Given the exploratory nature of our analyses, it is important to consider all potential findings that may explain mechanisms underlying the differential responses between SMT responders

and non-responders. Our results suggest that spinal degeneration (especially the discs) may be a potential treatment effect modifier for SMT in patients with non-specific LBP although further studies are warranted to test this hypothesis.

Distinct baseline ADC mean and post-SMT disc diffusion responses were noted between responders and non-responders. Our current data suggest that ADC mean of IVDs may be more sensitive to the microstructural intradiscal changes that may not be quantified by qualitative grading on MRI [40]. Since a lower ADC mean indicates greater disc degeneration [13, 41], our findings suggest that non-responders may have more degenerated L4-5 discs, which may be sources of LBP [42]. A prior ovine study has found that a spinal segment with a degenerated disc displays less movement during SMT as compared to a non-degenerated spine [43]. Such distinct mechanical responses may explain the differential SMT responses. Additionally, Vieira-Pellenz found that male LBP patients with degenerative discs who demonstrated greater increases in body height (a proxy of improved disc height) after a single session of SMT were associated with less LBP prevalence [44]. Given these results, further studies with larger sample sizes are warranted to determine the causal relations among spine degeneration, post-SMT mechanical responses and LBP.

Interestingly, our findings highlighted that responders were characterized by immediate post-SMT increases in ADC mean of discs associated with painful segments on session 1. This finding implies that favourable clinical outcomes arising from SMT may be partly due to improved diffusion of discs associated with painful segments. Since the significant relation between level-specific post-SMT increases in ADC mean of discs and improved clinical outcomes was only found in SMT responders, it corroborates the importance of conducting further studies to clarify the association between spinal degeneration and SMT outcomes, especially when SMT non-responders displayed a trend of more lumbar degenerative features. Notwithstanding the strong association between post-SMT increase in ADC mean of discs on session 1 and clinical improvements on session 3, improved disc diffusion might not be the sole reason for the differential clinical responses of our participants. Since 26.8% of discs at the painful segments of the responders showed no change/decrease in the post-SMT ADC mean (Table 4), their clinical improvements may be related to the mechanical and/or neurophysiological effects of SMT on spinal tissues other than improved disc diffusion. Similarly, if IVDs were not the sole source of LBP or the dosage of SMT was insufficient to elicit significant clinical improvements in the non-responders, enhanced post-SMT disc diffusion would not improve clinical outcomes. Furthermore, SMT is unlikely to benefit non-biomechanical pathologies, e.g. inflammation of vertebral subchondral bone [45] or

increased nociceptive innervation in IVDs/vertebral endplates [46].

Additionally, our findings underscore the presence of treatment-specific types of LBP. As non-specific LBP is a constellation of heterogeneous conditions that have varied mechanisms of pain generation [47, 48], the application of a single LBP treatment to patients with non-specific LBP would be expected to result in diverse clinical outcomes as is the case in our work and that of others [48, 49]. Although there have been many suggestions as to why this diverse range of responses exist to a given treatment [50–53], this secondary analysis suggests that a selective response to a particular intervention does exist and has a potential physiological explanation consisting of distinct biological and biomechanical characteristics. Our findings indicate a novel research direction for investigating the influence of structural features in affecting treatment responses, which may guide certain clinical decision in future. However, it is not recommended to order routine MRI scans for the sole purpose of determining the application of SMT.

The current study has inherent limitations. First, as the current exploratory, secondary analysis was not powered for the comparison of MRI findings between responders and non-responders. As such, no statistically significant difference can be drawn, and our results should be interpreted with caution. Future research with a larger sample size is warranted to validate the current findings. Second, although SMT responders and non-responders were dichotomized based on a 30% decrease in baseline mODI scores at the third visit, the absolute change in mODI score of a SMT responder could be as small as 6 points if the participant's baseline mODI score was only 20 points. While 6 points were still larger than some suggested minimal detectable changes in patients with LBP (ranging from 3.2 to 5.2 points) [18], this value was much smaller than other suggested minimal detectable changes (15–19 points) [19]. Therefore, some SMT responders' mODI improvements might not demonstrate minimal clinically important differences had other cut-off methods been used. Third, since the present study investigated the association between post-SMT changes in ADC mean of IVDs on session 1 and the clinical outcomes on session 3, the relation between serial changes in ADC mean of IVDs and the respective clinical outcomes remains unknown. As prior research suggests that SMT may have a positive dose–response effect on clinical outcomes [54], it is conceivable that SMT may have similar cumulative effects on the disc diffusion changes. Future research should measure both the post-SMT ADC mean and clinical outcomes at successive time points to clarify this relation. Fourth, while our results appeared to suggest distinct MRI spinal findings in SMT responders and non-responders, MRI scans are rarely ordered in clinical practice for patients with non-specific LBP. Future studies should investigate whether

differential spinal degenerative features exist in plain radiographs (i.e. X-rays) between SMT responders and non-responders, which may inform clinical decision-making and further assist in more refined patient selection for further imaging assessments.

Conclusions

There was no significant difference in structural features between SMT responders and non-responders. However, SMT responders demonstrate a trend of a lower prevalence of severely degenerated facets and relatively high baseline ADC values of the L4-5 discs. Interestingly, SMT responders demonstrate post-SMT increases in ADC values of discs associated with painful segments. These results suggest that SMT responses may be related to some underlying structural responses. Our findings provide a new hypothesis/direction for further investigating the underlying nature of the differential response to SMT, possible mechanisms of SMT and the existence of treatment-specific forms of LBP.

Acknowledgements The authors would like to thank the Canadian Chiropractic Research Foundation and Alberta Innovates-Health Solutions Graduate Research Allowance for funding this project. Arnold Wong was supported by the Golden Key Graduate Scholar Award. The authors also thank Magnetic Imaging Consultants for providing scan services, and the River Valley Health Clinic for providing professional spinal manipulation and clinical space. The authors also express gratitude to Mr. Karl Brandt and Ms. Carolyn Berendt for assisting the coding and decoding of data files.

Compliance with ethical standards

Conflict of interest The authors have no potential conflict of interest.

Appendix 1: Clinical prediction rule

Clinical characteristics of the clinical prediction rule for identifying people who benefit from spinal manipulative therapy

Clinical characteristics	Definition of a positive finding
1. Duration of the current episode of low back pain	Less than 16 days
2. Distal symptoms	No symptoms distal to the knee
3. Fear avoidance beliefs questionnaire work subscale	Less than 19 points
4. Lumbar stiffness	At least one lumbar segment is determined to be hypomobile by the examiner using a manual posteroanterior spinal mobility test

Clinical characteristics	Definition of a positive finding
5. Hip internal rotation range of motion	At least one hip with 35° or greater as measured by an inclinometer in prone

Appendix 2: Spinal manipulative therapy procedure

The supine participant crossed and put his/her fingers behind the neck. The clinician stood opposite to the side to be manipulated and side bent the participant’s trunk towards the side of the pelvis to be manipulated, and rotated the trunk in the opposite direction. Then the clinician applied a high-velocity, low-amplitude thrust to the pelvis in a posteroinferior direction. The clinician delivered spinal manipulative therapy to each side at each given session. If the first attempt did not result in cavitation, a second spinal manipulation was allowed for each side. A maximum of two spinal manipulations would be given to each side within a session. In the current study, only two out of 64 sessions required a second spinal manipulation on the one side (one for responder and one for non-responder).

Appendix 3: Intra-observer reliability of the dichotomized degeneration variables

Magnetic resonance imaging findings	Dichotomized category	Kappa (CI)
Facet joint	Normal (grade 0) Abnormal (grade 1, 2 or 3)	0.84 (0.68–1.00)
Disc degeneration	Normal (Pfirrmann grade 1); Abnormal (Pfirrmann grade 2, 3, 4 or 5)	0.80 (0.62–0.98)
Modic changes	Normal (no change) Abnormal (types 1, 2, 3, mixed 1/2 or mixed 2/3)	0.81 (0.62–1.00)

Cohen’s kappa coefficient was interpreted as poor (<0.00), slight (0.00–0.20), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80) and almost perfect (0.81–1.0) agreement [52]

CI confidence interval

Appendix 4: Intra-observer reliability of apparent diffusion coefficient measurement

Disc	ICC _{3,1} (95% CI)	Mean ADC (SD) ^a	SEM	MDC ₉₅	N
L1–2	0.98 (0.96–0.99)	2.181 (0.224)	0.032	0.089	15
L2–3	0.97 (0.94–0.99)	1.966 (0.136)	0.023	0.064	16
L3–4	0.99 (0.97–0.99)	2.010 (0.087)	0.009	0.025	16
L4–5	0.98 (0.95–0.99)	2.039 (0.111)	0.016	0.044	15
L5–S1	0.98 (0.96–0.99)	2.059 (0.246)	0.035	0.097	14

CI confidence interval, ICC intra-class correlation coefficient, SD standard deviation, SEM standard error of measurement, MDC₉₅ minimal detectable change at the 95% confidence interval, N number of disc

^aMean and standard deviation of apparent diffusion coefficient (ADC) values were expressed in units of 10⁻³ mm²/s

References

- Global Burden of Disease Study 2013 Collaborators (2015) Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 386:743–800. [https://doi.org/10.1016/S0140-6736\(15\)60692-4](https://doi.org/10.1016/S0140-6736(15)60692-4)
- Haldeman S, Kopansky-Giles D, Hurwitz EL et al (2012) Advancements in the management of spine disorders. *Best Pract Res Clin Rheumatol* 26:263–280. <https://doi.org/10.1016/j.berh.2012.03.006>
- Wong AYL, Parent EC, Dhillon SS et al (2015) Do participants with low back pain who respond to spinal manipulative therapy differ biomechanically from nonresponders, untreated controls or asymptomatic controls? *Spine* 40:1329–1337. <https://doi.org/10.1097/BRS.0000000000000981>
- Fritz JM, Koppenhaver SL, Kawchuk GN et al (2011) Preliminary investigation of the mechanisms underlying the effects of manipulation: exploration of a multi-variate model including spinal stiffness, multifidus recruitment, and clinical findings. *Spine* 36:1772–1781. <https://doi.org/10.1097/BRS.0b013e318216337d>
- Koppenhaver SL, Fritz JM, Hebert JJ et al (2011) Association between changes in abdominal and lumbar multifidus muscle thickness and clinical improvement after spinal manipulation. *J Orthop Sports Phys Ther* 41:389–399. <https://doi.org/10.2519/jospt.2011.3632>
- Long A, Donelson R, Fung T (2004) Does it matter which exercise? A randomized control trial of exercise for low back pain. *Spine* 29:2593–2602
- Beneciuk JM, Robinson ME, George SZ (2015) Subgrouping for patients with low back pain: a multidimensional approach incorporating cluster analysis and the STarT Back Screening Tool. *J Pain* 16:19–30. <https://doi.org/10.1016/j.jpain.2014.10.004>
- Hebert JJ, Koppenhaver SL, Walker BF (2011) Subgrouping patients with low back pain: a treatment-based approach to classification. *Sports Health* 3:534–542. <https://doi.org/10.1177/1941738111415044>
- Jensen TS, Karppinen J, Sorensen JS et al (2008) Vertebral endplate signal changes (Modic change): a systematic literature review of prevalence and association with non-specific low back pain. *Eur Spine J* 17:1407–1422. <https://doi.org/10.1007/s00586-008-0770-2>
- Kjaer P, Korsholm L, Bendix T et al (2006) Modic changes and their associations with clinical findings. *Eur Spine J* 15:1312–1319. <https://doi.org/10.1007/s00586-006-0185-x>
- Jensen RK, Leboeuf-Yde C, Wedderkopp N et al (2012) Is the development of Modic changes associated with clinical symptoms? A 14-month cohort study with MRI. *Eur Spine J* 21:2271–2279. <https://doi.org/10.1007/s00586-012-2309-9>
- Jensen OK, Nielsen CV, Sørensen JS, Stengaard-Pedersen K (2014) Type 1 Modic changes was a significant risk factor for 1-year outcome in sick-listed low back pain patients: a nested cohort study using magnetic resonance imaging of the lumbar spine. *Spine J* 14:2568–2581. <https://doi.org/10.1016/j.spine.2014.02.018>
- Antoniou J, Demers CN, Beaudoin G et al (2004) Apparent diffusion coefficient of intervertebral discs related to matrix composition and integrity. *Magn Reson Imaging* 22:963–972. <https://doi.org/10.1016/j.mri.2004.02.011>
- Huang Y-C, Urban JPG, Luk KDK (2014) Intervertebral disc regeneration: do nutrients lead the way? *Nat Rev Rheumatol* 10:561–566. <https://doi.org/10.1038/nrrheum.2014.91>
- Flynn T, Fritz J, Whitman J et al (2002) A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. *Spine* 27:2835–2843. <https://doi.org/10.1097/01.BRS.0000035681.33747.8D>
- Haneline MT, Young M (2009) A review of intraexaminer and interexaminer reliability of static spinal palpation: a literature synthesis. *J Manip Physiol Ther* 32:379–386. <https://doi.org/10.1016/j.jmpt.2009.04.010>
- Ostelo RWJG, Deyo RA, Stratford P et al (2008) Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine* 33:90–94. <https://doi.org/10.1097/BRS.0b013e31815e3a10>
- Suarez-Almazor ME, Kendall C, Johnson JA, Skeith K, Vincent D (2000) Use of health status measures in patients with low back pain in clinical settings. Comparison of specific, generic and preference-based instruments. *Rheumatology* 39:783–790. <https://doi.org/10.1093/rheumatology/39.7.783>
- Davidson M, Keating JL (2002) A comparison of five low back disability questionnaires: reliability and responsiveness. *Phys Ther* 82:8–24. <https://doi.org/10.1093/ptj/82.1.8>
- Hagg O, Fritzell P, Nordwall A (2003) The clinical importance of changes in outcome scores after treatment for chronic low back pain. *Eur Spine J* 12:12–20. <https://doi.org/10.1007/s00586-002-0464-0>
- DePalma MJ, Ketchum JM, Saullo T (2011) What is the source of chronic low back pain and does age play a role? *Pain Med* 12:224–233. <https://doi.org/10.1111/j.1526-4637.2010.01045.x>
- Kjaer P, Leboeuf-Yde C, Korsholm L et al (2005) Magnetic resonance imaging and low back pain in adults: a diagnostic imaging study of 40-year-old men and women. *Spine* 30:1173–1180. <https://doi.org/10.1097/01.brs.0000162396.97739.76>
- Weishaupt D, Zanetti M, Boos N, Hodler J (1999) MR imaging and CT in osteoarthritis of the lumbar facet joints. *Skelet Radiol* 28:215–219. <https://doi.org/10.1007/s002560050503>
- Kettler A, Wilke HJ (2006) Review of existing grading systems for cervical or lumbar disc and facet joint degeneration. *Eur Spine J* 15:705–718. <https://doi.org/10.1007/s00586-005-0954-y>
- Pfirschmann C, Metzendorf A, Zanetti M et al (2001) Magnetic resonance classification of lumbar intervertebral disc degeneration.

- Spine 26:1873–1878. <https://doi.org/10.1097/00007632-200109010-00011>
26. Jensen TS, Sørensen JS, Kjaer P (2007) Intra- and interobserver reproducibility of vertebral endplate signal (modic) changes in the lumbar spine: the Nordic Modic Consensus Group classification. *Acta Radiol* 48:748–754. <https://doi.org/10.1080/02841850701422112>
 27. Modic MT, Steinberg PM, Ross JS et al (1988) Degenerative disk disease: assessment of changes in vertebral body marrow with MR imaging. *J Bone Joint Surg Am* 166:193–199. <https://doi.org/10.1148/radiology.166.1.3336678>
 28. Maataoui A, Vogl TJ, Khan MF (2015) Magnetic resonance imaging-based interpretation of degenerative changes in the lower lumbar segments and therapeutic consequences. *World J Radiol* 7:194–197. <https://doi.org/10.4329/wjr.v7.i8.194>
 29. Määttä JH, Karppinen JI, Luk KDK et al (2015) Phenotype profiling of Modic changes of the lumbar spine and its association with other MRI phenotypes: a large-scale population-based study. *Spine J* 15:1933–1942. <https://doi.org/10.1016/j.spine.2015.06.056>
 30. Gjorup T (1988) The kappa coefficient and the prevalence of a diagnosis. *Methods Inf Med* 27:184–186
 31. Kovacs FM, Arana E, Royuela A et al (2014) Disc degeneration and chronic low back pain: an association which becomes non-significant when endplate changes and disc contour are taken into account. *Neuroradiology* 56:25–33. <https://doi.org/10.1007/s00234-013-1294-y>
 32. Kalichman L, Hodges P, Li L et al (2010) Changes in paraspinous muscles and their association with low back pain and spinal degeneration: CT study. *Eur Spine J* 19:1136–1144. <https://doi.org/10.1007/s00586-009-1257-5>
 33. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33:159–174. <https://doi.org/10.2307/2529310>
 34. Portney LG, Watkins MP (2009) Foundations of clinical research applications to practice, 3rd edn. Pearson Education Inc, New Jersey
 35. Fritz CO, Morris PE, Richler JJ (2012) Effect size estimates: current use, calculations, and interpretation. *J Exp Psychol Gen* 141:2–18. <https://doi.org/10.1037/a0024338>
 36. Fluss R, Faraggi D, Reiser B (2005) Estimation of the Youden index and its associated cutoff point. *Biom J* 47:458–472. <https://doi.org/10.1002/bimj.200410135>
 37. Eliasziw M, Young SL, Woodbury MG, Fryday-Field K (1994) Statistical methodology for the concurrent assessment of interrater and intrarater reliability: using goniometric measurements as an example. *Phys Ther* 74:777–788
 38. Healey JF (2007) The essentials of statistics: a tool for social research. Nelson Education, Toronto. <https://doi.org/10.1179/107735201800339434>
 39. Erdfelder E, Faul F, Buchner A (1996) GPower: a general power analysis program. *Behav Res Methods Instrum Comput* 1996(28):1–11. <https://doi.org/10.3758/BF03203630>
 40. Teraguchi M, Yoshimura N, Hashizume H et al (2014) Prevalence and distribution of intervertebral disc degeneration over the entire spine in a population-based cohort: the Wakayama Spine Study. *Osteoarthritis Cartil* 22:104–110. <https://doi.org/10.1016/j.joca.2013.10.019>
 41. Yu HJ, Bahri S, Gardner V, Muftuler LT (2015) In vivo quantification of lumbar disc degeneration: assessment of ADC value using a degenerative scoring system based on Pfirrmann framework. *Eur Spine J* 24(11):2442–2448. <https://doi.org/10.1007/s00586-014-3721-0>
 42. Weiler C, Lopez-Ramos M, Mayer HM et al (2011) Histological analysis of surgical lumbar intervertebral disc tissue provides evidence for an association between disc degeneration and increased body mass index. *BMC Res Notes* 4(1):497. <https://doi.org/10.1186/1756-0500-4-497>
 43. Colloca CJ, Gunzburg R, Freeman BJ et al (2012) Biomechanical quantification of pathologic manipulable spinal lesions: an in vivo ovine model of spondylolysis and intervertebral disc degeneration. *J Manip Physiol Ther* 35:354–366. <https://doi.org/10.1016/j.jmpt.2012.04.018>
 44. Vietra-Peltenz F, Olivia-Pascual-Vaca A, Rodriguez-Blanco C et al (2014) Short-term effect of spinal manipulation on pain perception, spinal mobility, and full height recovery in male subjects with degenerative disk disease: a randomized controlled trial. *Arch Phys Med Rehabil* 95:1613–1619. <https://doi.org/10.1016/j.apmr.2014.05.002>
 45. Albert HB, Lambert P, Rollason J et al (2013) Does nuclear tissue infected with bacteria following disc herniations lead to Modic changes in the adjacent vertebrae? *Eur Spine J* 22:690–696. <https://doi.org/10.1007/s00586-013-2674-z>
 46. Samartzis D, Liebenberg EC, Fong DY (2014) Innervation of pathologies in the lumbar vertebral end plate and intervertebral disc. *Spine J* 14:513–521. <https://doi.org/10.1016/j.spine.2013.06.075>
 47. Dankaerts W, O’Sullivan PB, Straker LM et al (2006) The inter-examiner reliability of a classification method for non-specific chronic low back pain patients with motor control impairment. *Man Ther* 11:28–39. <https://doi.org/10.1016/j.math.2005.02.001>
 48. Delitto A (2005) Research in low back pain: time to stop seeking the elusive “magic bullet”. *Phys Ther* 85:206–208
 49. McCarthy CJ, Arnall FA, Strimpakos N et al (2013) The biopsychosocial classification of non-specific low back pain: a systematic review. *Phys Ther Rev* 9:17–30. <https://doi.org/10.1179/108331904225003955>
 50. O’Sullivan P (2005) Diagnosis and classification of chronic low back pain disorders: maladaptive movement and motor control impairments as underlying mechanism. *Man Ther* 10:242–255. <https://doi.org/10.1016/j.math.2005.07.001>
 51. McKenzie RA, May S (2003) Mechanical diagnosis and therapy: the lumbar spine, 2nd edn. Spinal Publications, Waikanae. <https://doi.org/10.2519/jospt.2004.34.3.105>
 52. Waddell G, Newton M, Henderson I et al (1993) A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain* 52:157–168
 53. Mondloch M (2001) Does how you do depend on how you think you’ll do? A systematic review of the evidence for a relation between patients’ recovery expectations and health outcomes. *CMAJ* 165:174–179
 54. Bialosky JE, Bishop MD, Price DD et al (2009) The mechanisms of manual therapy in the treatment of musculoskeletal pain: a comprehensive model. *Man Ther* 14:531–538. <https://doi.org/10.1016/j.math.2008.09.001>