



Paraspinal Muscle in Chronic Low Back Pain: Comparison Between Standard Parameters and Chemical Shift Encoding-Based Water–Fat MRI

Nico Sollmann, MD, PhD,^{1,2,3,4*}  Noah B. Bonnheim, PhD,⁵ Gabby B. Joseph, PhD,¹ Ravi Chachad, MSc,¹ Jiamin Zhou, MSc,¹ Zehra Akkaya, MD,¹ Amir M. Pirmoazen, MD,¹  Jeannie F. Bailey, PhD,⁵ Xiaojie Guo, MPH,⁵ Ann A. Lazar, PhD,⁶ Thomas M. Link, MD, PhD,¹ Aaron J. Fields, PhD,⁵ and Roland Krug, PhD¹

Background: Paraspinal musculature (PSM) is increasingly recognized as a contributor to low back pain (LBP), but with conventional MRI sequences, assessment is limited. Chemical shift encoding-based water–fat MRI (CSE-MRI) enables the measurement of PSM fat fraction (FF), which may assist investigations of chronic LBP.

Purpose: To investigate associations between PSM parameters from conventional MRI and CSE-MRI and between PSM parameters and pain.

Study Type: Prospective, cross-sectional.

Population: Eighty-four adults with chronic LBP (44.6 ± 13.4 years; 48 males).

Field Strength/Sequence: 3-T, T1-weighted fast spin-echo and iterative decomposition of water and fat with echo asymmetry and least squares estimation sequences.

Assessment: T1-weighted images for Goutallier classification (GC), muscle volume, lumbar indentation value, and muscle-fat index, CSE-MRI for FF extraction (L1/2–L5/S1). Pain was self-reported using a visual analogue scale (VAS). Intra- and/or interreader agreement was assessed for MRI-derived parameters.

Statistical Tests: Mixed-effects and linear regression models to 1) assess relationships between PSM parameters (entire cohort and subgroup with GC grades 0 and 1; statistical significance $\alpha = 0.0025$) and 2) evaluate associations of PSM parameters with pain ($\alpha = 0.05$). Intraclass correlation coefficients (ICCs) for intra- and/or interreader agreement.

Results: The FF showed excellent intra- and interreader agreement (ICC range: 0.97–0.99) and was significantly associated with GC at all spinal levels. Subgroup analysis suggested that early/subtle changes in PSM are detectable with FF but not with GC, given the absence of significant associations between FF and GC (*P*-value range: 0.036 at L5/S1 to 0.784 at L2/L3). Averaged over all spinal levels, FF and GC were significantly associated with VAS scores.

View this article online at wileyonlinelibrary.com. DOI: 10.1002/jmri.28145

Received Dec 6, 2021, Accepted for publication Feb 25, 2022.

*Address reprint requests to: N.S., 185 Berry Street, San Francisco, CA 94107. E-mail: nico.sollmann@tum.de

Funding information This research was supported by the National Institutes of Health (NIH) under the award number P30AR075055 and through the NIH HEAL Initiative under award numbers U19-AR076737 and UH2-AR076719 as well as by the German Academic Exchange Service (Deutscher Akademischer Austauschdienst, DAAD: N.S.), the Joachim Herz Foundation (N.S.), and the Rolf W. Günther Foundation (N.S.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or its NIH HEAL Initiative.

From the ¹Department of Radiology and Biomedical Imaging, University of California San Francisco, San Francisco, California, USA; ²Department of Diagnostic and Interventional Radiology, University Hospital Ulm, Ulm, Germany; ³Department of Diagnostic and Interventional Neuroradiology, School of Medicine, Klinikum rechts der Isar, Technical University of Munich, Munich, Germany; ⁴TUM-Neuroimaging Center, Klinikum rechts der Isar, Technical University of Munich, Munich, Germany; ⁵Department of Orthopaedic Surgery, University of California San Francisco, San Francisco, California, USA; and ⁶Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, California, USA

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Data Conclusion: In the absence of FF, GC may be the best surrogate for PSM quality. Given the ability of CSE-MRI to detect muscle alterations at early stages of PSM degeneration, this technique may have potential for further investigations of the role of PSM in chronic LBP.

Level of Evidence: 2

Technical Efficacy Stage: 2

J. MAGN. RESON. IMAGING 2022;56:1600–1608.

Low back pain (LBP) is the most common and costly musculoskeletal condition and may be experienced by individuals of all ages.¹ The global point prevalence of activity-limiting LBP was 7.3% in 2015, with more than 500 million people affected worldwide.² The management of LBP is challenged by the fact that structural causes can only be identified in a minority of affected patients, while large numbers of patients are diagnosed with nonspecific LBP.³

When imaging beyond simple radiographs is clinically indicated, cross-sectional imaging by MRI is frequently used.^{4–6} It is the modality of choice compared to computed tomography (CT), mostly due to its high soft tissue contrast and radiation-free acquisition technique. Furthermore, MRI provides cross-sectional images that can assess multiple spinal structures including intervertebral discs, nerve roots, facet joints, and paraspinal musculature (PSM).^{4,5,7} While many MRI findings appear with similar prevalence in subjects with and without LBP, some have higher prevalence in patients with LBP (such as vertebral endplate lesions, disc bulges and herniations, or spondylolysis).^{1,6,8} A structure that is increasingly recognized as a potential contributor to LBP is the PSM, but findings from imaging are inconclusive and the distinct role of PSM in LBP is unclear.^{9–11} One potential reason is that conventional T1- and T2-weighted MRI sequences limit analysis to mostly qualitative morphologic assessment and may not provide quantitative or objective information on PSM composition. A more robust characterization of PSM by using advanced quantitative MRI could potentially help elucidate the role of PSM in LBP.

A number of image-based parameters have been explored as potential noninvasive biomarkers for PSM. For example, CT attenuation of the multifidus and erector spinae muscles from regions of interest (ROIs) placed in the center of the muscle mass have been compared between patients with and without LBP, but no significant differences have been found.¹² Furthermore, a number of MRI-based studies have extracted parameters mainly from conventional axial T1- or T2-weighted sequences, such as the muscle volume or cross-sectional area (CSA), signal intensity (SI), semiquantitative scoring using the Goutallier classification (GC), and recently, the lumbar indentation value (LIV).^{13–21} Muscle volume, CSA, and LIV are measures of muscle geometry or size that may show alterations in patients with LBP.^{16,19–21} In contrast, the semiquantitative GC scoring system and the muscle-fat index (MFI), which is calculated by dividing the

mean SI in a specific muscle by the SI in a homogenous region of fat outside the muscle, have been used to evaluate fatty degeneration of PSM and, thus, may reflect estimates of PSM quality in LBP.^{13,17,19,21} A limitation of PSM assessment using conventional T1- and T2-weighted imaging is that parameters derived from these sequences may not capture subtle changes in muscle tissue composition. Furthermore, semiquantitative grading depends on the reader's judgment and experience, and application of such grading systems across different MRI systems with nonharmonized pulse sequence protocols may be prone to bias.

One promising technique to quantitatively and objectively assess tissue composition is chemical shift encoding-based water–fat MRI (CSE-MRI).^{22,23} Specifically, CSE-MRI can provide an accurate measure of the fat fraction (FF) of PSM.^{24–26} A recent study used CSE-MRI to demonstrate that cartilage endplate damage at level L4/L5 was predictive of patient-reported pain and disability when adjacent to PSM with greater FF.²⁴ However, associations between the FF of PSM and parameters derived from conventional MRI (such as GC, muscle volume, LIV, and MFI) remain unclear in patients with chronic LBP, obscuring our understanding of how alterations in PSM composition relate to pain. Thus, the purpose of this study was to investigate the associations between PSM parameters derived from conventional MRI (GC, muscle volume, LIV, and MFI) and the FF as the reference standard from CSE-MRI, and to investigate whether PSM parameters were associated with self-reported pain in patients with chronic LBP.

Materials and Methods

This study and its procedures were HIPAA compliant and were approved by the institutional review board and conducted in accordance with its regulations. Written informed consent was obtained from all study participants, and the analyses were performed on de-identified data.

Study Design and Patient Cohort

This study was part of the Back Pain Consortium (BACPAC) Research Program, which is a translational and patient-centered effort to address the need for effective and personalized therapies for chronic LBP (<https://heal.nih.gov/research/clinical-research/back-pain>). A goal of BACPAC is to examine biomedical mechanisms within a biopsychosocial context by using interdisciplinary methods and to explore innovative technologies (including in-vivo imaging) to develop an integrated model of chronic LBP. The study protocol

for the BACPAC Spine Imaging Working Group includes multi-sequence lumbar MRI and questionnaires for self-report measures of pain and disability. Participant evaluations for this study were conducted between October 2019 and August 2021.

The following inclusion criteria were defined for enrollment: 1) age of ≥ 18 years and 2) chief complaint of chronic LBP (pain between inferior border of ribcage and gluteal fold) as defined by the National Institutes of Health Pain Consortium Research Task Force.²⁷ The exclusion criteria were as follows: 1) current or prior history of spine infection, spine tumor, autoimmune disorder (including ankylosing spondylitis, rheumatoid arthritis, polymyalgia rheumatica, psoriatic arthritis, or lupus), or cancer (except non-melanoma skin cancer); 2) any history of vertebral fractures, cauda equina syndrome, or radiculopathy with functional motor deficit (strength of $< 4/5$ on manual motor testing using the Medical Research Council scale²⁸); 3) life expectancy of 2 years or less, worker's compensation, personal injury litigation; 4) pregnancy; 5) prior or planned spine surgery; 6) disc herniation and leg pain ≥ 4 or $\geq 50\%$ for LBP (considering the visual analogue scale [VAS]²⁹); 7) diagnosis, based on radiographic evidence, of clinically relevant lumbar vertebral abnormalities (spondylolisthesis with more than 2 mm of translation or with pars fracture at the involved level, spondylolysis, lumbar scoliosis with a Cobb angle of greater than 15°); and 8) obesity with a body mass index (BMI) $> 40 \text{ kg/m}^2$.

Overall, a total of 84 patients with chronic LBP were prospectively enrolled and underwent multisequence MRI acquisitions for this study's purposes. Completion of the questionnaires and imaging were achieved on an outpatient basis.

Assessment of Pain

To assess self-reported pain, study participants completed the VAS.²⁹ Pain scoring by an 11-point VAS questionnaire and MRI studies were obtained on the same day.

Magnetic Resonance Imaging

Subjects underwent an MRI examination in supine position with a 3-T scanner (GE 3T Discovery MR750; GE Healthcare, Waukesha, WI, USA) using an eight-channel phased-array spine coil. A T1-weighted fast spin-echo (FSE) sequence (axial plane) and an iterative decomposition of water and fat with echo asymmetry and least squares estimation (IDEAL) sequence (axial plane) were used for this study (Table 1).

The IDEAL sequence is a commonly used CSE-MRI technique that provides robust and homogenous water-fat separation. It was processed using the vendor's routines, including phase error correction and a complex-based water-fat decomposition considering a pre-calibrated six-peak fat spectrum and a single T2*. Axial FF maps were generated from pixel-wise calculation of the ratio of the fat signal over the sum of fat and water signals.

Image Evaluation and Segmentations

Images were opened in MITK (version 2021.02; [[http://mitk.org/wiki/The_Medical_Imaging_Interaction_Toolkit_\(MITK\)](http://mitk.org/wiki/The_Medical_Imaging_Interaction_Toolkit_(MITK))]; German Cancer Research Center, Division of Medical and Biological Informatics, Medical Imaging Interaction Toolkit, Heidelberg, Germany) for multiparametric evaluation of PSM measures.

TABLE 1. Sequence Parameters

	T1-Weighted FSE	IDEAL
Plane	Axial	Axial
TE (in msec)	14.1	2, 3, 4, 5, 6, and 7
TR (in msec)	594	6.2
FOV (in cm)	18	28
Matrix	288 × 192	180 × 160
Slice thickness (in mm)	4	4
Slice spacing (in mm)	1	No slice spacing
Number of slices	42	54
Receiver bandwidth (kHz)	41.7	125
Frequency encoding direction	R/L	R/L
Signal averages (NEX)	2	1
Flip angle	90	3
Acquisition time (min:sec)	3:47	3:05

This table shows the sequence parameters for the T1-weighted FSE sequence and the IDEAL sequence, which covered the lumbar spine.

FOV = field of view; FSE = fast spin-echo; IDEAL = iterative decomposition of water and fat with echo asymmetry and least-squares estimation; TE = echo time; TR = repetition time.

For the GC and LIV determination, the axial T1-weighted FSE images were used. The semiquantitative GC provides a classification system for grading of the extent of fatty degeneration for muscles such as PSM.^{17,19,30} The GC consists of five grades: 0—normal muscle; 1—some fatty streaks; 2—less than 50% fatty muscle atrophy; 3—50% fatty muscle atrophy; and 4—greater than 50% fatty muscle atrophy (Fig. 1).³⁰ Grading was performed at consecutive axial slices centered on the intervertebral disc at each level (L1/2, L2/3, L3/4, L4/5, and L5/S1), considering the co-lateral multifidus and erector spinae muscles. In addition, the LIV, defined as the distance equal to the length of the bulge of the muscle to the attachment of the spinous process, was measured for each level (Fig. 2).^{17,19} Assessments were performed by a radiologist with 9 years of experience (N.S.).

Using information from manual segmentations, the muscle volume, MFI, and FF were extracted (Fig. 3). The FF of PSM is considered the reference standard for this study. Respective polygonal ROIs were directly drawn on the axial FF maps derived from IDEAL sequences, while the T1-weighted FSE images were opened

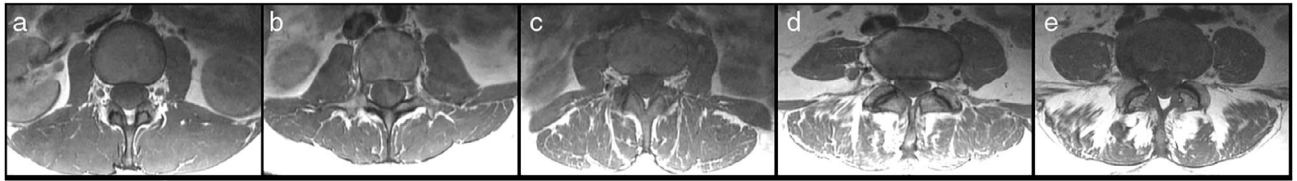


FIGURE 1: Goutallier classification (GC) for paraspinal musculature (PSM). Representative axial T1-weighted fast spin-echo (FSE) images for the different GC scores: 0—normal muscle (a); 1—some fatty streaks (b); 2—less than 50% fatty muscle atrophy (c); 3—50% fatty muscle atrophy (d); and 4—greater than 50% fatty muscle atrophy (e).

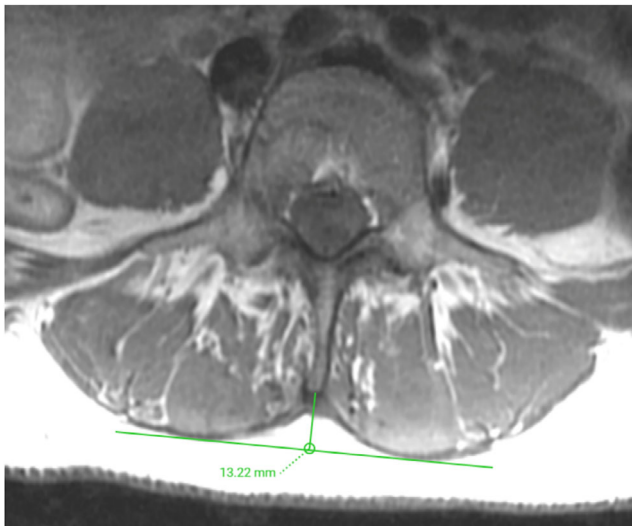


FIGURE 2: Determination of the lumbar indentation value (LIV). Representative axial T1-weighted fast spin-echo (FSE) image demonstrating the measurement of the LIV at the level of L3/L4, which amounted to 13.22 mm. The LIV is measured as the distance equal to the length of the bulge of the muscle to the attachment of the spinous process.

simultaneously for anatomical cross-correlation and depiction of muscle borders and PSM substructure. The ROIs enclosed the multifidus and erector spinae muscles of both sides in two consecutive slices centered on the intervertebral disc at each level (L1/2, L2/3, L3/4, L4/5, and L5/S1). Particular attention was paid to exclude subcutaneous fat or other structures not belonging to the muscles of interest.²⁴ The level-wise muscle volume as well as FF of multifidus and erector spinae muscles were extracted from the segmentations in the FF maps. In addition, using the level-wise segmentation masks with the axial T1-weighted FSE images, the MFI was determined. In detail, the MFI was defined as the mean SI of the target muscle (i.e., multifidus and erector spinae muscles) divided by the mean SI of homogenous subcutaneous fat.^{13,21,31,32} To extract the SI of subcutaneous fat, further ROIs were placed around the outer margins of the erector spinae muscles in the subcutaneous fat located posteriorly. The manual segmentations were performed by four readers with 9 years (N.S.), 5 years (N.B.), 3 years (R.C.), and 3 years of experience (J.Z.). Each of those readers performed placement of ROIs in one quarter of the cases (i.e., 21 subjects per reader).

Inter- and Intra-Reader Agreement

The entire dataset was evaluated by two other radiologists with 15 years (Z.A.) and 4 years (A.P.) of experience to assess interreader

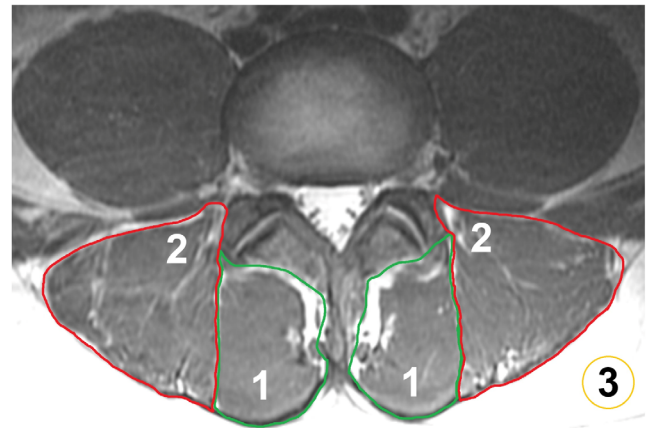


FIGURE 3: Segmentation of paraspinal musculature (PSM) and subcutaneous fat. Manual segmentations of the multifidus muscles (1) and erector spinae muscles (2) of both sides are shown at level L3/L4, together with a circular region of interest (ROI) placed in the subcutaneous fat (3). The segmentations of the multifidus and erector spinae muscles were used for the extraction of the level-wise fat fraction (FF) and muscle volume of PSM (based on the axial slices of an iterative decomposition of water and fat with echo asymmetry and least-squares estimation [IDEAL] sequence). Furthermore, segmentations of these muscles together with the ROIs in subcutaneous fat were used to calculate the muscle-fat index (MFI) as the mean signal intensity (SI) of PSM divided by the mean SI of homogenous subcutaneous fat (based on the axial slices of a T1-weighted fast spin-echo [FSE] sequence).

agreement for GC and LIV determination. Furthermore, to assess interreader agreement for segmentation-derived measures (FF, volume, and MFI of PSM), four randomly selected patient cases were manually segmented by each of the four readers who were involved in the manual segmentations, again using ROI placements at each spinal level (L1/2, L2/3, L3/4, L4/5, and L5/S1). Moreover, one reader performed measurements in these four randomly selected patients a second time to evaluate intrareader agreement (N.S.). A minimum of 4 weeks after initial segmentations was considered to minimize recall bias.

Statistical Analysis

Statistical analysis was performed using STATA (version 16; StataCorp LP, College Station, TX, USA) and Prism (version 6; GraphPad Prism, San Diego, CA, USA). Descriptive statistics were calculated for cohort characteristics, self-reported pain by VAS scores, as well as for the FF of PSM and the different other parameters derived from MRI data evaluation (GC, muscle volume, LIV,

and MFI). Correlation analyses using Spearman’s rho were calculated between the average values of FF among all spinal levels (L1–S1) and age and BMI, respectively.

Mixed-effects models (accounting for multiple measurements/levels per patient, adjusted for age, sex, and BMI) were performed to assess the relationships between predictor variables (GC, muscle volume, LIV, and MFI) and the outcome variable (FF of PSM), considering the mean value per level for GC and LIV as derived from assessments of the three readers. An interaction with spinal level was included in each model with each predictor along with the main effects (level and predictor). This analysis was performed for the whole cohort (84 patients) as well as for a subset of patients showing absent to minor fatty degeneration of PSM according to GC (grades 0 and 1; 14 patients), representing none to early degenerative changes of PSM. To adjust for multiple comparisons, we used the Bonferroni method, resulting in a level of statistical significance of $\alpha = 0.0025$. Adjusted β coefficients, P -values, and/or 95% confidence intervals (95% CIs) are reported for these models.

To assess the relationship between MRI-derived parameters (GC and FF of PSM) and self-reported pain (VAS), linear regression models were computed, with average values of GC and FF of PSM among all spinal levels as variables in the model. The models were adjusted for age, sex, and BMI, and adjusted β coefficients, P -values, and/or 95% CIs are reported. The level of statistical significance was set at $\alpha = 0.05$ for these analyses.

To investigate inter- and intrareader agreement, intraclass correlation coefficients (ICCs) were computed. By using the FF, muscle volumes, and MFI values of PSM derived from segmentations, level-wise interreader agreement was evaluated. Furthermore, using the repeated measurements by one reader in four patients, level-wise intrareader agreement was calculated. For GC and LIV, interreader agreement was evaluated by using the semiquantitative scores and distance measurements as provided by three readers.

Results

Patient Cohort

A total of 84 patients were enrolled in the study (mean age \pm standard deviation [SD]: 44.6 ± 13.4 years, range: 20–75 years, 48 males and 36 females). The average BMI was 25.5 ± 3.8 kg/m² (range: 18.9–37.8 kg/m²). The median pain score according to VAS was 7 points.

Level-Wise Variation of Fat Fractions

The FF of multifidus and erector spinae muscles was highest at L4/L5 (mean FF \pm SD: $17.8 \pm 9.1\%$, range: 3.3%–45.7%) and L5/S1 (mean FF \pm SD: $21.6 \pm 9.7\%$, range: 1.7%–41.5%; Fig. 4). The FF values were significantly different between spinal levels (Fig. 4). There was a weak to moderate correlation between the FF averaged among all spinal levels and BMI (Spearman’s rho = 0.25) and age (Spearman’s rho = 0.50), respectively.

Associations Between Fat Fractions and Standard Parameters

The FF of PSM was significantly associated with scorings of the GC for each level (Table 2). For every one-point increase

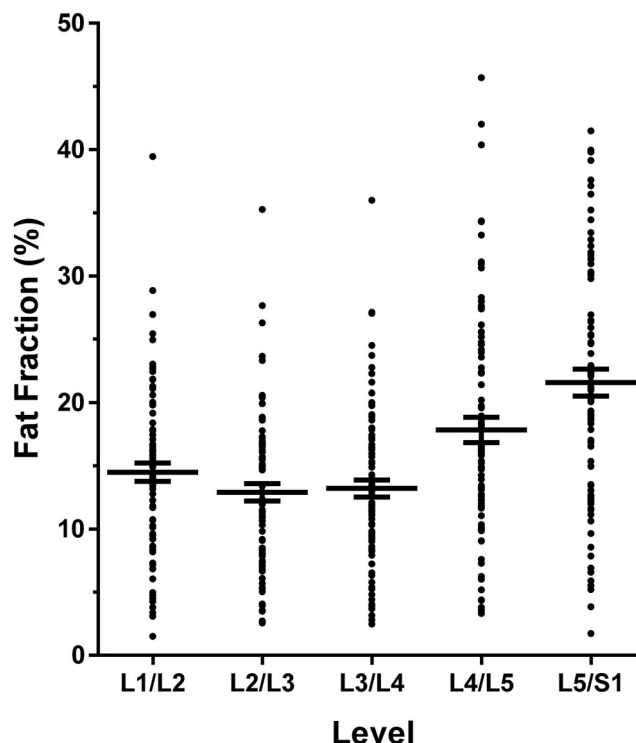


FIGURE 4: Anatomical variation of the fat fraction (FF) of paraspinal musculature (PSM). This graph shows the mean with the standard error of the mean (horizontal lines) for the level-wise FF of PSM (in %) as derived from the segmentations of the multifidus and erector spinae muscles on levels L1/L2, L2/L3, L3/L4, L4/L5, and L5/S1. The difference of the FF of PSM of all levels was significant ($\alpha = 0.05$) in relation to L5/S1 as the level with the highest FF of PSM.

in the GC, the FF increased by 3.66% at level L1/L2, 3.65% at L2/L3, 3.79% at L3/L4, 6.64% at L4/L5, and by 4.85% at L5/S1 (Table 2). For muscle volume and LIV, significant associations with the FF were only observed for one spinal level each (L5/S1 for muscle volume: β coefficient: 0.001; L1/L2 for LIV: β coefficient: 0.39; Table 2).

Patients assigned with a score of 0 according to the GC (0—normal muscle) showed a mean FF of $7.4 \pm 4.2\%$, which increased to $30.1 \pm 9.2\%$ for patients assigned with a score of 3 on the GC (3—50% fatty muscle atrophy; Fig. 5). A score of 4 on the GC (4—greater than 50% fatty muscle atrophy) was only assigned in two patients (27.6% and 26.5%).

Considering only the subcohort of patients with absent to minor fatty degeneration of PSM according to the GC (grades 0 and 1; 14 patients), none of the level-wise FF of PSM was significantly associated with the GC after adjustment for multiple comparisons (L1/L2: $p = 0.451$, L2/L3: $p = 0.784$, L3/L4: $p = 0.265$, L4/L5: $p = 0.269$, L5/S1: $p = 0.036$), while muscle volume showed a significant association with the FF of PSM for only one level (L5/S1: β coefficient: 0.0007). Neither the LIV nor MFI were significantly associated with the FF of PSM for any of the investigated spinal

TABLE 2. Associations Between the FF of PSM and Standard Parameters

	FF (Outcome)									
	L1/L2		L2/L3		L3/L4		L4/L5		L5/S1	
	β	95% CI	β	95% CI	β	95% CI	β	95% CI	β	95% CI
GC	3.66	<0.001	3.65	<0.001	3.79	<0.001	6.64	<0.001	4.85	<0.001
Volume	0.0001	0.063	0.0001	0.134	0.0001	0.110	0.0002	0.0001	0.0003	0.0008
LIV	0.39	0.001	0.32	0.011	0.23	0.054	0.27	0.036	0.20	0.119
MFI	0.03	0.766	0.09	0.345	0.13	0.175	0.20	0.010	0.13	0.056

This table lists the β coefficients, *P*-values, and 95% CIs for the associations between FF and GC, muscle volume, LIV, and MFI by spinal level (mixed-effects regression models adjusted for age, sex, and body mass index). Using correction for multiple comparisons, the level of statistical significance was set at $\alpha = 0.0025$ (statistically significant *P*-values are shown in bold). FF = fat fraction; PSM = paraspinal musculature; 95% CI = 95% confidence interval; GC = Goutallier classification; LIV = lumbar indentation value; MFI = muscle-fat index.

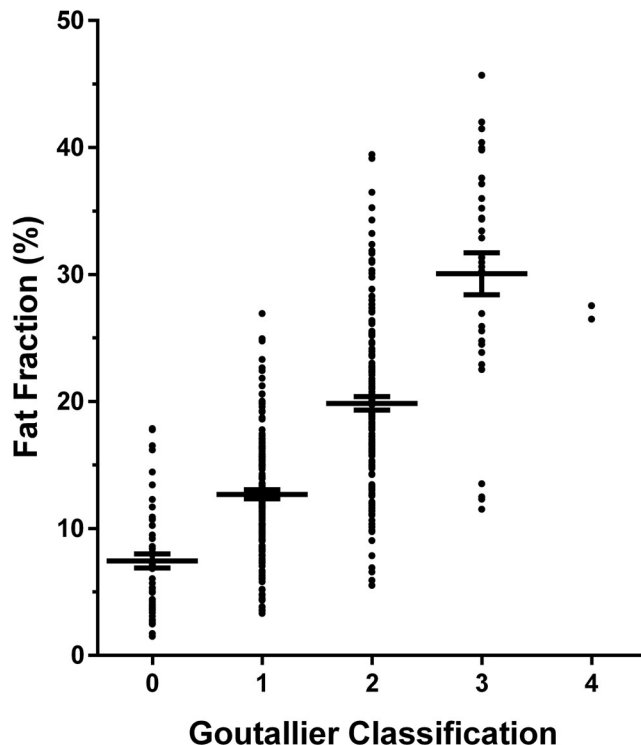


FIGURE 5: Fat fraction (FF) of paraspinal musculature (PSM) in relation to the Goutallier classification (GC). This graph shows the mean with the standard error of the mean (horizontal lines) for the FF of PSM (in %) as derived from the segmentations of the multifidus and erector spinae muscles in relation to the semiquantitative scoring according to the GC (0—normal muscle; 1—some fatty streaks; 2—less than 50% fatty muscle atrophy; 3—50% fatty muscle atrophy; and 4—greater than 50% fatty muscle atrophy).

TABLE 3. Interreader Agreement

	ICC (For Interreader Agreement)				
	L1/L2	L2/L3	L3/L4	L4/L5	L5/S1
FF	0.99	0.98	0.98	0.97	0.98
GC	0.66	0.65	0.69	0.73	0.79
Volume	0.95	0.96	0.91	0.90	0.81
LIV	0.98	0.97	0.97	0.93	0.84
MFI	0.92	0.90	0.89	0.89	0.78

This table lists the ICC for the evaluation of interreader agreement considering the FF of PSM, GC, muscle volume, LIV, and MFI by spinal level. ICC = intraclass correlation coefficient; FF = fat fraction; PSM = paraspinal musculature; GC = Goutallier classification; LIV = lumbar indentation value; MFI = muscle-fat index.

levels after adjustment for multiple comparisons (LIV: L1/L2: $p = 0.520$, L2/L3: $p = 0.771$, L3/L4: $p = 0.802$, L4/L5: $p = 0.783$, L5/S1: $p = 0.193$; MFI: L1/L2: $p = 0.820$, L2/

L3: $p = 0.908$, L3/L4: $p = 0.769$, L4/L5: $p = 0.540$, L5/S1: $p = 0.034$).

Associations Between Fat Fractions and Pain

The mean FF (for all spinal levels) was significantly associated with VAS scores (β coefficient: 7.60, 95% CI: 3.42–11.77). Similarly, the mean GC (for all spinal levels) was significantly associated with VAS scores (β coefficient: 8.59, 95% CI: 4.27–12.90).

Inter- and Intra-reader Agreement

Interreader agreement regarding the FF of PSM was excellent between readers, with the level-wise ICC ranging between 0.97 (L4/L5) and 0.99 (L1/L2; Table 3). Likewise, excellent intrareader agreement was also observed for the determination of the FF of PSM regarding all spinal levels (ICC ranging between 0.98 for L3/4 and 0.99 for all other levels). Level-wise comparison of interreader agreements for the other parameters is displayed in Table 3.

Discussion

This study's aim was to explore associations between standard parameters derived from conventional lumbar MRI using axial T1-weighted sequences (GC, muscle volume, LIV, and MFI) and the FF of PSM, which was calculated from CSE-MRI. Furthermore, we investigated associations of the GC and FF of PSM with self-reported pain in patients with chronic LBP. The main findings were as follows: 1) the FF of PSM was significantly associated with the GC for all investigated spinal levels (L1/2, L2/3, L3/4, L4/5, and L5/S1), but not in the subgroup of patients with absent to only minor fatty changes of PSM (GC grades 1 and 2); and 2) the FF and the GC of PSM were significantly associated with self-reported pain as measured by VAS.

The FF derived from CSE-MRI provides a noninvasive, quantitative, and objective marker of a tissue's relative fat content.^{22,23} The IDEAL sequence used in this study is a clinically feasible and common CSE-MRI method, which is based on a combination of the iterative least-squares approach for water-fat separation and the optimal sampling scheme under which the number of signal averages is independent of the water-fat ratios.²² For the musculoskeletal system, previous work has introduced CSE-MRI for quantification of fat, primarily investigating vertebral bodies of the lumbar spine (eg, in osteoporosis or for characterization of Modic-type endplate changes).^{33,34} More recently, axially prescribed CSE-MRI has also been applied to study the FF of PSM.^{24–26} Importantly, CSE-MRI has been validated against tissue histopathology and MR spectroscopy (MRS), with the muscle FF showing significant correlations to the amount of fat tissue from skeletal muscle biopsies as well as with those from single-voxel MRS.^{35,36} Given increasing interest in PSM related to improved phenotyping, prognosis, and treatment of

chronic LBP, CSE-MRI seems a promising method for potentially improved assessment of muscle quality.¹¹ In chronic LBP, a recent CSE-MRI study has explored the FF related to vertebral endplate pathology and pain, revealing that cartilage endplate damage at level L4/L5 was predictive of patient-reported symptoms when adjacent to PSM that showed a high FF as a potential indicator of decreased muscle quality.²⁴

Imaging by CSE-MRI is not routine in patients with chronic LBP, and the technique is not yet widely available. In the absence of CSE-MRI acquisitions, other parameters have been used to estimate PSM characteristics on clinical routine T1- or T2-weighted images in patients with chronic LBP.^{13–21} Our data suggest that the GC is the most accurate method for assessing fatty infiltration of PSM, as the GC score significantly associated with the FF of PSM from CSE-MRI. Few previous studies have used the GC for PSM assessment in patients with chronic LBP and the relationship between GC and FF of PSM remains unclear.^{17,19} The GC is a standardized and easy to apply method for seamless grading of fatty muscle degeneration based on routine T1-weighted images, and the significant associations with the FF of PSM for all lumbar levels could be regarded as evidence to use this score over even more common alternative imaging parameters such as geometric measures (muscle volume and related CSA or LIV) or SI-based measures (MFI). Furthermore, a previous study has reported a level-wise weighted kappa for intra- and interreader agreement of 0.71–0.93 and 0.76–0.85 for GC scoring of PSM for a protocol using routine T1-weighted imaging for levels L4/L5 and L5/S1 from a 1.5-T system.³⁷ In the present study, we revealed at least substantial agreement for GC scoring of the PSM based on T1-weighted FSE sequences, using a level-wise evaluation with ICCs.

In this study, the mean FF was considerably lower than the respective muscle fat description of the corresponding GC grades based on visual assessments. Similar results were previously reported for the rotator cuff muscles and the calf.^{38,39} A reasonable explanation for the observed trend toward overestimation of fat content using the GC relates to the often spatially heterogeneous muscle fat infiltration patterns, given that the GC grading based on visual reading of T1-weighted sequences cannot differentiate a voxel with an intermediate FF from voxels with high FF. Furthermore, incidental consideration of connective tissue as well as small vascular structures during visual reading and scoring may further contribute to overestimations of fat content.

Despite the significant associations between FF and GC in the entire cohort, our subgroup analysis revealed that GC grades 0 and 1 were not significantly associated with the FF, yet only including 14 patients. These findings suggest that the GC scoring system may not be able to detect subtle changes in muscle composition occurring at the early stages of muscle degeneration. However, detection of early

degenerative changes of PSM in patients with chronic LBP may have important treatment implications, as therapies targeting the PSM may provide clinical benefit and also affect other spinal pathologies.²⁴ For example, alterations in joint-level biomechanics related to alterations in PSM quality may be a critical compounding factor in painful endplate pathology.⁴⁰ Thus, a sensitive in-vivo measurement such as the FF derived from CSE-MRI may facilitate phenotyping and timely intervention in suitable patients, and it may have potential to enable the detection of nuanced compositional changes to PSM in advance of changes detectable using the GC scoring system for routine T1-weighted imaging.

We found that both FF of PSM and GC were significantly associated with scores on the VAS, supporting the concept that PSM quality plays a role in chronic LBP. However, we did not integrate the presence of known pain generators such as vertebral endplate lesions, disc bulges and herniation, and spondylolysis in our present analyses.^{1,6,8} This may preclude more definite conclusions as to the contribution of PSM composition to chronic LBP. In this regard, several mechanisms have been proposed for the pathophysiology of PSM, including reflex inhibition, denervation, modified motor control, disuse/muscle unloading mechanisms, and recently, also inflammatory mechanisms.¹¹ One task of the BACPAC Research Program is to develop and translate MRI-based methods that provide accurate and standardized noninvasive measures of chronic LBP, thus targeting a multiparametric model incorporating both conventional and advanced quantitative MRI for a better understanding of chronic LBP and improved patient selection. Associations of the FF and GC with self-reported pain motivate implementation of measures of PSM quality in such models.

Limitations

For the determination of the MFI, we used a clinically feasible approach by directly measuring the SI from PSM and subcutaneous fat. However, the extraction of values from conventional T1-weighted images to estimate fatty muscle degeneration has not yet been standardized and has been performed with various approaches.^{13,21,31,32} This hinders direct comparisons between studies, and thresholding methods for T1-weighted images might potentially result in stronger associations of metrics from T1-weighted images with the FF from CSE-MRI. Furthermore, the finding of associations between the FF and GC for PSM might not be generalizable to all patients with LBP, given that the cohort of this study was selective for nonspecific chronic LBP without common findings such as disc herniation with nerve compression and sciatica or lumbar vertebral abnormalities (eg, spondylolisthesis), often entailing altered biomechanical loads at the spine.

Conclusion

This study used lumbar MRI to derive parameters representative of PSM quality from conventional axial T1-weighted imaging (GC, muscle volume, LIV, and MFI) and CSE-MRI (FF) in patients with chronic LBP. The FF of PSM was significantly associated with the GC for all levels of the lumbar spine, while muscle volume, LIV, and MFI were not. Furthermore, both the FF and grading using the GC for PSM were significantly associated with self-reported pain according to the VAS. Hence, in the absence of CSE-MRI for PSM assessment, semiquantitative GC grading may be the best surrogate.

Acknowledgment

The authors acknowledge Ronit Gupta for his help in data preparation that contributed to the results presented in this work.

Open access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

References

- Hartvigsen J, Hancock MJ, Kongsted A, et al. What low back pain is and why we need to pay attention. *Lancet* 2018;391(10137):2356-2367.
- GBD 2015 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: A systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016;388(10053):1545-1602.
- Maher C, Underwood M, Buchbinder R. Non-specific low back pain. *Lancet* 2017;389(10070):736-747.
- Chou D, Samartzis D, Bellabarba C, et al. Degenerative magnetic resonance imaging changes in patients with chronic low back pain: A systematic review. *Spine* 2011;36(21 Suppl):S43-S53.
- Wassenaar M, van Rijn RM, van Tulder MW, et al. Magnetic resonance imaging for diagnosing lumbar spinal pathology in adult patients with low back pain or sciatica: A diagnostic systematic review. *Eur Spine J* 2012;21(2):220-227.
- Brinjikji W, Diehn FE, Jarvik JG, et al. MRI findings of disc degeneration are more prevalent in adults with low back pain than in asymptomatic controls: A systematic review and meta-analysis. *AJNR Am J Neuroradiol* 2015;36(12):2394-2399.
- Park MS, Moon SH, Kim TH, et al. Paraspinal muscles of patients with lumbar diseases. *J Neurol Surg A Cent Eur Neurosurg* 2018;79(4):323-329.
- Brinjikji W, Luetmer PH, Comstock B, et al. Systematic literature review of imaging features of spinal degeneration in asymptomatic populations. *AJNR Am J Neuroradiol* 2015;36(4):811-816.
- Ranger TA, Cicuttini FM, Jensen TS, et al. Are the size and composition of the paraspinal muscles associated with low back pain? A systematic review. *Spine J* 2017;17(11):1729-1748.
- Goubert D, Oosterwijck JV, Meeus M, Danneels L. Structural changes of lumbar muscles in non-specific low back pain: A systematic review. *Pain Physician* 2016;19(7):E985-E1000.

11. Hodges PW, Bailey JF, Fortin M, Battie MC. Paraspinal muscle imaging measurements for common spinal disorders: Review and consensus-based recommendations from the ISSLS degenerative spinal phenotypes group. *Eur Spine J* 2021;30(12):3428-3441.
12. Kalichman L, Hodges P, Li L, Guermazi A, Hunter DJ. Changes in paraspinal muscles and their association with low back pain and spinal degeneration: CT study. *Eur Spine J* 2010;19(7):1136-1144.
13. D'Hooge R, Cagnie B, Crombez G, Vanderstraeten G, Dolphens M, Danneels L. Increased intramuscular fatty infiltration without differences in lumbar muscle cross-sectional area during remission of unilateral recurrent low back pain. *Man Ther* 2012;17(6):584-588.
14. Hebert JJ, Kjaer P, Fritz JM, Walker BF. The relationship of lumbar multifidus muscle morphology to previous, current, and future low back pain: A 9-year population-based prospective cohort study. *Spine* 2014;39(17):1417-1425.
15. Hildebrandt M, Fankhauser G, Meichtry A, Luomajoki H. Correlation between lumbar dysfunction and fat infiltration in lumbar multifidus muscles in patients with low back pain. *BMC Musculoskelet Disord* 2017;18(1):12.
16. Sions JM, Elliott JM, Pohlig RT, Hicks GE. Trunk muscle characteristics of the multifidi, erector spinae, psoas, and quadratus lumborum in older adults with and without chronic low back pain. *J Orthop Sports Phys Ther* 2017;47(3):173-179.
17. Tamai K, Chen J, Stone M, et al. The evaluation of lumbar paraspinal muscle quantity and quality using the Goutallier classification and lumbar indentation value. *Eur Spine J* 2018;27(5):1005-1012.
18. Teichtahl AJ, Urquhart DM, Wang Y, et al. Fat infiltration of paraspinal muscles is associated with low back pain, disability, and structural abnormalities in community-based adults. *Spine J* 2015;15(7):1593-1601.
19. Virk S, Wright-Chisem J, Sandhu M, et al. A novel magnetic resonance imaging-based lumbar muscle grade to predict health-related quality of life scores among patients requiring surgery. *Spine* 2021;46(4):259-267.
20. Wan Q, Lin C, Li X, Zeng W, Ma C. MRI assessment of paraspinal muscles in patients with acute and chronic unilateral low back pain. *Br J Radiol* 2015;88(1053):20140546.
21. Goubert D, De Pauw R, Meeus M, et al. Lumbar muscle structure and function in chronic versus recurrent low back pain: A cross-sectional study. *Spine J* 2017;17(9):1285-1296.
22. Reeder SB, Pineda AR, Wen Z, et al. Iterative decomposition of water and fat with echo asymmetry and least-squares estimation (IDEAL): Application with fast spin-echo imaging. *Magn Reson Med* 2005;54(3):636-644.
23. Reeder SB, Hu HH, Sirlin CB. Proton density fat-fraction: A standardized MR-based biomarker of tissue fat concentration. *J Magn Reson Imaging* 2012;36(5):1011-1014.
24. Bailey JF, Fields AJ, Ballatori A, et al. The relationship between endplate pathology and patient-reported symptoms for chronic low back pain depends on lumbar paraspinal muscle quality. *Spine* 2019;44(14):1010-1017.
25. Sollmann N, Dieckmeyer M, Schlaeger S, et al. Associations between lumbar vertebral bone marrow and paraspinal muscle fat compositions— an investigation by chemical shift encoding-based water-fat MRI. *Front Endocrinol (Lausanne)* 2018;9:563.
26. Sollmann N, Zoffl A, Franz D, et al. Regional variation in paraspinal muscle composition using chemical shift encoding-based water-fat MRI. *Quant Imaging Med Surg* 2020;10(2):496-507.
27. Deyo RA, Dworkin SF, Amtmann D, et al. Report of the NIH Task Force on research standards for chronic low back pain. *J Pain* 2014;15(6):569-585.
28. Medical Research Council. Aids to the examination of the peripheral nervous system, memorandum no. 45. London: Her Majesty's Stationery Office; 1981.
29. Langley GB, Sheppard H. The visual analogue scale: Its use in pain measurement. *Rheumatol Int* 1985;5(4):145-148.
30. Goutallier D, Postel JM, Bernageau J, Lavau L, Voisin MC. Fatty muscle degeneration in cuff ruptures. Pre- and postoperative evaluation by CT scan. *Clin Orthop Relat Res* 1994;304:78-83.
31. Fu CJ, Chen WC, Lu ML, Cheng CH, Niu CC. Comparison of paraspinal muscle degeneration and decompression effect between conventional open and minimal invasive approaches for posterior lumbar spine surgery. *Sci Rep* 2020;10(1):14635.
32. Hyun SJ, Kim YJ, Rhim SC. Patients with proximal junctional kyphosis after stopping at thoracolumbar junction have lower muscularity, fatty degeneration at the thoracolumbar area. *Spine J* 2016;16(9):1095-1101.
33. Sollmann N, Loffler MT, Kronthaler S, et al. MRI-based quantitative osteoporosis imaging at the spine and femur. *J Magn Reson Imaging* 2021;54:12-35.
34. Fields AJ, Battie 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* 2019;28(10):2266-2274.
35. Fischer MA, Nanz D, Shimakawa A, et al. Quantification of muscle fat in patients with low back pain: Comparison of multi-echo MR imaging with single-voxel MR spectroscopy. *Radiology* 2013;266(2):555-563.
36. Güttsches AK, Rehmann R, Schreiner A, Rohm M, Forsting J, Froeling M, Tegenthoff M, Vorgerd M, Schlaefke L. Quantitative muscle-MRI correlates with histopathology in skeletal muscle biopsies. *J Neuromuscul Dis* 2021;8(4):669-678.
37. Battaglia PJ, Maeda Y, Welk A, Hough B, Kettner N. Reliability of the Goutallier classification in quantifying muscle fatty degeneration in the lumbar multifidus using magnetic resonance imaging. *J Manipulative Physiol Ther* 2014;37(3):190-197.
38. Alizai H, Nardo L, Karampinos DC, et al. Comparison of clinical semi-quantitative assessment of muscle fat infiltration with quantitative assessment using chemical shift-based water/fat separation in MR studies of the calf of post-menopausal women. *Eur Radiol* 2012;22(7):1592-1600.
39. Nardo L, Karampinos DC, Lansdown DA, et al. Quantitative assessment of fat infiltration in the rotator cuff muscles using water-fat MRI. *J Magn Reson Imaging* 2014;39(5):1178-1185.
40. Fields AJ, Liebenberg EC, Lotz JC. Innervation of pathologies in the lumbar vertebral end plate and intervertebral disc. *Spine J* 2014;14(3):513-521.