



# Texture analysis of paraspinal musculature in MRI of the lumbar spine: analysis of the lumbar stenosis outcome study (LSOS) data

Manoj Mannil<sup>1</sup> · Jakob M. Burgstaller<sup>2</sup> · Arjun Thanabalasingam<sup>1</sup> · Sebastian Winklhofer<sup>3</sup> · Michael Betz<sup>4</sup> · Ulrike Held<sup>2</sup> · Roman Guggenberger<sup>1</sup>

Received: 8 December 2017 / Revised: 12 February 2018 / Accepted: 16 February 2018 / Published online: 1 March 2018  
© ISS 2018

## Abstract

**Objective** To evaluate association of fatty infiltration in paraspinal musculature with clinical outcomes in patients suffering from lumbar spinal stenosis (LSS) using qualitative and quantitative grading in magnetic resonance imaging (MRI).

**Materials and methods** In this retrospective study, texture analysis (TA) was performed on postprocessed axial T2 weighted (w) MR images at level L3/4 using dedicated software (MaZda) in 62 patients with LSS. Associations in fatty infiltration between qualitative Goutallier and quantitative TA findings with two clinical outcome measures, Spinal stenosis measure (SSM) score and walking distance, at baseline and regarding change over time were assessed using machine learning algorithms and multiple logistic regression models.

**Results** Quantitative assessment of fatty infiltration using the histogram TA feature “mean” showed higher interreader reliability (ICC 0.83–0.97) compared to the Goutallier staging ( $\kappa = 0.69–0.93$ ). No correlation between Goutallier staging and clinical outcome measures was observed. Among 151 TA features, only TA feature “mean” of the spinotransverse group showed a significant but weak correlation with worsened SSM ( $p = 0.046$ ). TA feature “S(3,3) entropy” showed a significant but weak association with worsened WD over 12 months ( $p = 0.046$ ).

**Conclusion** MR TA is a reproducible tool to quantitatively assess paraspinal fatty infiltration, but there is no clear association with the clinical outcome in asymptomatic LSS patients.

**Keywords** Magnetic resonance imaging · Texture analysis · LSOS · Lumbar stenosis · Paraspinal musculature · Fat quantification

## Abbreviations

CSA	cross-sectional surface area
FID	free induction decay
GLCM	gray level co-occurrence
ICC	intraclass correlation coefficients

ROC	receiver operating characteristic
TSE	turbo spin echo
w	weighted
AUC	area under the curve
BMI	body mass index
LSOS	Lumbar Stenosis Outcome Study
LSS	lumbar spinal stenosis
MRI	magnetic resonance imaging
ROI	regions of interest
SSM	spinal stenosis measure
TA	texture analysis
WD	walking distance

✉ Roman Guggenberger  
roman.guggenberger@usz.ch

<sup>1</sup> Institute of Diagnostic and Interventional Radiology, University Hospital Zurich, University of Zurich, Raemistrasse 100, CH-8091 Zurich, Switzerland

<sup>2</sup> Horten Centre for Patient Oriented Research and Knowledge Transfer, University of Zurich, Pestalozzistrasse 24, 8032 Zurich, Switzerland

<sup>3</sup> Department of Neuroradiology, University Hospital Zurich, University of Zurich, Frauenklinikstrasse 10, 8091 Zurich, Switzerland

<sup>4</sup> Department of Orthopaedics, Balgrist University Hospital University of Zurich, Forchstrasse 340, 8008 Zurich, Switzerland

## Introduction

Lumbar spinal stenosis (LSS) is the most frequent indication for spine surgery in patients older than 65 years [1]. Lower back pain constitutes one of the main contributing factors to

disability and absence from work in industrialized countries [2, 3]. In healthy, asymptomatic individuals, magnetic resonance imaging (MRI) may visualize disc degeneration or nerve root compression, spondylolysis, spondylolisthesis, and spinal stenosis [4]; however, these findings neither predict the risk of lower back pain, nor correlate with clinical symptoms [5, 6]. Even in symptomatic patients, the presence of different types of abnormalities on MRI demonstrates little correlation with self-reported pain and negligible effect on patient care or outcome [7].

The lack of correlation between pathology and pain is largely argued to be due to the multifactorial nature of the reported pain. Paraspinal musculature plays an important role in the stability and functional movements of the lumbar vertebral column and its denervation and atrophy have been previously reported in patients with lumbar disc herniation [8] or lumbar spinal stenosis (LSS) [9]. Fatty infiltration of trunk musculature is also a marker for degenerative disc and facet joint disease [10]. It has been shown that paraspinal muscle mass decreases while the proportion of fat increases with age [11]. These changes influence muscular microarchitecture and potentially impact the clinical outcome.

MRI is able to provide non-invasive information on muscle status, e.g., cross-sectional surface area (CSA), edema, or fatty infiltration [12]; however, it was recently shown that the latter can only be assessed qualitatively with moderate reproducibility [13] and negligible impact on clinical outcome [14].

In order to increase reproducibility and objectivity of medical imaging findings, recent efforts have focused on methods to quantify herein contained information regarding pixel distribution [15]. Texture analysis (TA) has emerged as a viable tool for the detection of distinct quantifiable differences of tissues that cannot be depicted by a qualitative visual assessment [16]. So far, TA in medical imaging of musculature is mainly limited to the differentiation between neurogenic and myogenic diseases [17]; however, it has been successfully applied in other fields, such as neuroimaging [18, 19], oncology [20, 21], and in skeletal applications [22, 23]. The aim of this study was to apply TA on paraspinal musculature in symptomatic LSS patients and correlate the findings with clinical outcome measures.

## Materials and methods

### Study participants

Study approval was granted by the local institutional review board and ethical committee (*Number blinded for review*). Sixty-two patients of the multi-centric Lumbar Stenosis Outcome Study (LSOS) were prospectively enrolled for this study. Additional background information about LSOS has been published previously [24–26]. Patients with the

diagnosis of LSS, neurogenic claudication, available magnetic resonance imaging examination, and an absence of vertebral fractures, edema, spine infection, scoliosis ( $>15^\circ$ ), or clinically relevant peripheral artery occlusive disease, were included.

As surgical interventions are known to risk ischemia and denervation of the paraspinal musculature [27], patients who had undergone prior surgical treatment were excluded from this trial. To assess the predictive value of the texture features in comparison to clinical outcome scores, patients who underwent surgery up to 12 months after the MRI examination were excluded as well. Similarly, patients who had undergone epidural injections up to 3 months prior or 12 months after the MRI examination were excluded from further analysis.

### Clinical outcome measures

All included patients were clinically assessed by use of two widely used clinical endpoints for LSS: First, every patient completed the spinal stenosis measure (SSM) questionnaire at baseline, and after 12 months. The SSM questionnaire is a LSS-specific clinical score and was used in its severity of symptoms subscale (SSM symptom severity scale). Each item is rated on a Likert scale (range 1–5, best–worst). Response options on the SSM symptom severity scale range from “no symptoms” to “very severe symptoms” [28, 29].

Second clinical outcome measure used at baseline, and after 12 months was walking distance (WD), which was extracted from the SSM function scale. Patients were assessed by their maximum WD on a plane terrain before stopping and resting. The distance in meters (m) was binned into four categories: (1)  $< 15$  m, (2) 15–200 m, (3) 200–3,000 m, and (4)  $> 3,000$  m. To account for change over time, SSM and WD were each dichotomized into two categories: (1) (any) worsening over 12 months, and (2) unchanged or improved outcome over 12 months.

### MR imaging

MR imaging in supine position was performed in all patients included in the LSOS. Due to the nature of a multicentre study with multiple participating institutions, imaging was performed on different MRI scanners with field strengths at 1.5 and 3 T with varying scanning parameters. However, certain quality criteria had to be fulfilled by all participating centres, which included standard protocols—sagittal T1w, sagittal T2w, and axial T2w turbo spin echo (TSE) images—and quality control by a board-certified radiologist. All imaging data from the participating clinical centres were collected at one place and saved on a picture archiving and communication system (PACS, IMPAX 6; AGFA Healthcare, Mortsels, Belgium).

For further analysis, we chose morphology-based axial 2D T2w TSE images with 3.5-mm slice thickness perpendicular

to the intervertebral disc at level L3/4 directly acquired with a standard matrix of  $256 \times 256$ . An axial image was chosen to average asymmetric paraspinal muscular findings between both sides. The level L3/4 was chosen as it showed on average most fatty infiltration with least susceptibility to scoliotic changes. For each individual, a single image in the axial plane at this level was saved in DICOM format and postprocessed for image normalization, pixel spacing and field of view.

### Qualitative Goutallier classification

The Goutallier classification is originally used to quantify the amount of fatty infiltration of the rotator cuff muscles, particularly in the context of rotator cuff tendon tears [30]. In musculoskeletal imaging it is widely used for the subjective assessment of fatty infiltration in skeletal muscles. The grading ranges from 0 to 4: (0) grade 0: normal muscle, (1) grade 1: some fatty streaks, (2) grade 2: less than 50% fat proportion of muscle mass, (3) grade 3: 50% fat proportion of muscle mass, and (4) grade 4: greater than 50% fat proportion of muscle mass [30]. According to Battaglia et al. [31], the Goutallier classification system can be reliably applied to quantify fatty degeneration in lumbar spinal musculature. For the qualitative readout, we graded the amount of fatty infiltration in the psoas muscles (negative control), and the paraspinal musculature—erector spinae group (longissimus and iliocostalis together), spinotransverse group (predominantly multifidus), and erector spinae and spinotransverse group combined as one region of interest [32].

Goutallier staging was performed twice for intra-observer analysis by the same reader (x.y.) blinded for review, with 4 years of experience in skeletal radiology; and by a second reader for inter-observer analysis (y.y.) who was also blinded for review, with 1 year of dedicated experience in skeletal radiology.

### Quantitative texture analysis (TA)

TA was performed in all 62 individuals using a freely available software package (MaZda, version 4.6, Institute of Electronics, Technical University of Lodz, Lodz, Poland) [33]. Normalization was performed between the mean and three standard deviations (“ $\pm 3\sigma$ ” method), which corrects small technical intra- and interscanner fluctuations [34]. Table 1 provides an overview of all TA categories and corresponding features calculated by the software.

To account for differences in spinal inclination, free-hand regions of interest (ROI) were drawn on the most horizontal portions of the lumbar spine at level L3/4. ROIs were drawn on axial T2w TSE DICOM spine MR images. ROI delineation was restricted to the aforementioned paraspinal muscle groups as recently proposed by Crawford et al. [32] (Fig. 1).

For testing the intra- and inter-observer reproducibility of the TA features, ROI delineation was repeated twice by the same reader for intra-observer (x.y.) blinded for review; and for inter-observer analysis by a second reader (y.y.) blinded for review. As several TA features require identical spatial resolution and pixel size to be comparable, images were rescaled according to the pixel spacing values obtained from the DICOM header. Rescaling was performed using an in-house MATLAB routine to the coarsest in-plane resolution of  $0.3 \times 0.3 \text{ mm}^2$ .

### Texture feature selection and feature reduction

The TA software calculated 208 features for each ROI (Table 1). Feature selection and dimension reduction was performed as follows: in a first analysis, we selected all 1st level TA features as recently described by Tabari et al. [35], including “mean”, “variance”, “skewness”, and “kurtosis”; secondly, we removed those TA features with a reduced intra- and inter-observer reproducibility, defined as intraclass correlation coefficient  $< 0.85$ ; ten, in order to best reflect total paraspinal muscle changes, the remaining features were averaged over both sides per patient. Consecutive dimension reduction was performed using a machine learning algorithm (Boruta package) [36], which uses random forest algorithms and performs a top-down search for relevant features by comparing original attributes’ importance ( $p = 0.01$ ) with importance achievable at random.

### Statistical analysis

Continuous variables were expressed as means  $\pm$  standard deviation, and categorical variables as frequencies and/or percentages. Intra- and inter-observer reproducibility of the fatty degeneration of paraspinal musculature was performed by use of Cohen’s kappa ( $\kappa$ ) statistics. Goutallier values were analyzed separately for each of the aforementioned muscle regions and sides. A  $\kappa$  value  $> 0.81$  was interpreted as excellent agreement, 0.61–0.80 as good, 0.41–0.60 as moderate, 0.21–0.40 as fair, and values  $\leq 0.20$  as poor agreement [37].

To determine the inter- and intra-reader reproducibility for the TA measurements, intraclass correlation coefficients (ICC) were calculated for each pair of variables. According to Landis and Koch, ICC of 0.61–0.80 were interpreted as substantial agreement, and 0.81–1.00 as excellent agreement [38].

The 1st-level-histogram TA feature mean accounts for the mean value of pixel intensities. We assumed higher mean histogram values for higher T2w signal intensities, as found in fatty infiltration (Fig. 2). To quantify the strength of a (linear) association between the TA feature mean and the subjective fatty infiltration of the paraspinal musculature according to Goutallier, Pearson correlation coefficients were calculated. A correlation coefficient  $\pm r$  between 0.31 and 0.5 was

**Table 1** Overview of all computed texture categories with corresponding features

Texture category	Texture feature
Histogram	Mean, variance, skewness, kurtosis
Gray-level co-occurrence matrix (GLCM) (computed for four directions [(a,0), (0,a), (a,a), (0,-a)] at five interpixel distances (a = 1–5); (6 bits/pixel)	Angular second moment, contrast, correlation, entropy, sum entropy, sum of squares, sum average, sum variance, inverse different moment, difference entropy, difference variance
Run-length matrix (RLM) (computed for four angles [vertical, horizontal, 0°, and 135°]); (6 bits/pixel)	Run-length non-uniformity, gray-level non-uniformity, long run emphasis, short run emphasis, fraction of image in runs
Absolute gradient (4 bits/pixel)	Gradient mean, variance, skewness, kurtosis, and non-zeros
Autoregressive model	Teta 1–4, sigma
Wavelet transform (calculated for seven subsampling factors (n = 1–7))	Energy of wavelet coefficients in low-frequency sub-bands, horizontal high-frequency sub-bands, vertical high-frequency sub-bands, and diagonal high-frequency sub-bands

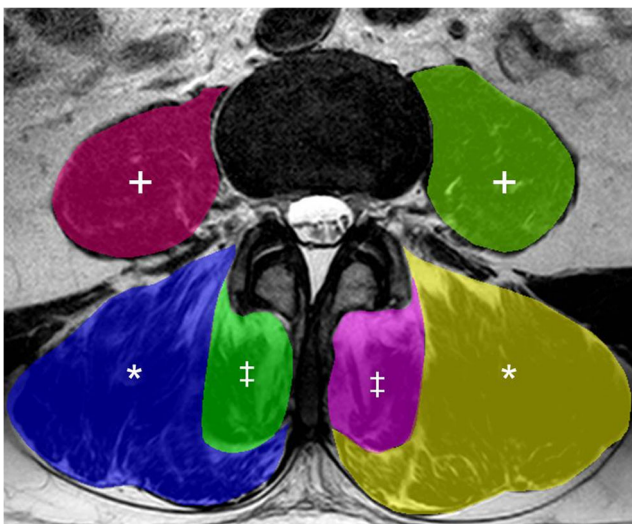
graded as weak, between 0.51 and 0.7 as moderate, and above 0.7 as strong linear relationship [39].

For significantly correlating TA features, receiver operating characteristics (ROCs) regarding tested clinical outcomes were calculated. To evaluate the likelihood that any observed differences between the Goutallier staging and each of the dichotomized two clinical outcome measures—SSM and WD—arose by chance, a Chi-squared test was performed. Additionally, a Pearson correlation was performed between the Goutallier staging and the two continuous clinical outcome measures, SSM and WD.

The impact of TA features on clinical outcomes was analyzed with two approaches. First, all first-level features were correlated with clinical endpoints by using multiple logistic regression models. Second, those TA features with good inter-

rater reliability ( $ICC \geq 0.85$ ) were preselected, and entered the automated feature selection algorithm as already described. Selected features by the machine learning algorithm were entered in (multiple) logistic regression models.

To test for statistical significance, a two-tailed *p* value below 0.05 was considered to indicate statistical significance with exception of the Boruta algorithm, where a *P* value threshold below 0.01 was used. All statistical analyses were conducted by using freely available software (R for Microsoft Windows, Version 3.2.3; <http://www.r-project.org>, The R Software Foundation, Vienna, Austria; Boruta package) [36]. Figures were edited with Adobe Illustrator CC (2015.3.1, Adobe Systems Inc., San Jose, CA).



**Fig. 1** Segmentation of paraspinal musculature in T2w axial MR images at level L3/4 for TA. Freehand-region of interest (ROI) delineation was restricted to the psoas muscles [+]; erector spinae group (longissimus and iliocostalis together) [\*], and spinotransverse muscles (predominantly multifidus) [‡] as proposed by Crawford et al. [32]

## Results

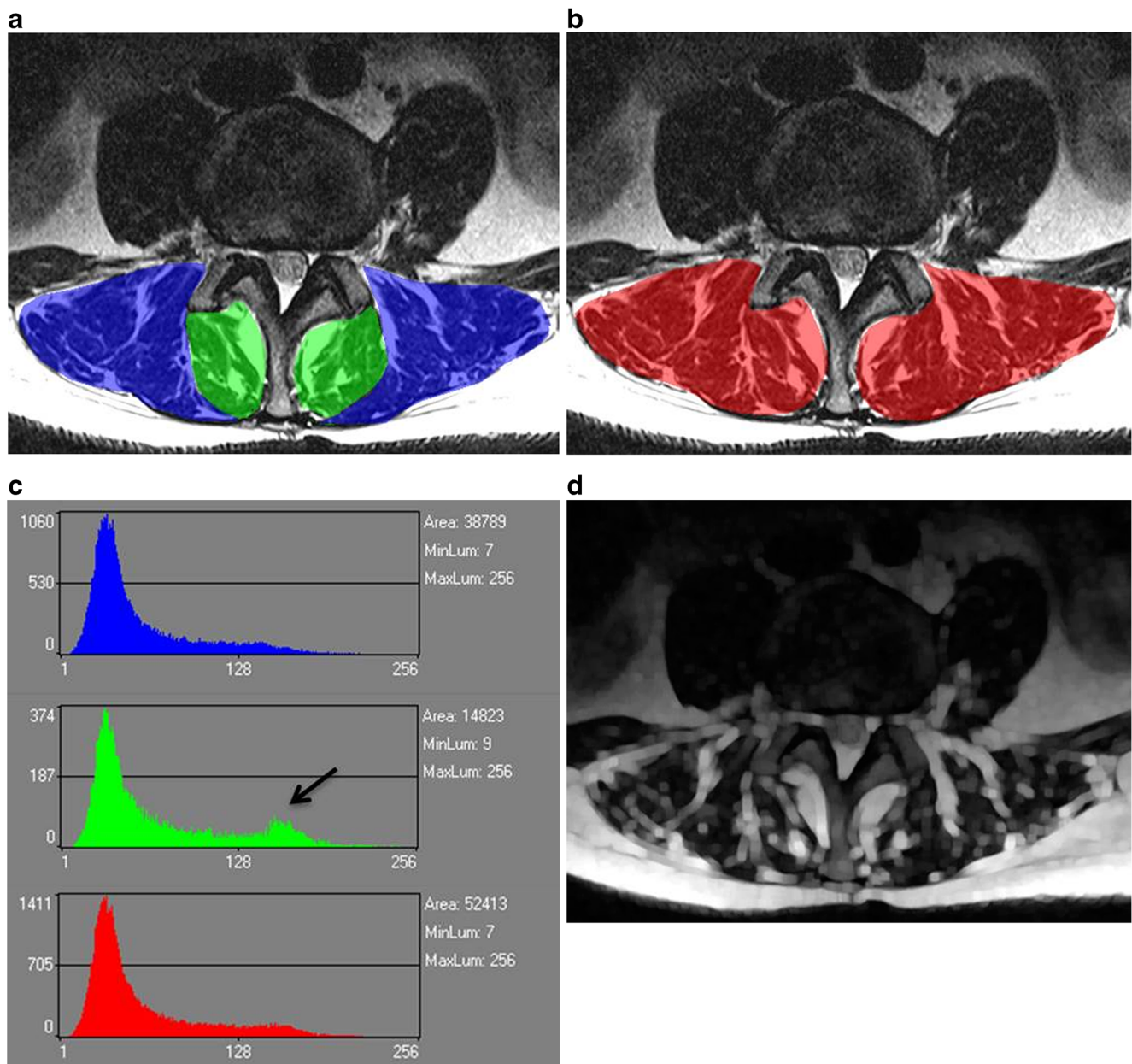
Sixty-two patients of the multi-centric LSOS were prospectively enrolled for this study. Patients' demographics are summarized in Table 2.

### Qualitative fatty infiltration assessment

We observed variability in the reproducibility of the subjective amount of fatty degeneration according to the qualitative Goutallier classification in paraspinal musculature, with  $\kappa$  values ranging from  $\kappa = 0.69$  to 1. Lowest reproducibility was observed on the overall assessment of erector spinae group and spinotransverse muscle group combined as one region of interest ( $\kappa = 0.69$ ) (Table 3).

### Quantitative fatty infiltration assessment

After exclusion of TA features with reduced ICC, 21 TA features remained for the spinotransverse group, 44 TA



**Fig. 2** **a** Region of interest segmentation for histogram analysis in paraspinal musculature in erector spinae group (blue), spinotransverse group (green), and **b** combined (red). **c** Depiction of the histogram for each of the three aforementioned groups. Note the second peak at higher pixel intensity values for the spinotransverse group (arrow),

corresponding to a higher degree of fatty infiltration. **d** Visualization of a texture map regarding the 90th percentile histogram values, corresponding to high T2w signal intensity, therefore an indirect measure of fatty infiltration

features remained for the erector spinae group, and 86 TA features remained for both groups combined for further analysis. ICC's for the quantitative first level/histogram TA feature mean showed consistently excellent reproducibility across regions and raters (Table 3). A hypothesized correlation between the 1st level histogram TA feature mean and the Goutallier classification was observed. We found a moderate correlation with coefficients ranging from  $r = 0.48 - 0.65$ .

### Correlation with clinical outcome measures

The Goutallier classification system showed no correlation with the clinical outcome measures SSM and WD dichotomized regarding change over 12 months compared to baseline ( $r = 0.09 - 0.14$ ;  $p = 0.08 - 0.87$ ). Without correcting for multiple comparison, the 1st level histogram texture feature mean of the spinotransverse musculature showed a significant correlation with dichotomized SSM change over 12 months ( $p =$

**Table 2** Descriptive Statistics of included LSOS participants

Parameter	Male	Female
<i>n</i>	33	29
Age [years]	71 ± 11	76 ± 6
BMI [kg/m <sup>2</sup> ]	29 ± 6	28 ± 7
SSM symptom score baseline	3 ± 0.8	3 ± 0.7
SSM symptom score after 12 months	2.5 ± 1	2.8 ± 1
WD baseline		
< 15 m	7 (21.2%)	12 (41.4%)
15–200 m	14 (42.4%)	12 (41.4%)
200–3,000 m	8 (24.2%)	2 (6.9%)
> 3,000 m	3 (9.1%)	2 (6.9%)
NA	1 (3%)	1 (3.4%)
WD after 12 months		
< 15 m	10 (30.3%)	9 (32.1%)
15–200 m	16 (48.5%)	10 (35.7%)
200–3,000 m	4 (12.1%)	6 (21.4%)
> 3,000 m	2 (6.1%)	3 (10.7%)
NA	1 (3%)	–

NA not available

0.046). The corresponding receiver operating characteristic (ROC) curve is depicted in Fig. 3a.

Application of the machine learning algorithm for feature reduction on the TA features with ICC ≥ 0.85 with regard to the clinical outcome measure SSM worsening after 12 months yielded no significantly contributing TA features. With regard to worsening in WD over 12 months and without adjusting for multiple comparisons, it identified a single significantly, yet weakly contributing TA feature, S(3,3) entropy ( $p = 0.046$ ). The corresponding ROC-curve is depicted in Fig. 3b.

**Table 3** Reliability analysis of qualitative and quantitative fatty infiltration

	Region	Right side	Left side
Goutallier Kappa $\kappa$	Spinotransverse group	1.00	1.00
Intrareader	Erector spinae group	1.00	1.00
	Combined	0.97	0.98
Goutallier Kappa $\kappa$	Spinotransverse group	0.93	0.88
Interreader	Erector spinae group	0.76	0.83
	Combined	0.69	0.88
TA feature mean ICC intrareader	Spinotransverse group	0.88	0.85
	Erector spinae group	0.92	0.93
	Combined	0.95	0.97
TA feature mean ICC interreader	Spinotransverse group	0.88	0.83
	Erector spinae group	0.91	0.93
	Combined	0.95	0.97

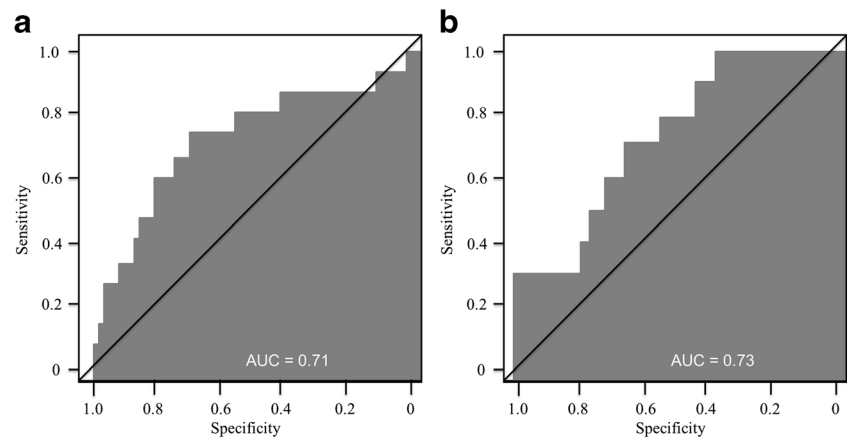
## Discussion

Our study is the first, to the best of our knowledge, to determine the influence of paraspinal musculature characteristics on clinical outcome using MR TA with machine learning algorithms in asymptomatic patients with lumbar spinal stenosis. Quantitative assessment of fatty infiltration using the histogram TA feature mean showed higher interreader reliability compared to the Goutallier. We observed no correlation between Goutallier staging and the two clinical outcome measures SSM and WD. Among 151 TA features only TA feature mean of the spinotransverse group showed a significant, yet weak correlation with worsened SSM and TA feature “S(3,3) entropy”, a measure of altered microarchitecture, showed a significant, yet weak association with worsened WD over 12 months. The moderate area under the curve (AUC) of 0.73/ 0.71 are possibly due to the multifactorial nature of LSS symptoms and the individual subjective assessment of functional impairment. These findings are in accordance with recent findings of the LSOS study group, where no relevant prognostic value of fatty degeneration on lumbar spinal stenosis treatment was observed [14].

In general, quantitative assessment of fatty infiltration in paraspinal musculature overcomes the impaired reproducibility of the qualitative Goutallier classification as recently reported by Winklhofer et al. for the entire lumbar paraspinal musculature (L1–5) [13]. We utilized the mean of a normalized histogram vector  $p(i)$ , defined as ratio between total number of pixels with gray level  $i$  and total number of pixels [40], to indirectly depict fatty infiltration in paraspinal musculature. Our approach had an overall excellent inter- and intrareader reproducibility. Park et al. [41] used a similar approach, while Gloor et al. [42] used MRI image histogram analysis to quantify fatty infiltration in muscular dystrophy. However, they used specialized sequences measuring the free induction decay (FID) signal in gradient-spoiled steady-state free precession sequences, while we performed TA on morphology-based T2w MR sequences, an integral part of standard MRI evaluation of LSS, according to the North American Spine Society [43].

In a machine learning-based approach, we first performed dimension reduction by discarding those TA features with impaired reproducibility due to minor differences in ROI delineation; however, in future studies automatic segmentation tools based on artificial neural networks might render this step redundant [44]. We observed a weak association between the gray level co-occurrence (GLCM) TA feature S(3,3) entropy and a deterioration in WD. GLCM TA features quantify joint probability pixel distributions. Entropy in particular accounts for a measure of randomness in pixel distribution. For instance, TA of a checkerboard would yield in low entropy levels. Entropy may therefore indicate clinically relevant micro-architectural changes of paraspinal musculature in

**Fig. 3** Receiver operating characteristic (ROC) curve for **a** texture feature mean and worsening of spinal stenosis measure (SSM) score after 12 months and **b** texture feature ‘S(3,3) entropy’ and worsening of walking distance (WD) after 12 months



patients with LSS that cannot be picked up by qualitative rating scales, e.g. Goutallier.

The following study limitations must be acknowledged. First, this was a retrospective study with inherent limitations. Second, we excluded individuals with recent epidural infiltrations and prior surgeries; therefore, our sample size is relatively low. Assessment of quantitative fatty infiltration was not compared to other imaging-related quantification techniques (e.g. chemical shift imaging [45]) or a histological gold standard. Moreover, quantitative assessment of fatty infiltration in paraspinal musculature was restricted to one spinal level (L3/4) and is likely limited in the presence of other T2w hyperintense findings, e.g. muscular edema. Finally, image acquisition was heterogeneous as different MR scanners were used in this multicentric study.

In conclusion, we were able to show, in a carefully selected patient cohort of asymptomatic patients with lumbar spinal stenosis without prior interventions, that MR TA serves as a viable tool to indirectly quantify paraspinal fatty infiltration with higher reproducibility than the conventional qualitative Goutallier grading system. TA might aid in detecting underlying microarchitectural changes; however, there is no clear association between paraspinal musculature TA findings with clinical outcome in asymptomatic LSS patients.

### Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

### References

- Deyo RA, Mirza SK, Martin BI, Kreuter W, Goodman DC, Jarvik JG. Trends, major medical complications, and charges associated with surgery for lumbar spinal stenosis in older adults. *JAMA*. 2010;303(13):1259–65.
- Waddell G. Low back pain: a twentieth century health care enigma. *Spine (Phila Pa 1976)*. 1996;21(24):2820–5.
- Lurie J, Tomkins-Lane C. Management of lumbar spinal stenosis. *BMJ*. 2016;352:h6234.
- Airaksinen O, Brox JI, Cedraschi C, Hildebrandt J, Klüber-Moffett J, Kovacs F, et al. Chapter 4. European guidelines for the management of chronic nonspecific low back pain. *Eur Spine J*. 2006;15 Suppl 2:S192–300.
- Savage RA, Whitehouse GH, Roberts N. The relationship between the magnetic resonance imaging appearance of the lumbar spine and low back pain, age and occupation in males. *Eur Spine J*. 1997;6(2):106–14.
- Brinjikji W, Luetmer PH, Comstock B, Bresnahan BW, Chen LE, Deyo RA, et al. Systematic literature review of imaging features of spinal degeneration in asymptomatic populations. *AJNR Am J Neuroradiol*. 2015;36(4):811–6.
- Beattie PF, Meyers SP, Stratford P, Millard RW, Hollenberg GM. Associations between patient report of symptoms and anatomic impairment visible on lumbar magnetic resonance imaging. *Spine (Phila Pa 1976)*. 2000;25(7):819–28.
- Rantanen J, Hurme M, Falck B, Alaranta H, Nykvist F, Lehto M, et al. The lumbar multifidus muscle five years after surgery for a lumbar intervertebral disc herniation. *Spine (Phila Pa 1976)*. 1993;18(5):568–74.
- Leinonen V, Maatta S, Taimela S, Herno A, Kankaanpaa M, Partanen J, et al. Impaired lumbar movement perception in association with postural stability and motor- and somatosensory-evoked potentials in lumbar spinal stenosis. *Spine (Phila Pa 1976)*. 2002;27(9):975–83.
- Sebro R, O'Brien L, Torriani M, Bredella MA. Assessment of trunk muscle density using CT and its association with degenerative disc and facet joint disease of the lumbar spine. *Skelet Radiol*. 2016;45(9):1221–6.
- 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):1136–44.
- Keller A, Gunderson R, Reikeras O, Brox JI. Reliability of computed tomography measurements of paraspinal muscle cross-sectional area and density in patients with chronic low back pain. *Spine (Phila Pa 1976)*. 2003;28(13):1455–60.
- Winklhofer S, Held U, Burgstaller JM, Finkenstaedt T, Bolog N, Ulrich N, et al. Degenerative lumbar spinal canal stenosis: intra- and inter-reader agreement for magnetic resonance imaging parameters. *Eur Spine J*. 2017;26(2):353–61.
- Betz M, Burgstaller JM, Held U, Andreisek G, Steurer J, Porchet F, et al. Influence of paravertebral muscle quality on treatment efficacy of epidural steroid infiltration or surgical decompression in lumbar spinal stenosis: analysis of the Lumbar Spinal Outcome Study (LSOS) data—a Swiss prospective multi-center cohort study.

- Spine (Phila Pa 1976). 2017;42(23):1792–8. <https://doi.org/10.1097/BRS.0000000000002233>.
15. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. 2016;278(2):563–77.
  16. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun*. 2014;5:4006.
  17. Sogawa K, Nodera H, Takamatsu N, Mori A, Yamazaki H, Shimatani Y, et al. Neurogenic and myogenic diseases: quantitative texture analysis of muscle US data for differentiation. *Radiology*. 2017;160826
  18. Hainc N, Stippich C, Stieltjes B, Leu S, Bink A. Experimental texture analysis in glioblastoma: a methodological study. *Invest Radiol*. 2017;52(6):367–73.
  19. Ingrisich M, Schneider MJ, Norenberg D, Negro de Figueiredo G, Maier-Hein K, Suchorska B, et al. Radiomic analysis reveals prognostic information in T1-weighted baseline magnetic resonance imaging in patients with glioblastoma. *Invest Radiol*. 2017;52(6):360–6.
  20. Hwang IP, Park CM, Park SJ, Lee SM, McAdams HP, Jeon YK, et al. Persistent pure ground-glass nodules larger than 5 mm: differentiation of invasive pulmonary adenocarcinomas from Preinvasive lesions or minimally invasive adenocarcinomas using texture analysis. *Invest Radiol*. 2015;50(11):798–804.
  21. Pickles MD, Lowry M, Gibbs P. Pretreatment prognostic value of dynamic contrast-enhanced magnetic resonance imaging vascular, texture, shape, and size parameters compared with traditional survival indicators obtained from locally advanced breast cancer patients. *Invest Radiol*. 2016;51(3):177–85.
  22. Rachidi M, Marchadier A, Gadois C, Lespessailles E, Chappard C, Benhamou CL. Laws' masks descriptors applied to bone texture analysis: an innovative and discriminant tool in osteoporosis. *Skelet Radiol*. 2008;37(6):541–8.
  23. Hofmann FC, Neumann J, Heilmeyer U, Joseph GB, Nevitt MC, McCulloch CE, et al. Conservatively treated knee injury is associated with knee cartilage matrix degeneration measured with MRI-based T2 relaxation times: data from the osteoarthritis initiative. *Skelet Radiol*. 2018;47(1):93–106.
  24. Ulrich NH, Burgstaller JM, Held U, Winklhofer S, Farshad M, Pichieri G, et al. The influence of single-level versus multilevel decompression on the outcome in multilevel lumbar spinal stenosis: analysis of the lumbar spinal outcome study (LSOS) data. *Clin Spine Surg*. 2017;Dec;30(10):E1367–75.
  25. Aichmair A, Burgstaller JM, Schwenkglens M, Steurer J, Porchet F, Brunner F, et al. Cost-effectiveness of conservative versus surgical treatment strategies of lumbar spinal stenosis in the Swiss setting: analysis of the prospective multicenter lumbar stenosis outcome study (LSOS). *Eur Spine J*. 2017;26(2):501–9.
  26. Burgstaller JM, Schuffler PJ, Buhmann JM, Andreisek G, Winklhofer S, Del Grande F, et al. Is there an association between pain and magnetic resonance imaging parameters in patients with lumbar spinal stenosis? *Spine (Phila Pa 1976)*. 2016;41(17):E1053–62.
  27. Bresnahan LE, Smith JS, Ogden AT, Quinn S, Cybulski GR, Simonian N, et al. Assessment of paraspinous muscle cross-sectional area following lumbar decompression: minimally invasive versus open approaches. *Clin Spine Surg*. 2017;30(3):E162–68.
  28. Stucki G, Liang MH, Fossel AH, Katz JN. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *J Clin Epidemiol*. 1995;48(11):1369–78.
  29. Stucki G, Daltroy L, Liang MH, Lipson SJ, Fossel AH, Katz JN. Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. *Spine (Phila Pa 1976)*. 1996;21(7):796–803.
  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. 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 Manip Physiol Ther*. 2014;37(3):190–7.
  32. Crawford RJ, Cornwall J, Abbott R, Elliott JM. Manually defining regions of interest when quantifying paravertebral muscles fatty infiltration from axial magnetic resonance imaging: a proposed method for the lumbar spine with anatomical cross-reference. *BMC Musculoskelet Disord*. 2017;18(1):25.
  33. Szczypinski PM, Strzelecki M, Materka A, Klepaczko A. MaZda: a software package for image texture analysis. *Comput Methods Prog Biomed*. 2009;94(1):66–76.
  34. Collewet G, Strzelecki M, Mariette F. Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. *Magn Reson Imaging*. 2004;22(1):81–91.
  35. Tabari A, Torriani M, Miller KK, Klibanski A, Kalra MK, Bredella MA. Anorexia nervosa: analysis of trabecular texture with CT. *Radiology*. 2016;160970
  36. Kurska MB, Rudnicki WR. Feature selection with the Boruta package. *J Stat Softw*. 2010;36(11):1–13.
  37. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20(1):37–46.
  38. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–74.
  39. Mukaka MM. Statistics corner: a guide to appropriate use of correlation coefficient in medical research. *Malawi Med J*. 2012;24(3):69–71.
  40. Aggarwal N, Agrawal RK. First and second order statistics features for classification of magnetic resonance brain images. *J Signal Inf Proces*. 2012;3(2):146–53.
  41. Park MJ, Cho JM, Jeon KN, Bae KS, Kim HC, Choi DS, et al. Mass and fat infiltration of intercostal muscles measured by CT histogram analysis and their correlations with COPD severity. *Acad Radiol*. 2014;21(6):711–7.
  42. Gloor M, Fasler S, Fischmann A, Haas T, Bieri O, Heinimann K, et al. Quantification of fat infiltration in oculopharyngeal muscular dystrophy: comparison of three MR imaging methods. *J Magn Reson Imaging*. 2011;33(1):203–10.
  43. North American Spine Society. Diagnosis and treatment of degenerative lumbar spinal stenosis: evidence-based clinical guidelines for multidisciplinary spine care. Burr Ridge, IL: North American Spine Society; 2011.
  44. Gassman EE, Powell SM, Kallemeyn NA, Devries NA, Shivanna KH, Magnotta VA, et al. Automated bony region identification using artificial neural networks: reliability and validation measurements. *Skelet Radiol*. 2008;37(4):313–9.
  45. Yanik B, Keyik B, Conkbayir I. Fatty degeneration of multifidus muscle in patients with chronic low back pain and in asymptomatic volunteers: quantification with chemical shift magnetic resonance imaging. *Skelet Radiol*. 2013;42(6):771–8.