European Radiology (2019) 29:22–30
https://doi.org/10.1007/s00330-018-5552-6

MUSCULOSKELETAL

Correlation of texture analysis of paraspinal musculature on MRI with different clinical endpoints: Lumbar Stenosis Outcome Study (LSOS)

Manoj Mannil[1] · Jakob M. Burgstaller[2] · Ulrike Held[2] · Mazda Farshad[3] · Roman Guggenberger[1]

Received: 15 March 2018 / Revised: 9 May 2018 / Accepted: 21 May 2018 / Published online: 14 June 2018

Abstract

Objectives The aim of this study was to apply texture analysis (TA) on paraspinal musculature in T2-weighted (T2w) magnetic resonance images (MRI) of symptomatic lumbar spinal stenosis (LSS) patients and correlate the findings with clinical outcome measures.

Methods Ninety patients were prospectively enrolled in the multi-centric Lumbar Stenosis Outcome Study (LSOS). All patients received a T2w MRI, from which we selected axial images perpendicular to the intervertebral disc at level L3/4 for TA. Regions-of-interest (ROI) were drawn of the paraspinal musculature and 304 TA features/ ROI were calculated. As clinical outcome measurements, we analysed three commonly applied measures: Spinal Stenosis Measure (SSM), Roland-Morris Disability Questionnaire (RMDQ), as well as the Numeric Rating Scale (NRS). We used two machine learning-based classifiers: Decision table, and k-nearest neighbours (k-NN).

Results We observed no meaningful correlation between TA in paraspinal musculature and the two clinical outcome measures SSM symptoms and SSM function, while a moderate correlation was observed regarding the outcome measures RMDQ (k-NN: r = 0.56) and NRS (Decision Table: r = 0.72).

Conclusions In conclusion, MR TA is a viable tool to quantify medical images and illustrate correlations of microarchitectural changes invisible to a human reader with potential clinical impact.

Key Points

- *TA is feasible on paraspinal musculature using MRI.*
- *TA on paraspinal musculature correlates with SSM and RMDQ.*
- *TA may enable a statement regarding clinical impact of imaging findings.*

Keywords Magnetic resonance imaging · Machine learning · Spine · Muscles

Abbreviations

AUC	Area under the curve
BMI	Body-mass-index
CSA	Cross-sectional muscle area
DICOM	Digital imaging and communications in medicine
GLCM	Grey-level co-occurrence matrix
k-NN	k-Nearest Neighbour
LBP	Lumbar back pain
LSS	Lumbar spinal stenosis
MRI	Magnetic resonance imaging
NRS	Numeric rating scale (pain)
PACS	Picture archiving and communication system
RLM	Run-length matrix
RMDQ	Roland-Morris-disability questionnaire
ROC	Receiver-operating characteristics
ROI	Region of interest
SSM	Spinal stenosis measure
TA	Texture analysis
TSE	Turbo spin echo

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00330-018-5552-6) contains supplementary material, which is available to authorized users.

✉ Roman Guggenberger
roman.guggenberger@usz.ch

[1] Institute of Diagnostic and Interventional Radiology, University Hospital Zurich, University of Zurich, Raemistrasse 100, 8091 Zurich, Switzerland

[2] Horten Centre for Patient Oriented Research and Knowledge Transfer, University of Zurich, Pestalozzistrasse 24, 8032 Zurich, Switzerland

[3] Department of Orthopaedics, Balgrist University Hospital Zurich, University of Zurich, Forchstrasse 340, 8008 Zurich, Switzerland

Introduction

Lower back pain constitutes one of the main contributing factors to disability and absence from work in industrialised countries [1, 2]. A leading cause for lower back pain constitutes lumbar spinal stenosis (LSS). LSS is a debilitating condition caused by multifactorial narrowing of the spinal canal or vertebral foramina of the lumbar spine. Symptoms of LSS include, but are not limited to (radiating) lower back pain (LBP), muscle weakness, paraesthesia, and pseudoclaudication [3, 4].

Magnetic Resonance Imaging (MRI) of asymptomatic individuals commonly depicts degenerative intervertebral disc changes, nerve root compressions, spondylolysis, spondylolisthesis, and/or foraminal stenoses [5, 6]. However, these imaging findings neither correlate with clinical symptoms, nor do they consistently predict the risk of lower back pain [7, 8]. Beattie et al [9] showed that even in symptomatic patients, abnormal features on MRI studies demonstrate little correlation with pain patterns. Similarly, Burgstaller et al [10] showed negligible correlation between imaging findings and pain severity. The missing causal link is hypothesised to be due to the multifactorial nature of the reported pain.

Paraspinal muscle denervation and consecutive atrophy with fatty degeneration is commonly seen in patients with LSS [11]. This observation is likely to have clinical impact as paraspinal musculature is known to play a key function in the functional integrity and stability of the vertebral column. A recent meta-analysis for instance, has shown evidence that a decrease in multifidus muscular cross-sectional area is associated with lower back pain and possesses even predictive value for up to 12 months [12]. Furthermore, paraspinal muscle mass is known to decrease with age, while the proportion of fat increases [4, 13]. These changes potentially influence muscular microarchitecture invisible to subjective visual assessment in MRI [14].

In the context of 'Radiomics', medical images can now be converted into minable data by means of mathematical algorithms that exploit inter-pixel relationships [15, 16]. They consider radiologic images as vast three-dimensional data sets in which each voxel of the underlying volume represents distinct physical measurements of tissue-dependent characteristics [17]. By help of pixel distribution patterns and clinical correlations thereof, new pathways to diagnosis may be established that not only increase the reproducibility in medical image reading, but possibly change our concept of classifying pathologies [18]. This approach can be incorporated with so-called 'texture analysis' (TA), which may be used for detecting abnormalities in radiological images which cannot be identified in a subjective visual assessment alone [19], while having a histopathologic correlation [20]. TA has shown promising results in the field of neuroimaging [21, 22], oncology [23, 24], and in skeletal applications [25, 26].

The purpose of this study was to apply TA on paraspinal musculature in symptomatic LSS patients without prior surgical or epidural interventions and correlate these findings with established clinical outcome measures.

Materials and methods

Study participants

Ninety patients of the multi-centric Lumbar Stenosis Outcome Study (LSOS) were consecutively enrolled. Study approval was granted by the local institutional review board and ethical committee (KEK-ZH-NR: 2010-0395/0). All study participants provided written informed consent. Adult patients with the diagnosis of LSS based on clinical examination and MR imaging, who suffered from symptomatic claudication and had available magnetic resonance imaging of diagnostic imaging quality were included. Individuals with (inter-)vertebral infections, spinal column fractures, severe scoliosis (>15°), and/or clinically relevant peripheral artery occlusive disease were excluded.

As surgical interventions are known to risk ischemia and denervation of the paraspinal musculature [27], patients who had undergone prior surgical treatment were also excluded from this trial. Similarly, patients who had undergone spinal infiltration up to three months prior to the MRI examination were excluded from further analysis. Additional background information about LSOS has been published previously [10, 28, 29].

Clinical outcome measures

All included patients were clinically assessed by use of established clinical endpoints for LSS imaging: First, every patient completed the Spinal Stenosis Measure (SSM) questionnaire within three months after the time of the MRI examination. The SSM questionnaire is a LSS-specific clinical score and was used in both its subscales: SSM symptoms, and SSM function. Each item is rated on a Likert scale (range 1-5, and range 1-4, respectively, best-worst) [30, 31].

A secondary clinical outcome measure was the Roland Morris Disability Questionnaire (RMDQ) in its original version and dichotomised to the recommended cut-off of 14 points [32]. The RMDQ is a disability measure in the form of a questionnaire, which is graded on a 24-point scale. The obtained score positively correlates with the degree of disability. The RMDQ has been shown to be reproducible and sensitive to change over time in individuals with LSS [33].

A third clinical outcome measure was the Numeric Rating Scale (NRS) regarding pain perception. It ranges from absence of pain (0 points) to worst possible pain (10 points) [34].

MR imaging

All patients included in the LSOS received MR imaging in a supine position. Imaging was performed on several MRI scanners with field strengths at 1.5 T and 3 T with varying scanning parameters due to multiple participating institutions. However, all participating centres incorporated a standardised scan protocol, which included sagittal T1 weighted (w), sagittal T2w, and axial T2w turbo spin echo (TSE) sequences, and a dedicated quality control by a board-certified radiologist.

All imaging data from the participating clinical centres were collected at a centralised picture archiving and communication system (PACS, IMPAX 6; AGFA Healthcare).

Axial 2-D T2w TSE images with 3.5 mm slice thickness perpendicular to the intervertebral disc at level L3/4 were directly acquired with a standard matrix of 256 x 256. Single slice DICOM images in the axial plane at this level were exported for postprocessing and consecutive qualitative readout and TA. Postprocessing included rescaling images according to the pixel spacing values obtained from the DICOM header to an in-plane resolution of 0.3 mm using a customised MATLAB routine.

Texture analysis

TA was performed in all 90 MR DICOM images of individuals with LSS using open source software (MaZda, version 4.6, Institute of Electronics, Technical University of Lodz) [35]. Normalisation was performed between the mean and three standard deviations ("±3σ" method) in order to correct for small technical intra- and interscanner fluctuations and other bias as shown in [36, 37].

TA was performed using features originating from five different groups:

(*1*) First-level TA features incorporated *histogram analyses*, which visualise surface texture by depicting intensity concentrations across the ROIs. (*2*) *Grey-Level Co-Occurrence Matrix (GLCM)* features are second level TA features, that are computed from intensities of pairs of pixels, where the spatial relationship of the two pixels in a pair is defined. (3) *Run-Length*-Matrix (RLM) depicts coarseness of a texture in a specific direction. It counts pixel-runs with specified grey-scale level and length. (*4*) *Autoregressive Model Features (ARM)*: This model assumes pixel intensities as a weighted sum of neighbouring pixel values. The transform method for TA is based on discrete (*5*) *HAAR Wavelet Features*: Wavelet images are scaled in vertical and horizontal directions, resulting in multiple frequency channels. Data on TA frequency components can be obtained by energies within the channels [38, 39]. 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 by one reader with 5 years of experience in musculoskeletal radiology (*M.M*) on axial T2 w TSE DICOM spine MR images. ROI delineation was restricted to the erector spinae, and spinotransverse paraspinal muscle groups as part of the recently proposed classification by Crawford et al [40] (Fig. 1).

Table 1 Overview of all computed texture categories with corresponding features

Texture category	Texture feature
Histogram	Mean, variance, skewness, kurtosis
Grey-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, grey-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	teta1 to 4, sigma
Wavelet transform (calculated for seven subsampling factors (n = 1 to 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

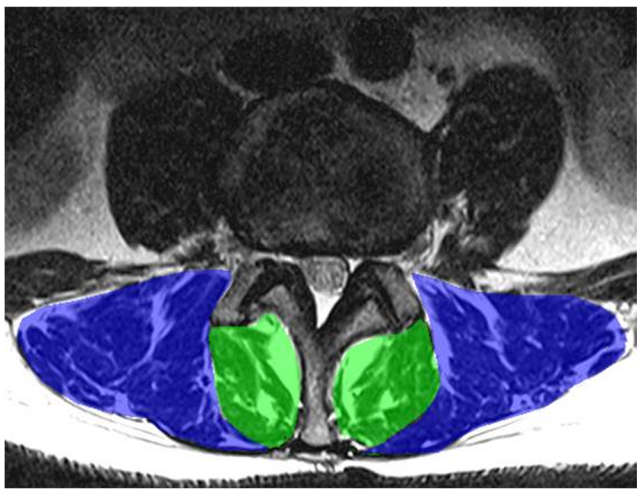

Fig. 1 Segmentation of paraspinal musculature in T2w axial MR images at level L3/4 for Texture Analysis. After prior rescaling to a standard resolution, freehand-region of interest (ROI) delineation, restricted to erector spinae group (longissimus and iliocostalis together) (blue), and spinotransverse muscles (predominantly multifidus) (green) was performed using dedicated software (MaZda, version 4.6, Institute of Electronics, Technical University of Lodz, Lodz, Poland)

Texture feature selection and feature reduction

The TA software calculated 304 TA features for each ROI (Table 1) as previously described by Baessler et al [41]. Feature selection and dimension reduction were performed by use of an 'AttributeSelectedClassifier', which evaluates a subset of attributes by considering the individual predictive ability of each feature along with the degree of redundancy among each other in a separate training data set.

Statistical analysis

Continuous variables were expressed as means ± standard deviation, and categorical variables as frequencies and/or percentages.

In order to overcome observed heteroscedasticity in clinical outcome measures (SSM, RMDQ, NRS) we performed a square root transformation before classification/ regression analysis.

To prevent overfitting, we split the dataset at random, in the recommended ratio of 2/3 (66.7%) training set (*model derivation cohort*) and 1/3 (33.3%) testing set (*validation cohort*) [42]. In detail, we used 10 fold cross validation for model derivation and applied the supervised classifiers to the remaining third of data in a separate data file containing previously unseen data.

The aforementioned 304 TA features were complemented with three additional clinical features for model derivation: (1) gender, (2) age, and (3) Body-Mass-Index (BMI). After dimension reduction, we used two common machine learning-based classifiers: Decision Table, and k-nearest neighbour (k-NN) [43, 44]. In detail, we used a Decision Table with majority classifier, and a k-NN classifier with (k= 1-10) and Euclidean distance function selected for implementation of a brute force search algorithm.

Best classification results were compared regarding their predictive ability of the aforementioned, square root transformed, clinical outcome measures.

Machine learning-based analyses were performed using non-commercial software (Weka, University of Waikato). Remaining analyses were conducted by using commercial software (SPSS v.21, IBM).

Results

Ninety patients of the multi-centric LSOS were prospectively enrolled for this retrospective, multicentre study. Patients' demographics organised by group are summarised in Table 2. The best classification/ regression results after exclusion of extreme outliers/ missing data are described below.

SSM symptoms

Machine learning-based regression of the SSM symptom score by use of Decision Table classifier resulted in coefficient of determination $r^2 = 0.02$ ($r = 0.13$) based on the validation cohort.

SSM function

Best regression results in SSM function were observed by use of k-NN with k = 1 and the Euclidean distance function selected for implementation of the brute force search algorithm. The obtained coefficient of determination was $r^2 = 0.15$ ($r = 0.39$).

Roland-Morris disability questionnaire

Dimension reduction returned 24 out of 307 features (7.8%). Best regression results of the continuous RMDQ were

Table 2 Descriptive statistics

	Model derivation cohort	Validation cohort
N	60	30
Gender (male: female) [%]	50: 50	63.3: 36.7
Age [years]	73.5 ± 9.7 (range: 47-88)	72.9 ± 7.2 (range: 58-87)
BMI [Kg/m^2]	26 ± 4.6 (range: 19-49)	28.6 ± 5.8 (range: 18-39)
NRS	6.6 ± 2 (range: 1-10)	4.7 ± 2.4 (range: 0-9)
RMDQ	12.2 ± 5.3 (range: 1-22)	8.9 ± 5.8 (range: 0-19)
SSM function	2.5 ± 0.7 (range: 1-3.8)	1.8 ± 0.7 (range: 1-3.4)
SSM symptoms	3.2 ± 0.6 (range: 1.6-4.7)	2.7 ± 4.1 (range: 1.4-2.8)

BMI = Body-Mass-Index; NRS = Numeric Rating Scale (Pain); RMDQ = Roland-Morris Disability Questionnaire

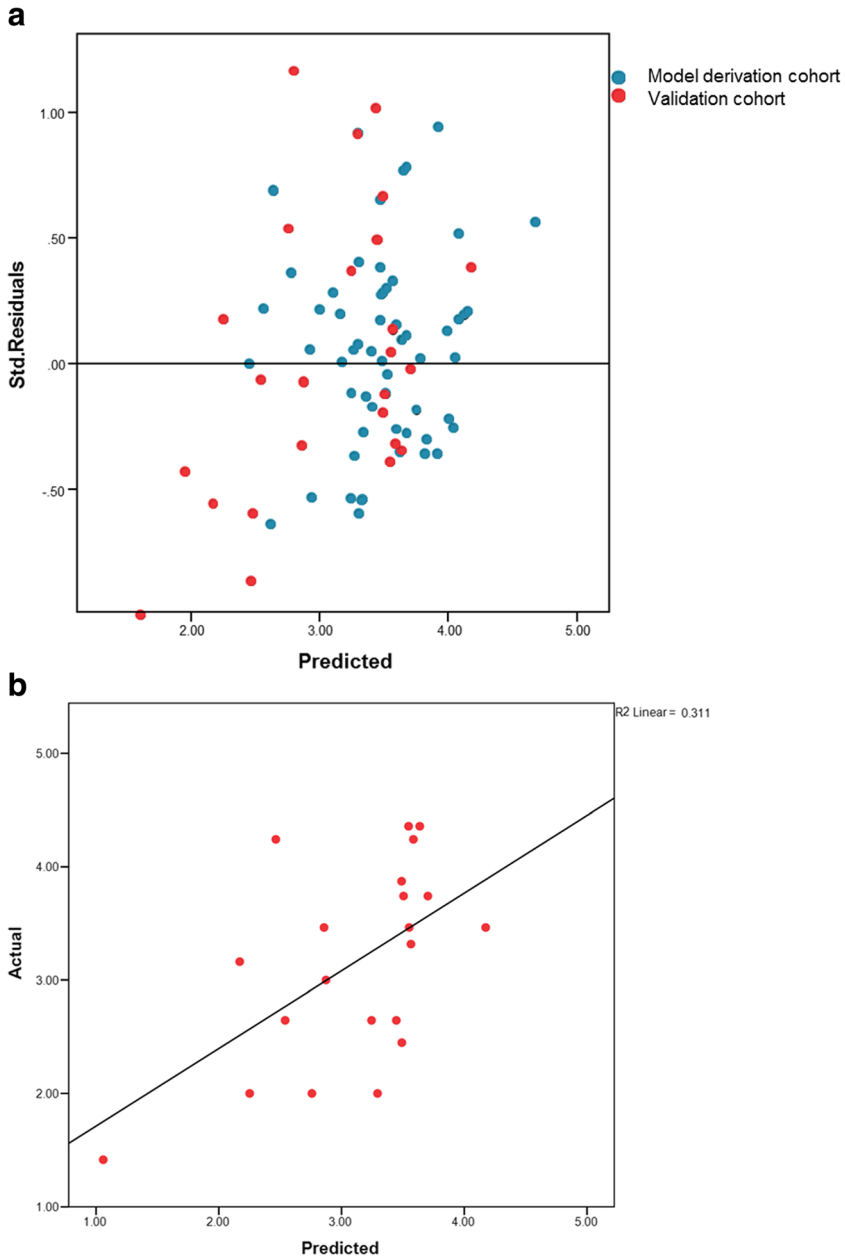

Fig. 2 **a** Residual plot: scatterplot depicting predicted values of variable Roland-Morris Disability Questionnaire (RMDQ) after square root transformation against its standardised Residuals. **b** Scatterplot of validation cohort illustrating the correlation of predicted square root transformed RMDQ values against their predicted values according to the results of the machine learning-based k-NN classifier

obtained by use of the k-NN classifier with k = 1 and the Euclidean distance function selected for implementation of the brute force search algorithm (Supplemental Table 1). The obtained coefficient of determination was $r^2 = 0.31$ ($r = 0.56$). A residual plot and a visualisation of the prediction results in the validation cohort are depicted in Fig. 2.

After dichotomisation of the continuous RMDQ at the recommended cut-off of 14 points, best classification results were observed by use of Decision Table. The detailed Decision table is depicted in Supplemental Table 2. Overall 70% of instances in the validation cohort could be correctly classified. The specificity was 61.1%. Receiver-Operating Characteristics (ROC) returned an area-under-the-curve of 0.7 (Fig. 3).

Numeric rating scale for pain perception

Best regression results in NRS were observed by use of the Decision Table. For use of this classifier, we transformed the continuous variable to a multicategorical one. The detailed Decision table is depicted in Supplemental Table 3. The obtained coefficient of determination was $r^2 = 0.51$ ($r = 0.72$). A residual plot and a visualisation of the prediction results in the validation cohort are depicted in Fig. 4.

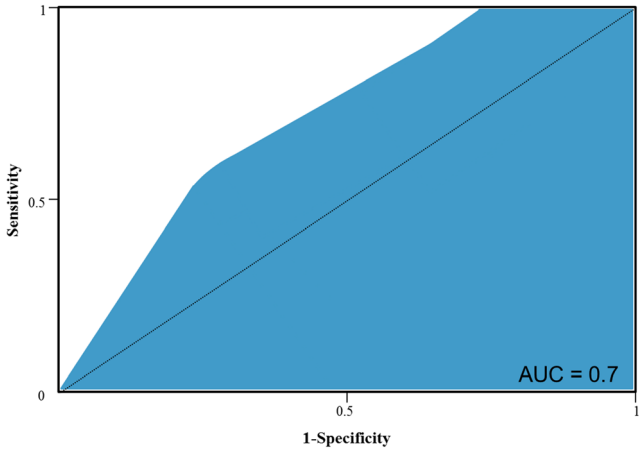

Fig. 3 Receiver-operating characteristics curve of dichotomised RMDQ showing an area-under-the-curve (AUC) of 0.7 at a cut-off of 14 points

Non-texture features

We used three non-texture features for model derivation: 'BMI', 'gender', and 'age'. While 'gender' and 'age' were not used in any outcome related classification, 'BMI' turned out to be a major contributing feature in both NRS and RMDQ prediction.

Summarised results of all classification/ regression analyses are depicted in Supplemental Table 4.

Discussion

Our study investigates the correlation of paraspinal muscle characteristics on clinical outcome measures using MR TA and machine learning-based algorithms. While we did not observe a meaningful correlation between TA and the two clinical outcome measures SSM symptoms and SSM function, a moderate correlation was observed regarding the two clinical outcome measures RMDQ and NRS.

The RMDQ, a LBP-specific self-report questionnaire regarding patients' functional activities, showed a positive correlation. Our generated model was based on the continuous score utilising histogram feature 'skewness', various GLCM features, one autoregressive model feature ('Teta4'), and the non-TA feature 'BMI'. Our previous investigations showed a correlation of histogram changes with fatty atrophy in paraspinal musculature [45]. Changes in GLCM represent structural changes of paraspinal musculature and were previously associated with a possible prognostic value of impaired walking distance in the context of LSS [45]. These structural muscle changes in patients suffering from LBP were previously reported by Goubert et al [46]. Similarly, Abbas et al found differences in cross-sectional muscle area (CSA) and in their density at the same vertebral level we used [47]. Additionally, Leinonen et al described severe denervation with accompanying abnormal activation of paraspinal musculature in affected populations [48]. The autoregressive model feature Teta4 represents a higher-order feature, which provides information about local interactions between neighbouring pixels with respect to their grey-level values and has in its subform 'Teta1' been previously reported to allow for the discrimination of scarred myocardial tissue [41].

After dichotomisation of the RMDQ results into mild and severe cases, dimension reduction returned similar features: one histogram feature ('MinNorm'), two GLCM features ('S10DifEntrp', 'S5,-5DifVarnc'), and 'BMI'. Similar to "skewness", the higher 'MinNorm' histogram values observed in severe cases correspond to higher histogram signal intensity values observed in fat. This corroborates also findings of prior studies on the association of fatty muscle infiltration with degenerative changes and lower back pain [13, 45, 49]. Obesity itself is known to be associated with a higher incidence of LSS [50] with lower functional status scores [51], but in isolation of little prognostic value [52, 53]. The importance of fatty infiltration is also presented in our Decision Table model by the non-TA feature 'BMI', which shows a positive correlation with NRS. This observed correlation of pain perception and fatty atrophy is in accordance with recent findings in a systematic review of Goubert et al [54].

The missing correlation with the commonly used rating scales SSM in its two sub-forms, shows that (1) these rating scales are not sufficient to characterise a specific relationship between paraspinal muscle TA and clinical symptoms, and highlight (2) the manifold pathophysiology of lower back pain.

Our observed clinical associations with TA findings are based on machine learning classifiers. These allow us to train a classifier and test the resulting algorithm and its test performance on previously unseen data. In comparison to conventional multivariate analysis, current machine learning algorithms are suitable to test multiple features despite limited sample size. To minimise so called 'overfitting', we split the dataset at random and implemented 10 fold cross validation for model derivation and applied the supervised classifiers to the separate validation data set, which contained previously unseen data.

The following limitations need to be acknowledged. First, this was a retrospective study with inherent limitations. Secondly, we excluded study participants with recent epidural infiltrations and prior surgeries in order to best pick up correlations between muscle architecture and clinical status at a certain time point without confounders. Therefore, our sample size is relatively low. Moreover, in

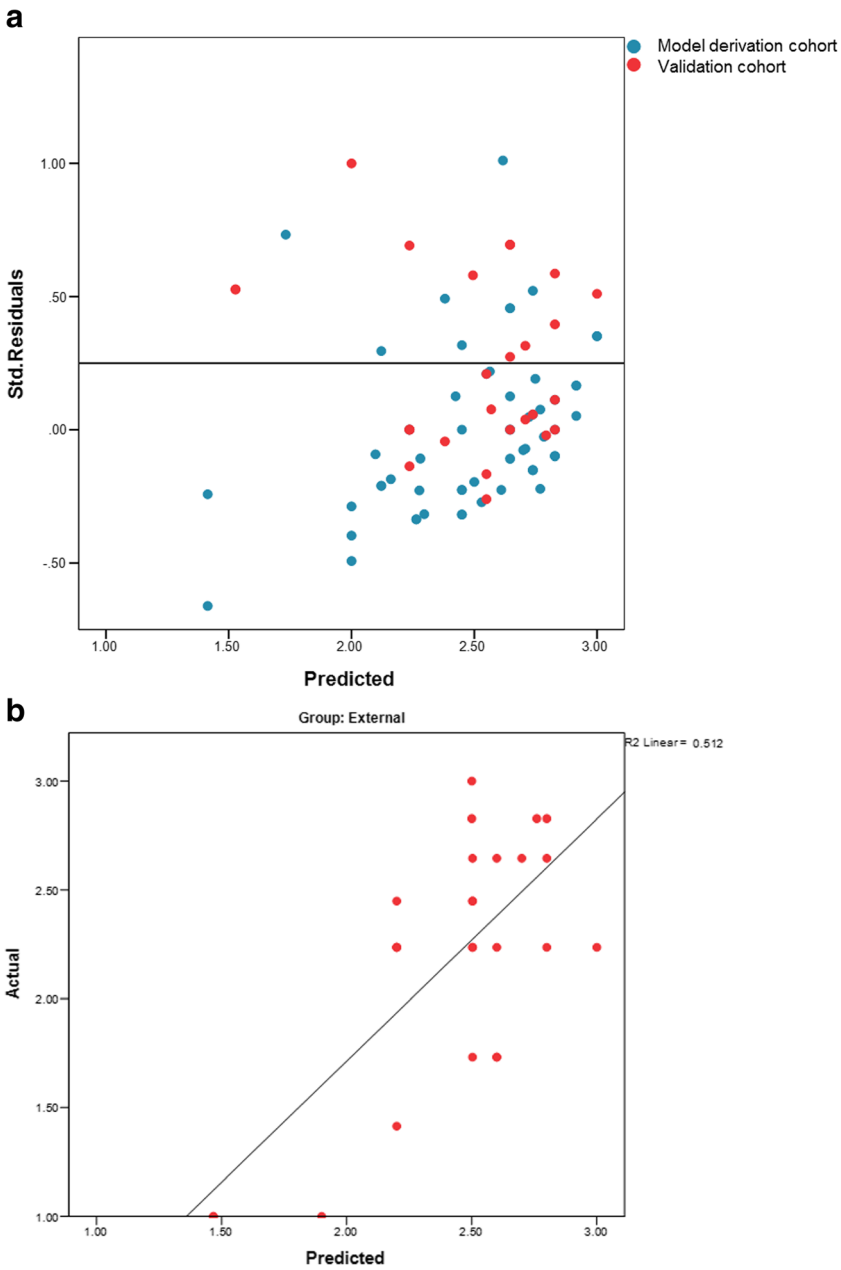

Fig. 4 **a** Residual plot: Scatterplot depicting predicted values of variable NRS after square root transformation against its standardised Residuals. **b** Scatterplot of validation cohort illustrating the correlation of predicted square root transformed NRS values against their predicted values according to the results of the machine learning-based Decision Table classifier

order to account for spinal curvature, TA was restricted to one spinal level (L3/4) and is likely limited in the presence of other T2 w hyperintense findings, e.g., muscular edema.

Moreover, we extracted commonly used TA features by use of an established software framework (MaZda); however, it does not allow the investigation of certain features, e.g., local binary patterns.

Finally, image acquisition was heterogeneous due to vendor-specific differences, despite a common scan protocol in this multicentre study. However, in order to account for imaging heterogeneity, data homogenisation prior to TA analyses was performed as described earlier.

In conclusion, MR TA is a viable tool to quantify medical images and illustrate changes invisible to a human reader. Its findings show correlation with clinical symptoms regarding pain and functional impairment.

Funding This study has received funding by Helmut Horten Foundation, Baugarten Foundation, Pfizer-Foundation for geriatrics & research in geriatrics, Symphasis Charitable Foundation and OPO Foundation funds.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Roman Guggenberger, MD.

Conflict of interest The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

Statistics and biometry Ulrike Held and Jakob M. Burgstaller kindly provided statistical advice for this manuscript.

Informed consent Written informed consent was obtained from all subjects (patients) in this study.

Ethical approval Institutional Review Board approval was obtained.

Study subjects or cohorts overlap http://www.lumbalstenose.ch/home/

Methodology
- retrospective
- observational
- multicentre study

References

1. Waddell G (1996) Low back pain: a twentieth century health care enigma. Spine (Phila Pa 1976) 21:2820–2825
2. Lurie J, Tomkins-Lane C (2016) Management of lumbar spinal stenosis. BMJ 352:h6234
3. Deyo RA, Mirza SK, Martin BI, Kreuter W, Goodman DC, Jarvik JG (2010) Trends, major medical complications, and charges associated with surgery for lumbar spinal stenosis in older adults. JAMA 303:1259–1265
4. Fortin M, Lazary A, Varga PP, Battie MC (2017) Association between paraspinal muscle morphology, clinical symptoms and functional status in patients with lumbar spinal stenosis. Eur Spine J. https://doi.org/10.1007/s00586-017-5228-y
5. Airaksinen O, Brox JI, Cedraschi C et al (2006) Chapter 4. European guidelines for the management of chronic nonspecific low back pain. Eur Spine J 15(Suppl 2):S192–S300
6. de Bruin F, ter Horst S, Bloem HL et al (2016) Prevalence of degenerative changes of the spine on magnetic resonance images and radiographs in patients aged 16-45 years with chronic back pain of short duration in the Spondyloarthritis Caught Early (SPACE) cohort. Rheumatology (Oxford) 55:56–65
7. Savage RA, Whitehouse GH, Roberts N (1997) The relationship between the magnetic resonance imaging appearance of the lumbar spine and low back pain, age and occupation in males. Eur Spine J 6:106–114
8. Brinjikji W, Luetmer PH, Comstock B et al (2015) Systematic literature review of imaging features of spinal degeneration in asymptomatic populations. AJNR Am J Neuroradiol 36:811–816
9. Beattie PF, Meyers SP, Stratford P, Millard RW, Hollenberg GM (2000) Associations between patient report of symptoms and anatomic impairment visible on lumbar magnetic resonance imaging. Spine (Phila Pa 1976) 25:819–828
10. Burgstaller JM, Schuffler PJ, Buhmann JM et al (2016) Is there an association between pain and magnetic resonance imaging parameters in patients with lumbar spinal stenosis? Spine (Phila Pa 1976) 41:E1053–E1062
11. Leinonen V, Maatta S, Taimela S et al (2002) Impaired lumbar movement perception in association with postural stability and motor- and somatosensory-evoked potentials in lumbar spinal stenosis. Spine (Phila Pa 1976) 27:975–983
12. Ranger TA, Cicuttini FM, Jensen TS et al (2017) Are the size and composition of the paraspinal muscles associated with low back pain? A systematic review. Spine J 17:1729–1748
13. Kalichman L, Hodges P, Li L, Guermazi A, Hunter DJ (2010) Changes in paraspinal muscles and their association with low back pain and spinal degeneration: CT study. Eur Spine J 19:1136–1144
14. Keller A, Gunderson R, Reikeras O, Brox JI (2003) Reliability of computed tomography measurements of paraspinal muscle cross-sectional area and density in patients with chronic low back pain. Spine (Phila Pa 1976) 28:1455–1460
15. Ingrisch M, Schneider MJ, Norenberg D et al (2017) Radiomic analysis reveals prognostic information in T1-weighted baseline magnetic resonance imaging in patients with glioblastoma. Investig Radiol. https://doi.org/10.1097/RLI.0000000000000349
16. Aerts HJ, Velazquez ER, Leijenaar RT et al (2014) Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 5:4006
17. Kolossvary M, Kellermayer M, Merkely B, Maurovich-Horvat P (2017) Cardiac computed tomography radiomics: a comprehensive review on radiomic techniques. J Thorac Imaging. https://doi.org/10.1097/RTI.0000000000000268
18. Barabasi AL, Gulbahce N, Loscalzo J (2011) Network medicine: a network-based approach to human disease. Nat Rev Genet 12:56–68
19. Tourassi GD (1999) Journey toward computer-aided diagnosis: role of image texture analysis. Radiology 213:317–320
20. De Certaines J, Larcher T, Duda D et al (2015) Application of texture analysis to muscle MRI: 1-What kind of information should be expected from texture analysis? EPJ Nonlinear Biomedical Physics 3
21. Li Y, Liu X, Xu K et al (2018) MRI features can predict EGFR expression in lower grade gliomas: a voxel-based radiomic analysis. Eur Radiol 28:356–362
22. Prasanna P, Patel J, Partovi S, Madabhushi A, Tiwari P (2016) Radiomic features from the peritumoral brain parenchyma on treatment-naive multi-parametric MR imaging predict long versus short-term survival in glioblastoma multiforme: preliminary findings. Eur Radiol. https://doi.org/10.1007/s00330-016-4637-3
23. Liu S, Liu S, Ji C et al (2017) Application of CT texture analysis in predicting histopathological characteristics of gastric cancers. Eur Radiol 27:4951–4959
24. Bates A, Miles K (2017) Prostate-specific membrane antigen PET/MRI validation of MR textural analysis for detection of transition zone prostate cancer. Eur Radiol 27:5290–5298
25. Gondim Teixeira PA, Leplat C, Chen B et al (2017) Contrast-enhanced 3T MR perfusion of musculoskeletal tumours: T1 value heterogeneity assessment and evaluation of the influence of T1 estimation methods on quantitative parameters. Eur Radiol 27:4903–4912
26. Lisson CS, Lisson CG, Flosdorf K et al (2017) Diagnostic value of MRI-based 3D texture analysis for tissue characterisation and discrimination of low-grade chondrosarcoma from enchondroma: a pilot study. Eur Radiol. https://doi.org/10.1007/s00330-017-5014-6
27. Bresnahan LE, Smith JS, Ogden AT et al (2016) Assessment of paraspinal muscle cross-sectional area following lumbar decompression: minimally invasive versus open approaches. Clin Spine Surg. https://doi.org/10.1097/BSD.0000000000000038
28. Ulrich NH, Burgstaller JM, Held U et al (2017) The influence of single-level versus multilevel decompression on the outcome in multisegmental lumbar spinal stenosis: analysis of the Lumbar Spinal Outcome Study (LSOS) data. Clin Spine Surg. https://doi.org/10.1097/BSD.0000000000000469
29. Aichmair A, Burgstaller JM, Schwenkglenks M et al (2017) 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 26:501–509
30. Stucki G, Liang MH, Fossel AH, Katz JN (1995) Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. J Clin Epidemiol 48: 1369–1378
31. Stucki G, Daltroy L, Liang MH, Lipson SJ, Fossel AH, Katz JN (1996) Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. Spine (Phila Pa 1976) 21: 796–803
32. Roland M, Morris R (1983) A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. Spine (Phila Pa 1976) 8:141–144
33. Stratford PW, Binkley J, Solomon P, Finch E, Gill C, Moreland J (1996) Defining the minimum level of detectable change for the Roland-Morris questionnaire. Phys Ther 76:359–365 discussion 366-358
34. Childs JD, Piva SR, Fritz JM (2005) Responsiveness of the numeric pain rating scale in patients with low back pain. Spine (Phila Pa 1976) 30:1331–1334
35. Szczypinski PM, Strzelecki M, Materka A, Klepaczko A (2009) MaZda–a software package for image texture analysis. Comput Methods Prog Biomed 94:66–76
36. Collewet G, Strzelecki M, Mariette F (2004) Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. Magn Reson Imaging 22:81–91
37. Orphanidou-Vlachou E, Vlachos N, Davies NP, Arvanitis TN, Grundy RG, Peet AC (2014) Texture analysis of T1 - and T2 - weighted MR images and use of probabilistic neural network to discriminate posterior fossa tumours in children. NMR Biomed 27:632–639
38. Tahir F, Fahiem MA (2014) A statistical-textural-features based approach for classification of solid drugs using surface microscopic images. Comput Math Methods Med 2014:791246
39. Szczypiński PM, Strzelecki M, Materka A, Klepaczko A (2009) MaZda - the software package for textural analysis of biomedical images. Springer, Berlin
40. Crawford RJ, Cornwall J, Abbott R, Elliott JM (2017) 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 18:25
41. Baessler B, Mannil M, Oebel S, Maintz D, Alkadhi H, Manka R (2017) Subacute and chronic left ventricular myocardial scar: accuracy of texture analysis on nonenhanced cine MR images. Radiology. https://doi.org/10.1148/radiol.2017170213:170213
42. Dobbin KK, Simon RM (2011) Optimally splitting cases for training and testing high dimensional classifiers. BMC Med Genet 4:31
43. Dwyer B, Hutchings K (1977) Flowchart optimisation in cope, a multi-choice decision table. Aust Comp J 9:92
44. Altman NS (1992) An introduction to kernel and nearest-neighbor nonparametric regression. Am Stat 46:175–185
45. Mannil M, Burgstaller JM, Thanabalasingam A et al (2018) Texture analysis of paraspinal musculature in MRI of the lumbar spine: analysis of the Lumbar Stenosis Outcome Study (LSOS) data. Skeletal Radiol. https://doi.org/10.1007/s00256-018-2919-3
46. Goubert D, De Pauw R, Meeus M et al (2017) Lumbar muscle structure and function in chronic versus recurrent low back pain: a cross-sectional study. Spine J 17:1285–1296
47. Abbas J, Slon V, May H, Peled N, Hershkovitz I, Hamoud K (2016) Paraspinal muscles density: a marker for degenerative lumbar spinal stenosis? BMC Musculoskelet Disord 17:422
48. Leinonen V, Maatta S, Taimela S et al (2003) Paraspinal muscle denervation, paradoxically good lumbar endurance, and an abnormal flexion-extension cycle in lumbar spinal stenosis. Spine (Phila Pa 1976) 28:324–331
49. Teichtahl AJ, Urquhart DM, Wang Y et al (2015) Physical inactivity is associated with narrower lumbar intervertebral discs, high fat content of paraspinal muscles and low back pain and disability. Arthritis Res Ther 17:114
50. Knutsson B, Sanden B, Sjoden G, Jarvholm B, Michaelsson K (2015) Body mass index and risk for clinical lumbar spinal stenosis: a cohort study. Spine (Phila Pa 1976) 40:1451–1456
51. Fanuele JC, Abdu WA, Hanscom B, Weinstein JN (2002) Association between obesity and functional status in patients with spine disease. Spine (Phila Pa 1976) 27:306–312
52. Rihn JA, Radcliff K, Hilibrand AS et al (2012) Does obesity affect outcomes of treatment for lumbar stenosis and degenerative spondylolisthesis? Analysis of the Spine Patient Outcomes Research Trial (SPORT). Spine (Phila Pa 1976) 37:1933–1946
53. Onyekwelu I, Glassman SD, Asher AL, Shaffrey CI, Mummaneni PV, Carreon LY (2017) Impact of obesity on complications and outcomes: a comparison of fusion and nonfusion lumbar spine surgery. J Neurosurg Spine 26:158–162
54. Goubert D, Oosterwijck JV, Meeus M, Danneels L (2016) Structural changes of lumbar muscles in non-specific low back pain: a systematic review. Pain Physician 19:E985–E1000