



Published in final edited form as:

*Spine (Phila Pa 1976)*. 2021 February 15; 46(4): 259–267. doi:10.1097/BRS.0000000000003833.

## A Novel Magnetic Resonance Imaging-based Lumbar Muscle Grade to Predict Health-related Quality of Life Scores Among Patients Requiring Surgery

Sohrab Virk, MD, MBA<sup>a</sup>, Joshua Wright-Chisem, MD<sup>a</sup>, Milan Sandhu, BS<sup>a</sup>, Avani Vaishnav, MD<sup>a</sup>, Todd J. Albert, MD<sup>a,b</sup>, Catherine Himo Gang, MPH<sup>a</sup>, Sheeraz Qureshi, MD, MBA<sup>a,b</sup>

<sup>a</sup>Hospital for Special Surgery, New York, NY

<sup>b</sup>Weill Cornell Medical College, New York, NY

### Abstract

**Study Design.**—Retrospective cross-sectional cohort.

**Objective.**—The aim of this study was to determine whether muscle health measurements are associated with health-related quality of life scores (HRQOLs) for patients with lumbar spine pathology.

**Summary of Background Data.**—Poor muscle health has been implicated as a source of pain/dysfunction for patients with lumbar spine pathology. Our aim was to quantify the relationship using muscle health measurements and HRQOLs.

**Methods.**—Three hundred and eight patients were included (mean age 57.7 ± standard deviation 18.2 years' old). We randomly selected patients into a derivation cohort (200) and validation cohort (108) to create our muscle health grade. We measured muscle health by the lumbar indentation value (LIV), goutallier classification (GC), and ratio of paralumbar muscle cross-sectional area over body mass index (PL-CSA/BMI). A muscle health grade was derived based on whether a measurement showed a statistically significant impact on visual analog scale back and leg pain (VAS-leg and VAS-leg), Oswestry Disability Index (ODI), short-form 12 physical health score (SF-12 PHS), short-form 12 mental health score (SF-12 MHS) and Patient-reported Outcomes Measurement Information System (PROMIS). A variety of statistical tools were used to determine whether there was a relationship between a measurement and HRQOLs.

**Results.**—In the derivation cohort, a muscle health grade was created based on the GC and PL-CSA/BMI ratio. For patients with a GC ≥ 2, one point was given. For patients with a PL-CSA/BMI ≥ 130, one point was given. Patients with 2 points were graded as “A” and 0 or 1 point were graded “B.” Within the validation cohort of patients, there was a statistically significant higher PROMIS (mean 34.5 ± standard deviation 12.6 vs. 27.6 ± 14.0,  $P=0.002$ ), ODI (38.8 ± 18.3 vs. 45.8 ± 18.1,

---

Address correspondence and reprint requests to Sohrab Virk, MD, MBA, Hospital for Special Surgery, New York, NY 10065; virks@hss.edu.

**Level of Evidence:** 3

Relevant financial activities outside the submitted work: consultancy, grants, royalties, stocks.

$P=0.05$ ) and SF-12 PHS ( $34.7 \pm 11.3$  vs.  $29.1 \pm 6.3$ ,  $P=0.002$ ) for patients with a good muscle health grade of “A.”

**Conclusion.**—This study offers an objective measurement of muscle health that correlates with HRQOLs for patients with lumbar spine pathology.

### Keywords

disability; health-related quality of life scores; low back pain; lumbar grading system; lumbar spine; muscle; muscle health; pain; paralumbar muscle health

---

The prevalence of low back and spinal pathology has been estimated between 65% and 85% with one-fourth of adults in the United States stating that they have had low back pain within the last three months.<sup>1-4</sup> Recent studies show that sarcopenia or decreased muscle strength/quality is associated with low back pain.<sup>5-8</sup> The manner with which this decrease in muscle health/quality impacts common spinal conditions causing low back pain, such as lumbar spinal stenosis or a disc herniation, remains unclear.

Although there have been studies examining how certain conditions impact patient function in relation to lumbar muscle health, there has been limited data examining how muscle health impacts commonly reported health-related quality of life scores.<sup>9-11</sup> Previous studies have identified changes in paralumbar muscles with age and low back pain.<sup>12,13</sup> Wagner et al also found that larger psoas cross-sectional area was protective against severe disability as measured by the Oswestry Disability Index (ODI).<sup>14</sup> The authors of this study suspect that a portion of disability and pain that patients with lumbar pathology suffer from is due to decreased muscle health.

To study this relationship, we have examined three preoperative markers for paralumbar muscle health and examined whether these markers impact health-related quality of life scores (HRQOLs). We performed this experiment in two steps. We first derived a muscle health score based upon an analysis of muscle health measurements on a cohort of patients. We then applied this score to another validation cohort of patients to ensure our score accurately predicted health related quality of life scores.

## METHODS

We performed a retrospective review of patient charts and imaging between January 2017 and April 2019. This was done with the approval of the senior author's home institutional review board (IRB). We included consecutive operative patients that presented with debilitating back and/or radiating leg pain to a spine clinic in the Northeast United States. Patients with <6 weeks of symptoms were not included within our study. Only patients who had received non-operative management of their symptoms with treatments including, but not limited to, physical therapy, acupuncture, chiropractor, activity modification, and so on, were included in our analysis. Patients were diagnosed and indicated for surgery with the following primary diagnoses: spondylolisthesis, disc herniation, foraminal stenosis, central stenosis, degenerative disc disease. Patients with debilitating back and/or radiating low back pain were indicated for surgery if their back pain/radiating leg pain matched anatomic

findings of stenosis and/or instability on radiographs and magnetic resonance imaging. We excluded patients with an acute traumatic injury causing their symptoms and those with a cancer diagnosis. Any patient with a previous lumbar spine surgery was also excluded.

Basic demographic information was collected on each patient including age, sex, and body mass index (BMI). At the patient's preoperative visit there were six scores collected on each patient. These HRQOLs included patient-reported outcome measurement information system physical health (PROMIS) scores, ODI, short-form 12 mental health score (SF-12 MHS), short-form 12 physical health score (SF-12 PHS), visual analog scale back (VAS back), and visual analog scale leg (VAS leg) scores. PROMIS physical function was the sole PROMIS measure used within this study.

### Muscle Health Measurements

Muscle health was measured in a variety of manners. We used a previously validated measurement of muscle degeneration named the lumbar indentation value (LIV).<sup>12</sup> An example measurement is shown in Figure 1. This measurement is the perpendicular distance between the spinous process and a line tangential to the paralumbar muscle bulges.

To obtain the cross-sectional area of paralumbar muscles (PL-CSA), we used a specially designed free online software.<sup>12,15</sup> The image-processing software platform (Image J, National Institutes of Health, Bethesda, MD) utilized allowed us to trace out the boundary of the combined multifidus and erector spinae muscles. A sample of this measurement is shown in Figure 2. We used a methodology of measuring that was similar to one performed by Takayama *et al.*<sup>12</sup> The muscle area was outlined manually with a cursor on the right and left side of the spinous process and was added together. The area of multifidus and erector spinae was outlined from the facet/lamina/spinous process to posteriorly and laterally to the thoracolumbar fascia. Figure 2 shows the combined area of erector spinae/multifidus measured on one axial slice. These measurements were then added together. To correct for a patient's habitus when evaluating the absolute value of PL-CSA, we divided PL-CSA (mm<sup>2</sup>) with the BMI of the patient (kg/m<sup>2</sup>). To date, no study has used this ratio to quantify paralumbar muscle health.

The Goutallier classification of the paralumbar muscle was graded on a 1 to 4 scale based on a qualitative assessment of fat atrophy of the muscle.<sup>16</sup> A One of 4 was defined as having minimal to no fatty streaks in muscle, a 2 of 4 as having fat evident but more muscle present, a 3 of 4 meant equal fat and muscle, and a 4 of 4 was more fat than muscle.

All measurements were done on an axial T2 weight cut from the MRI at L4-L5. We selected L4-L5 based upon previous research that showed lumbar muscle degeneration is most prevalent at this level when present at all.<sup>12,17</sup> All measurements were done using Sectra Medical Imaging (Sectra AB, Linkoping, Sweden). The majority of MRIs were done with a 1.5T-scanner. The repetition time and echo time were 3635 ms and 100 ms, respectively with 3 to 3.5 mm cuts. The matrix size was 416 × 224. Patients were placed in a supine position with the lumbar spine in a neutral position. A portion of MRI scans was done at outside institutions and we could not control for the specifications of these outside MRIs, although any MRIs that were not done in a supine position were automatically excluded.

All measurements were performed by either a fellowship-trained spine surgeon, resident in orthopedic surgery, or medical student. Measurements performed by the medical student and orthopaedic surgery resident were checked by a fellowship-trained spine surgeon to ensure proper accuracy. Medical students and residents were trained on the methodology of measuring paralumbar muscle (cross-sectional area, lumbar indentation value, and Goutallier classification) over the course of a day. Relevant articles detailing this measuring methodology were also provided to the medical student and orthopaedic surgery resident to help with understanding the accurate method to measure paralumbar muscle.<sup>12,18,19</sup> An intraclass correlation coefficient was calculated for 20 measurements between researchers for CSA and LIV. A kappa value was calculated for 20 measurements of Goutallier classification between researchers.

### Development and Validation of Muscle Health Grade

To create our muscle health grade we randomly divided our patient population into two cohorts. One cohort represented two-thirds of our entire patients and was used to develop our muscle health grade (derivation cohort). The second group was used to validate our findings (validation cohort). A schematic of how this score was created is shown in Figure 3.

### Statistical Analysis

To compare preoperative HRQOL scores to the evaluation of muscle health several statistical tools were utilized. The HRQOL values for each Goutallier class were compared using an analysis of variance (ANOVA) calculation. A linear regression analysis was used to compare LIV with the preoperative HRQOL. When multiple analyses were performed on the same set of data, a bonferoni correction factor was to correct the p values reported. We compared the ratio PL-CSA/BMI to HRQOL's using a linear regression analysis. We also tested the diagnostic performance of each measurement using a receiver-operating characteristic curve (ROC curve). A ROC curve was created based on having a PROMIS score of 40 as the state variable. This PROMIS score was selected as it represented one standard deviation from the average level of physical function for the general United States population.<sup>20,21</sup> We choose 40 as our cutoff because any amount of lumbar spine related pathology frequently disables patients significantly and the majority of patients with a diagnosis such as lumbar spinal stenosis has PROMIS physical function scores <40 (*i.e.*, more than one standard deviation below the general population).<sup>22</sup> If a parameter had an acceptable area under the curve (AUC) for its corresponding ROC curve then the optimal cutoff point to stratify "good" *versus* "poor" muscle was determined from the ROC curve. The Youden index (specificity + sensitivity - 1) was used to determine the optimal cutoff value for each variable as well.<sup>23</sup> A patient was then given one point for each muscle health measurement that corresponded to a "good" muscle health score.

We then compared HRQOLs from each scoring category within our muscle health score based on either a student *t* test or an ANOVA analysis in both the derivation and validation cohort. A statistically significant value was any correlation or difference with a p value of 0.05 or less.

## RESULTS

There were 308 patients included within our analysis. The basic demographic information, muscle health measurements, and the breakdown of the patient's pathology are shown in Table 1. A breakdown of HRQOL scores is also provided with Table 1. No statistically significant differences were found between the derivation and validation cohorts of patients in terms of any of the HRQOL scores or muscle health measurements. The intraclass correlation coefficient for CSA was 0.99 and for LIV was 0.98 showing excellent reliability between researchers. The kappa value for goutallier classification was 0.78 showing good reliability.

### Derivation Cohort Results—Individual Scores and HRQOL Values

The most common Goutallier classification for our patients was 2. There were 31 patients with a muscle health score of 1, 112 patients with a score of 2, 34 patients with a score of goutallier of 3, and 23 patients with a score of 4. The HRQOL score that correlated with Goutallier classification was PROMIS ( $P=0.01$ ). The AUC for the goutallier classification was fair (AUC=0.75). A cutoff of 2 was chosen to represent “good” muscle health as it represented a sensitivity of 0.95 and a high Youden index of 0.39. The ROC curve for goutallier classification is shown in Figure 4A.

The LIV was correlated with all six HRQOL scores using a linear regression analysis. No statistically significant linear relationships were found. As above, we created a ROC curve for the LIV and found an AUC of 0.54. Therefore, we did not include the LIV within our muscle health score as it failed to correlate with HRQOLs using a linear regression analysis and our ROC curve showed that LIV had no discrimination capacity for the PROMIS score.

The PL-CSA/BMI ratio was correlated with all six HRQOL scores using a linear regression analysis. We found that there was a statistically significant linear relationship between PL-CSA/BMI and PROMIS ( $P<0.001$ ,  $r=0.29$ ) and VAS-leg scores ( $P=0.02$ ,  $r=0.16$ ). As above, we created an ROC curve for PL-CSA/BMI. The AUC for PL-CSA/BMI was fair (AUC=0.73). Of note, a “fair” AUC has been deemed acceptable in previously published studies of predictive models.<sup>24–26</sup> A cutoff of 130 was chosen to represent “good” muscle health as it represented a sensitivity of 0.90 and had a high Youden index of 0.33. The ROC curve for PL-CSA is shown in Figure 4B.

### Derivation Cohort—Creating the Muscle Health Grade

To create our muscle health grade, we combined our results from the above analysis to grade paralumbar muscle based on goutallier classification and PL-CSA/BMI. A Goutallier classification of 2 was deemed to be “good” muscle health and a patient would be given one point. Similarly, if a patient had a PL-CSA/BMI of  $\geq 130$  they were graded as “good” and given a point. Table 2 outlines the combined components of the muscle health grade which were able to model HRQOLs. Within our scoring rubric a score of 0 or 1 corresponded to a “poor” muscle health grade of “B”. A score of 2 corresponded to “good” muscle health grade of “A.”

### Validation of Muscle Health Grade

We applied the same grading system defined above to a separate cohort of patients. There were 108 patients included within the validation cohort of patients. Within the validation cohort, patients with a PL-CSA/BMI >130 had a statistically significant better SF-12 PHS ( $P=0.02$ ), ODI ( $P=0.02$ ), and VAS-leg ( $P=0.02$ ). In this same validation cohort, patients with a GC of 2 had an improved SF-12 PHS ( $P=0.01$ ) and PROMIS ( $P=0.001$ ).

There were 32 patients with a muscle health score of 0, 27 patients with a score of 1, and 49 patients with a score 3. The breakdown of HRQOL scores for each grade of muscle health is shown in Table 3 for the validation cohort. There is a statistically significant difference in SF-12 PHS, ODI, and PROMIS in relation to a better muscle health grade for the derivation cohort.

When applying the muscle health grade to the entire cohort of patients (308) we found these findings remained consistent. There were statistically significant better VAS-leg ( $P=0.02$ ), SF-12 PHS ( $P=0.02$ ), ODI ( $P=0.02$ ), and PROMIS scores ( $P<0.001$ ) for patients with “good” muscle health.

## DISCUSSION

We found that there are significant correlations between poor muscle health and worse HRQOL scores for patients suffering from lumbar spine pathology. Poor muscle health as quantified by a low muscle health grade corresponds to worse functional status as measured by ODI, PROMIS, and SF-12 PHS. These findings offer insight into the role that paralumbar muscle health might play into overall disability of patients as measured by HRQOL measurements.

This muscle health grade provides a clinical tool and insight on how sarcopenia is associated with disability associated with common spinal disorders.<sup>5,27–30</sup> It is important for spinal surgeons to understand this link given the high prevalence of sarcopenia in patients, especially the elderly.<sup>31</sup> Estimates of sarcopenia range from 6% to 24% of the population and is markedly high (22.6%–26.8%) in the elderly.<sup>31–34</sup> There are a variety of ways to measure sarcopenia, but a commonly used modality is hand-grip strength testing.<sup>35</sup> This testing may not be done routinely in the office of a spinal surgeon, whereas surgeons are routinely evaluating lumbar MRIs. We hope that this study provides surgeons a quick surrogate measurement of sarcopenia and the related disability of sarcopenia by evaluating the Goutallier classification and estimating the overall paralumbar muscle area. Of note, the cross-sectional area of paralumbar muscle can be measured by first downloading once the imagej software from the National Institute of Health website for free. Saving an axial MRI at the L4-L5 disc space as a JPEG and using the imagej software to draw out the area of muscle of interest. This can be done in a relatively short amount of time. Furthermore, this study also provides a link between the findings related to sarcopenia/deconditioning and commonly reported HRQOLs.

Our findings of decreased muscle health related to poor functional scores ties into previously reported literature on patients with lumbar spinal stenosis having hyperactive paralumbar



muscles.<sup>10</sup> Leinonen *et al* found that patients with lumbar spinal stenosis had surprisingly good muscle endurance. This could be from increased activity of lumbar muscles in patients with lumbar spinal stenosis working to maintain lumbar lordosis.<sup>36</sup> Further study is required to understand how paralumbar muscle health changes throughout the disease process. How hyperactive paralumbar muscles impact muscle health measurements on MRI is also a potential area of further research. The authors acknowledge that lumbar spine pathology and disability is a complex relationship and muscle health is not the sole determinant of disability. The authors contend, however, that muscle health may be a subtle but important factor contributing to discomfort/disability for patients with lumbar spine pathology as measured by HRQOLs. Furthermore, the differences in HRQOLs between the cohort of patients with “good” *versus* “poor” muscle health approaches or is above minimally clinically important difference (MCID) figures for SF-12 physical health score, PROMIS physical function, and ODI.<sup>37–39</sup>

The relationship between disc degeneration and muscle health has been explored and offers an important area of future research. Several studies have outlined the correlation between disc degeneration and paralumbar muscle degeneration.<sup>13,40,41</sup> The pathophysiology of this process of disc degeneration and muscle degeneration may to some degree be related to changes in biologic signaling between these structures.<sup>42,43</sup> The authors feel this is a vital area of new research that may lead to future treatments for patients suffering from lumbar spine pathology.

An important finding of this study is the use of a scaling factor in order to properly compare paralumbar cross sectional area across patients. The PL-CSA/BMI figure was found to correlate with pre-operative HRQOLs. We believe this is a unique measurement of muscle health as it factors in the patient’s habitus in order to properly gauge the relative size of paralumbar muscle cross-sectional area. Previous studies to date have not incorporated a scaling factor to cross-sectional area.<sup>44–47</sup>

There are several important limitations to the present study. It is a retrospective study of patients with known lumbar pathology and therefore we do not have a control group of normal patients to compare our various muscle health measurements. We attempted to control for this variable by separating our patient population into derivation and validation cohorts. This allowed us to develop a score then apply its results to a separate patient population. Future study will require us to apply this scoring system to a normative control population of patients with lumbar MRI. There has also been literature demonstrating the link between chronic pain and changes in paralumbar muscle.<sup>6,48</sup> We did not control for this potential source for sarcopenia, although it may in part be reflected in the ODI within our 2 cohorts of patients. We also did not factor in other HRQOL measures that may better reflect the quality of a patient’s pain, time course of dysfunction/disability, sites of pain, among others. The precise nature of low back pain and/or radiating leg pain may significantly impact muscle health and this is likely a topic for future research. The authors acknowledge that the dichotomous nature of our scoring system does limit the usefulness of our muscle health grade for clinical use. Body mass index was used to correct for the generalized fat percentage that a patient might be living with. There is, however, evidence that BMI does not reflect perfectly the location of excess mass.<sup>49</sup> It was included within

our study as it is a commonly measured patient characteristic and we hoped that its wide use would facilitate interpretation and utility of our muscle health grade. Further study on the distribution of fat in the abdomen and above the fascia in the lumbar spine and its impact on commonly used HRQOLs for the spine will likely be needed. There is an inherent limitation related to the LIV measurement in that a patient's position on the MRI scanner may impact the LIV. Pressure from the MRI scanner may distort the surface topology of paralumbar muscle and impact the LIV measurement. The authors also acknowledge that this methodology of creating a scoring system from validation/derivations may be unfamiliar to some readers. There have been numerous publications in a variety of journals that have used this methodology to create a score or prediction model for a clinical entity.<sup>50–52</sup> Our analysis also includes a population of patients that is similar in size to previously published studies using this methodology.

In conclusion, this is the first study to show a link between a portion of HRQOL measurements and muscle health for patients with lumbar spine pathology requiring surgery. It also provides a relatively easy to obtain muscle health score that correlates strongly with functional HRQOLs. We hope these data inform clinicians on the role of paralumbar muscle health in disability related to lumbar spine pathology.

## References

1. Deyo RA, Mirza SK, Martin BIJS. Back pain prevalence and visit rates: estimates from US national surveys, 2002. *Spine (Phila Pa 1976)* 2002;2006;31:2724–7.
2. Shahidi B, Parra CL, Berry DB, et al. Contribution of lumbar spine pathology and age to paraspinous muscle size and fatty infiltration. *Spine (Phila Pa 1976)* 2017;42:616–23. [PubMed: 27517512]
3. Andersson GBJT. Epidemiological features of chronic low-back pain. *Lancet* 1999;354:581–5. [PubMed: 10470716]
4. Hartvigsen J, Hancock MJ, Kongsted A, et al. What low back pain is and why we need to pay attention. *Lancet (London, England)* 2018;391:2356–67. [PubMed: 29573870]
5. Tanishima S, Hagino H, Matsumoto H, et al. Association between sarcopenia and low back pain in local residents prospective cohort study from the GAINA study. *BMC Musculoskelet Disord*; 18:452– 2017. [PubMed: 29141602]
6. Wan Q, Lin C, Li X, et al. MRI assessment of paraspinous muscles in patients with acute and chronic unilateral low back pain. *Br J Radiol* 2015;88:; 20140546-. [PubMed: 26105517]
7. Goubert D, Oosterwijck JV, Meeus M, et al. Structural changes of lumbar muscles in non-specific low back pain: a systematic review. *Pain Phys* 2016;19:E985–1000.
8. Teichtahl AJ, Urquhart DM, Wang Y, et al. Fat infiltration of paraspinous muscles is associated with low back pain, disability, and structural abnormalities in community-based adults. *Spine J* 2015;15:1593–601. [PubMed: 25828477]
9. Chen Y-Y, Pao J-L, Liaw C-K, et al. Image changes of paraspinous muscles and clinical correlations in patients with unilateral lumbar spinal stenosis. *Eur Spine J* 2014;23:999–1006. [PubMed: 24395004]
10. Leinonen V, Maatta S, Taimela S, et al. Paraspinous muscle denervation, paradoxically good lumbar endurance, and an abnormal flexion-extension cycle in lumbar spinal stenosis. *Spine (Phila Pa 1976)* 2003;28:324–31. [PubMed: 12590204]
11. Chen YY, Pao JL, Liaw CK, et al. Image changes of paraspinous muscles and clinical correlations in patients with unilateral lumbar spinal stenosis. *Eur Spine J* 2014;23:999–1006. [PubMed: 24395004]
12. Takayama K, Kita T, Nakamura H, et al. New predictive index for lumbar paraspinous muscle degeneration associated with aging. *Spine (Phila Pa 1976)* 2016;41:E84–90. [PubMed: 26335668]

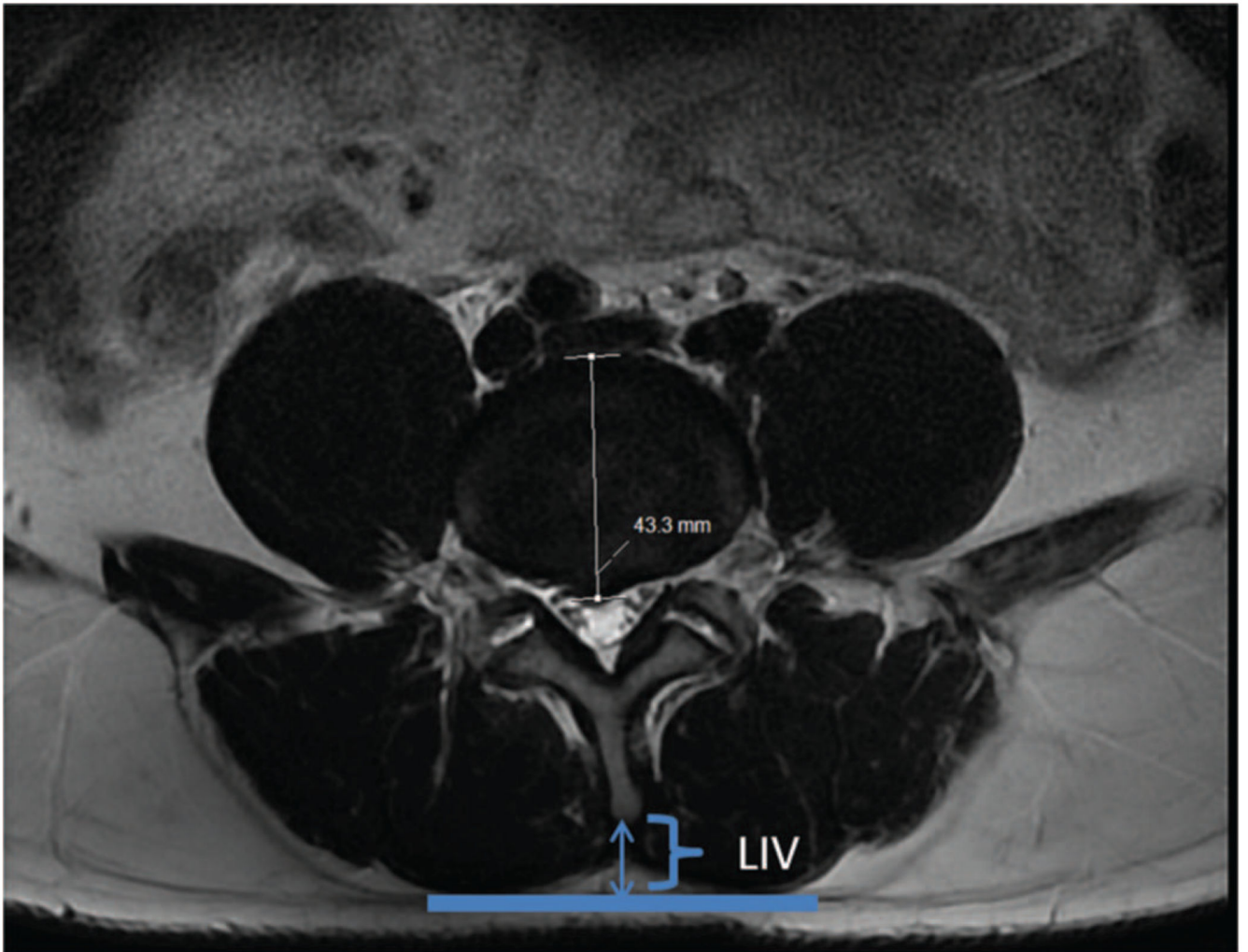


13. Kalichman L, Carmeli E, Been E. The association between imaging parameters of the paraspinal muscles, spinal degeneration, and low back pain. *BioMed Res Int* 2017;2017:2562957. [PubMed: 28409152]
14. Wagner SC, Sebastian AS, McKenzie JC, et al. Severe lumbar disability is associated with decreased psoas cross-sectional area in degenerative spondylolisthesis. *Global Spine J* 2018;8: 716–21. [PubMed: 30443482]
15. Fortin M, Battié MC. Quantitative paraspinal muscle measurements: inter-software reliability and agreement using OsiriX and Image J. *Phys Ther* 2012;92:853–64. [PubMed: 22403091]
16. 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;27:1005–12. [PubMed: 29396765]
17. Lee JC, Cha JG, Kim Y, et al. Quantitative analysis of back muscle degeneration in the patients with the degenerative lumbar flat back using a digital image analysis: comparison with the normal controls. *Spine (Phila Pa 1976)* 2008;33:318–25. [PubMed: 18303466]
18. Bang WS, Lee DH, Kim KT, et al. Relationships between vitamin D and paraspinal muscle: human data and experimental rat model analysis. *Spine J* 2018;18:1053–61. [PubMed: 29355791]
19. Battaglia PJ, Maeda Y, Welk A, et al. Reliability of the Goutallier classification in quantifying muscle fatty degeneration in the lumbar multifidus using magnetic resonance imaging. *J Manipulative Physiol Ther* 2014;37:190–7. [PubMed: 24630770]
20. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473–83. [PubMed: 1593914]
21. Oude Voshaar MAH, ten Klooster PM, Glas CAW, et al. Validity and measurement precision of the PROMIS physical function item bank and a content validity-driven 20-item short form in rheumatoid arthritis compared with traditional measures. *Rheumatology* 2015;54:2221–9. [PubMed: 26224306]
22. Patel AA, Dodwad SM, Boody BS, et al. Validation of Patient Reported Outcomes Measurement Information System (PROMIS) Computer Adaptive Tests (CATs) in the surgical treatment of lumbar spinal stenosis. *Spine (Phila Pa 1976)* 2018;43: 1521–8. [PubMed: 29557925]
23. Ruopp MD, Perkins NJ, Whitcomb BW, et al. Youden Index and optimal cut-point estimated from observations affected by a lower limit of detection. *Biom J* 2008;50:419–30. [PubMed: 18435502]
24. Wingert NC, Gotoff J, Parrilla E, et al. The ACS NSQIP risk calculator is a fair predictor of acute periprosthetic joint infection. *Clin Orthop Relat Res* 2016;474:1643–8. [PubMed: 26911971]
25. Cissen M, Wely MV, Scholten I, et al. Measuring sperm DNA fragmentation and clinical outcomes of medically assisted reproduction: a systematic review and meta-analysis. *PLoS One* 2016;11:e0165125. [PubMed: 27832085]
26. Sgroi M, Loitsch T, Reichel H, et al. Diagnostic value of clinical tests for supraspinatus tendon tears. *Arthroscopy* 2018;34:2326–33. [PubMed: 29802066]
27. Morley JE, Anker SD, von Haehling S. Prevalence, incidence, and clinical impact of sarcopenia: facts, numbers, and epidemiology— update 2014. *J Cachexia Sarcopenia Muscle* 2014;5:253–9. [PubMed: 25425503]
28. Park S, Kim HJ, Ko BG, et al. The prevalence and impact of sarcopenia on degenerative lumbar spinal stenosis. *Bone Joint J* 2016;98-b:1093–8. [PubMed: 27482023]
29. Park S-M, Kim G-U, Kim H-J, et al. Low handgrip strength is closely associated with chronic low back pain among women aged 50 years or older: a cross-sectional study using a national health survey. *PLoS One* 2018;13:e0207759. [PubMed: 30475880]
30. Rizzoli R, Reginster J-Y, Arnal J-F, et al. Quality of life in sarcopenia and frailty. *Calcif Tissue Int* 2013;93:101–20. [PubMed: 23828275]
31. Kenny AM, Prestwood KM, Iannuzzi-Sucich M. Prevalence of sarcopenia and predictors of skeletal muscle mass in healthy, older men and women. *J Gerontol A Biol Sci Med Sci* 2002;57:M772–7. [PubMed: 12456735]
32. Iii LJM, Khosla S, Crowson CS, et al. Epidemiology of sarcopenia. *J Am Geriatr Soc* 2000;48:625–30. [PubMed: 10855597]
33. Baumgartner RN, Koehler KM, Gallagher D, et al. Epidemiology of sarcopenia among the elderly in New Mexico. *Am J Epidemiol* 1998;147:755–63. [PubMed: 9554417]

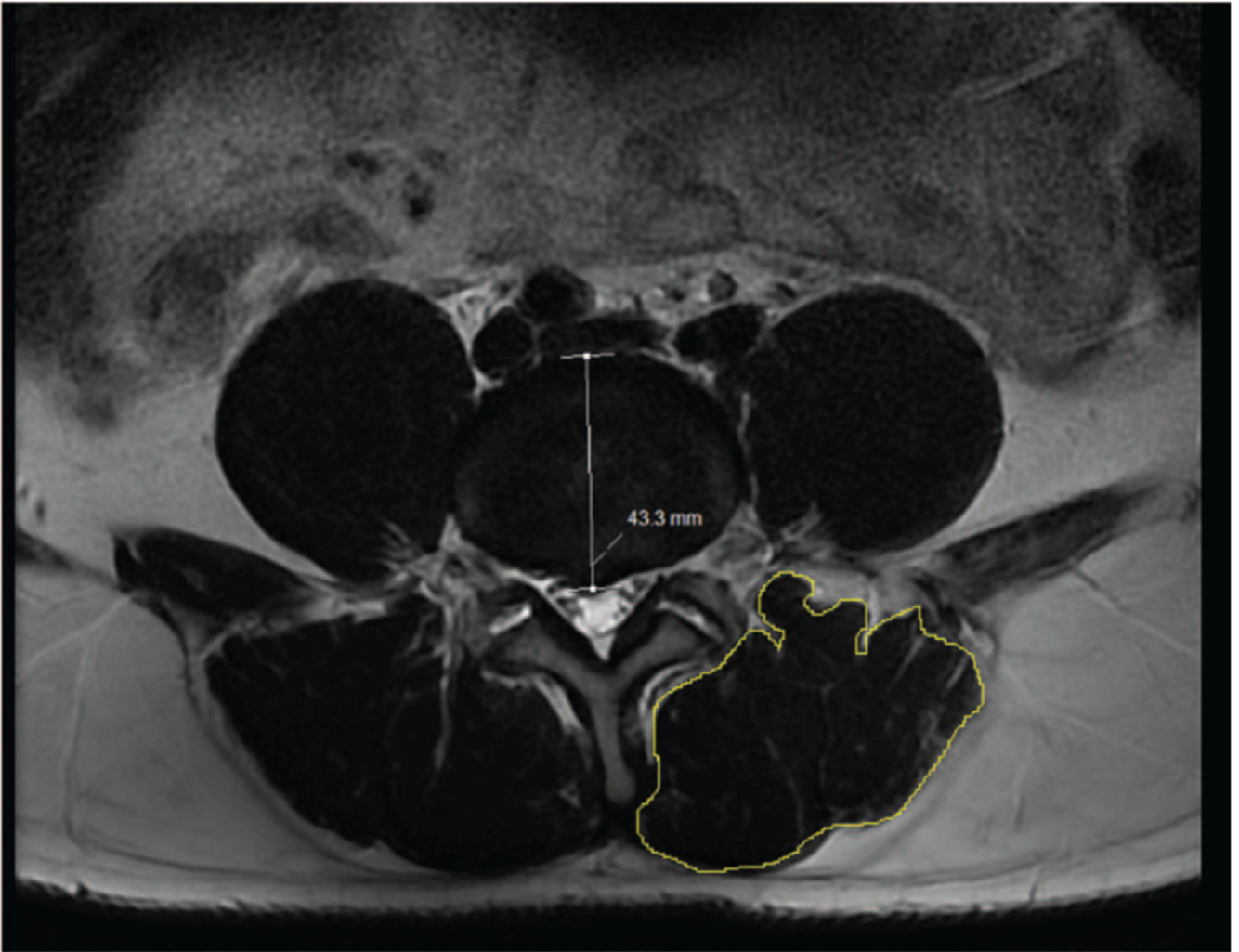
34. Tank LB, Movsesyan L, Mouritzen U, et al. Appendicular lean tissue mass and the prevalence of sarcopenia among healthy women. *Metabolism* 2002;51:69–74. [PubMed: 11782875]
35. Sousa-Santos AR, Amaral TF. Differences in handgrip strength protocols to identify sarcopenia and frailty—a systematic review. *BMC Geriatr* 2017;17:238. [PubMed: 29037155]
36. Kim H-J, Chung S, Kim S, et al. Influences of trunk muscles on lumbar lordosis and sacral angle. *Eur Spine J* 2006;15:409–14. [PubMed: 16151709]
37. Diaz-Arribas MJ, Fernandez-Serrano M, Royuela A, et al. Minimal clinically important difference in quality of life for patients with low back pain. *Spine (Phila Pa 1976)* 2017;42:1908–16. [PubMed: 28658040]
38. Hung M, Saltzman CL, Kendall R, et al. What Are the MCIDs for PROMIS, NDI, and ODI instruments among patients with spinal conditions?. *Clin Orthop Relat Res* 2018;476:2027–36. [PubMed: 30179950]
39. Copay AG, Glassman SD, Subach BR, et al. Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study questionnaire Short Form 36, and pain scales. *Spine J* 2008;8:968–74. [PubMed: 18201937]
40. Parkkola R, Korman M. Lumbar disc and back muscle degeneration on MRI: correlation to age and body mass. *J Spinal Disord* 1992;5:86–92. [PubMed: 1571617]
41. Kalichman L, Hodges P, Li L, et al. Changes in paraspinal muscles and their association with low back pain and spinal degeneration: CT study. *Eur Spine J* 2010;19:1136–44. [PubMed: 20033739]
42. Shahidi B, Fisch KM, Gibbons MC, et al. Increased fibrogenic gene expression in multifidus muscles of patients with chronic versus acute lumbar spine pathology. *Spine (Phila Pa 1976)* 2020;45:E189–95. [PubMed: 31513095]
43. Hodges PW, James G, Blomster L, et al. Multifidus muscle changes after back injury are characterized by structural remodeling of muscle, adipose and connective tissue, but not muscle atrophy: molecular and morphological evidence. *Spine (Phila Pa 1976)* 2015;40:1057–71. [PubMed: 25943090]
44. Jorgensen MJ, Marras WS, Gupta P. Cross-sectional area of the lumbar back muscles as a function of torso flexion. *Clin Biomech (Bristol, Avon)* 2003;18:280–6. [PubMed: 12689777]
45. Lee HI, Song J, Lee HS, et al. Association between cross-sectional areas of lumbar muscles on magnetic resonance imaging and chronicity of low back pain. *Ann Rehab Med* 2011;35:852–9.
46. Ghiasi MS, Arjmand N, Shirazi-Adl A, et al. Cross-sectional area of human trunk paraspinal muscles before and after posterior lumbar surgery using magnetic resonance imaging. *Eur Spine J* 2016;25:774–82. [PubMed: 25967560]
47. Choi MK, Kim SB, Park CK, et al. Cross-sectional area of the lumbar spine trunk muscle and posterior lumbar interbody fusion rate: a retrospective study. *Clin Spine Surg* 2017;30:E798–803. [PubMed: 27623301]
48. Crossman K, Mahon M, Watson PJ, et al. Chronic low back pain-associated paraspinal muscle dysfunction is not the result of a constitutionally determined “adverse” fiber-type composition. *Spine (Phila Pa 1976)* 2004;29:628–34. [PubMed: 15014272]
49. Nuttall FQ. Body mass index: obesity, bmi, and health: a critical review. *Nutr Today* 2015;50:117–28. [PubMed: 27340299]
50. Yilgor C, Sogunmez N, Boissiere L, et al. Global Alignment and Proportion (GAP) score: development and validation of a new method of analyzing spinopelvic alignment to predict mechanical complications after adult spinal deformity surgery. *J Bone Joint Surg Am* 2017;99:1661–72. [PubMed: 28976431]
51. James MT, Pannu N, Hemmelgarn BR, et al. Derivation and external validation of prediction models for advanced chronic kidney disease following acute kidney injury. *JAMA* 2017;318:1787–97. [PubMed: 29136443]
52. Carabini LM, Zeeni C, Moreland NC, et al. Development and validation of a generalizable model for predicting major transfusion during spine fusion surgery. *J Neurosurg Anesthesiol* 2014;26:205–15. [PubMed: 24714381]

**Key Points**

- Paralumbar muscle health correlates with HRQOL measures for patients with lumbar spine pathology
- Patients with a low ratio of paralumbar cross-sectional area over BMI have lower HRQOLs.
- A novel grading system using paralumbar cross-sectional area over BMI and Goutallier classification correlates strongly to HRQOLs.
- This study demonstrates the role of paralumbar muscle in the complex physiology of pain/disability from lumbar spine pathology.

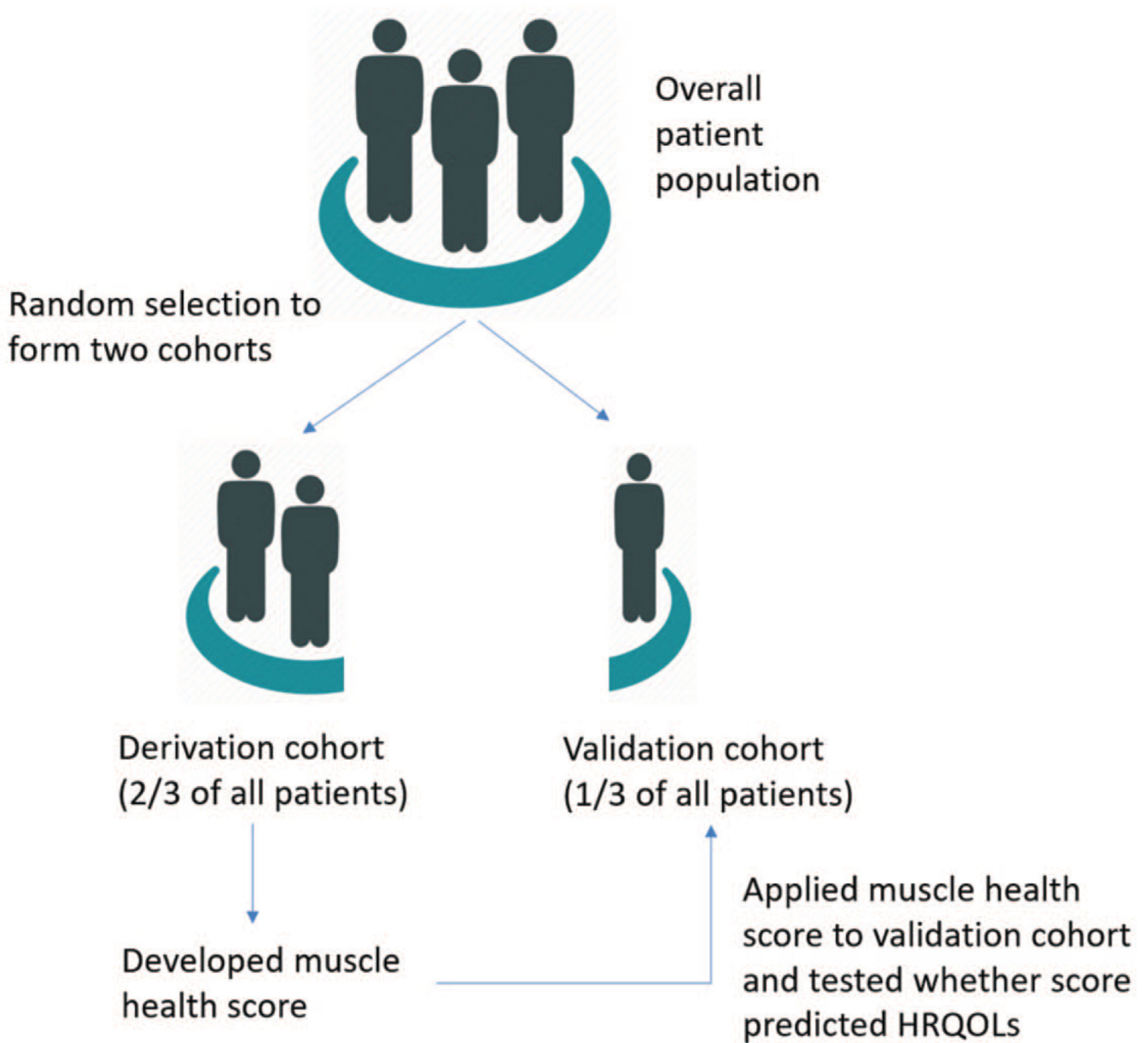


**Figure 1.**  
The lumbar indentation value (LIV) is the perpendicular distance between the spinous process and a line tangential to bilateral paralumbar muscle bulges.



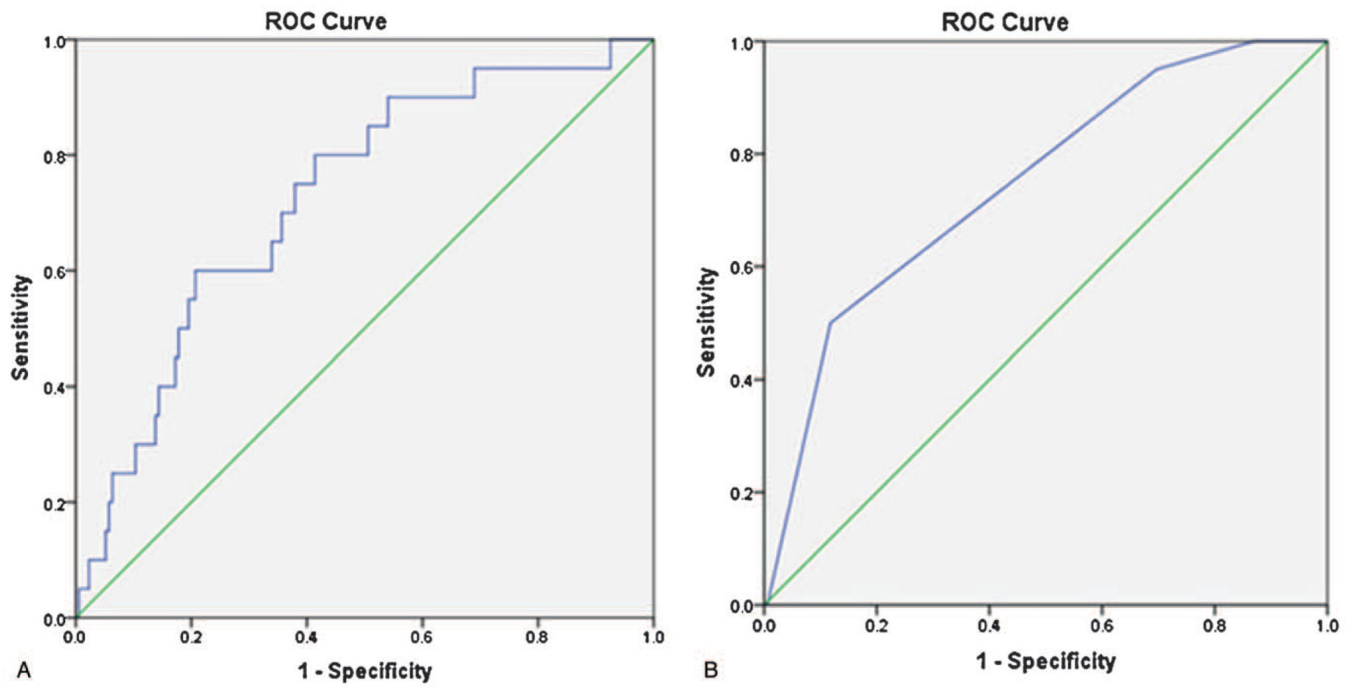
**Figure 2.** This is a sample measurement of the cross-sectional measurement of paraspinal muscle (PL-CSA) using T2-weighted imaging. This is an axial slice from the L4-L5 disc space. The yellow outline delineates the area of measurement which is drawn by cursor on both the right and left sides of the spine.





**Figure 3.** Method for deriving and validating the muscle health score.





**Figure 4.** These graphs represent the receiver-operating character curves for the goutallier classification (**A**) and paralumbar muscle (PL-CSA) (**B**).

Baseline Demographic and Health-related Quality of Life Scores for Both the derivation and validation cohort of Patients

TABLE 1.

Variable	Derivation Cohort	Validation Cohort	P
N	200	108	
Age, y, mean $\pm$ standard deviation	56.6 $\pm$ 18.8	59.9 $\pm$ 17.3	0.15
Female (no. of patients)	89 (46%)	51 (47%)	0.65
Body mass index, kg/m <sup>2</sup> , mean $\pm$ standard deviation	26.9 $\pm$ 5.5	27.1 $\pm$ 4.7	0.73
Diagnosis			
Spondylolisthesis (no. of patients)	55 (28%)	27 (25%)	0.64
Disc herniation (no. of patients)	65 (33%)	32 (30%)	0.60
Central stenosis (no. of patients)	96 (43%)	51 (47%)	0.90
Foraminal stenosis (no. of patients)	62 (31%)	41 (37%)	0.22
Degenerative disc disease (no. of patients)	78 (39%)	44 (41%)	0.77
PreOp HRQOL (mean $\pm$ standard deviation)			
VAS leg	4.9 $\pm$ 3.3	5.5 $\pm$ 3.3	0.16
VAS back	5.1 $\pm$ 3.9	5.2 $\pm$ 3.2	0.96
SF-12 MHS	47.1 $\pm$ 13.5	50.0 $\pm$ 41.9	0.39
SF-12 PHS	32.3 $\pm$ 9.3	31.4 $\pm$ 9.6	0.44
ODI	41.6 $\pm$ 19.3	40.2 $\pm$ 19.4	0.58
PROMIS	30.0 $\pm$ 14.7	29.1 $\pm$ 15.3	0.63
Muscle measurements			
Total cross-sectional area of paraspinal muscle, mm <sup>2</sup> , mean $\pm$ standard deviation	3852.8 $\pm$ 1073.4	3851.6 $\pm$ 1248.3	0.99
Lumbar indentation value, mm, mean $\pm$ standard deviation	14.1 $\pm$ 6.0	14.2 $\pm$ 6.3	0.88
Goutallier classification			
1 (no. of patients)	31 (15%)	23 (21%)	0.20
2 (no. of patients)	112 (56%)	48 (44%)	0.06
3 (no. of patients)	34 (16%)	26 (24%)	0.13
4 (no. of patients)	23 (12%)	11 (10%)	0.72

ODI indicates Oswestry Disability Index; PROMIS, Patient-reported Outcomes Measurement Information System; SF-12 PHS, short-form 12 physical health score; SF-12 MHS, short-form 12 mental health score; VAS, Visual Analog Scale.

**TABLE 2.**

## Components of the Muscle Health Score

Variable	Score
Goutallier Classification	
3-4	0
1-2	1
Paralumbar cross-sectional area/BMI	
<130	0
>130	1
Muscle Grade	Total score
Poor muscle health	0-1
Good muscle health	2

BMI indicates body mass index.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

The Breakdown of Health-related Quality of Life Scores For Each Grade Is Provided. In general, Scores Improved With Each Improved Muscle Grade

**TABLE 3.**

Muscle Health Grade	N	VAS Back	VAS Leg	SF-12 MHS	SF-12 PHS	ODI	PROMIS
A (score=0, 1)	59	5.9 ± 3.1	6.1 ± 3.2	52.6 ± 45.3	29.1 ± 6.3	45.8 ± 18.1	27.6 ± 14.0
B (score=2)	49	5.4 ± 3.2	4.9 ± 3.5	47.0 ± 37.7	34.7 ± 11.3	38.8 ± 18.3	34.5 ± 12.6
<i>P</i>		0.35	0.06	0.50	0.002	0.05	0.009

All continuous data points are presented as a mean value plus/minus a standard deviation. ODI indicates Oswestry Disability Index; PROMIS, Patient-reported Outcomes Measurement Information System; SF-12 PHS, short-form 12 physical health score; SF-12 MHS, short-form 12 mental health score; VAS, Visual Analog Scale.