An Objective and Reliable Method for Identifying Sarcopenia in Lumbar Spine Surgery Patients: Using Morphometric Measurements on Computed Tomography Imaging

Marko Tomov[1], Mohammed Ali Alvi[2], Mohamed Elminawy[2], Bradford Currier[1], Michael Yaszemski[1], Ahmad Nassr[1], Paul Huddleston[1], Arjun Sebastian[1], Mohamad Bydon[2], Brett Freedman[1]

[1]Department of Orthopedic Surgery, Mayo Clinic, Rochester, MN, USA
[2]Department of Neurosurgery, Mayo Clinic, Rochester, MN, USA

Study Design: A retrospective observational study.

Purpose: Establish a quantifiable and reproducible measure of sarcopenia in patients undergoing lumbar spine surgery based on morphometric measurements from readily available preoperative computed tomography (CT) imaging.

Overview of Literature: Sarcopenia—the loss of skeletal muscle mass—has been linked with poor outcomes in several surgical disciplines; however, a reliable and quantifiable measure of sarcopenia for future assessment of outcomes in spinal surgery patients has not been established.

Methods: A cohort of 90 lumbar spine fusion patients were compared with 295 young, healthy patients obtained from a trauma da¬tabase. Cross-sectional vertebral body (VB) area, as well as the areas of the psoas and paravertebral muscles at mid-point of pedicles at L3 and L4 for both cohorts, was measured using axial CT imaging. Total muscle area-to-VB area ratio was calculated along with intraclass correlation coefficients for interobserver and intraobserver reliability. Finally, T-scores were calculated to help identify those patients with considerably diminished muscle-to-VB area ratios.

Results: Both muscle mass and VB areas were considerably larger in males compared with those in females, and the ratio of these two measures was not enough to account for large differences. Thus, a gender-based comparison was made between spine patients and healthy control patients to establish T-scores that would help identify those patients with sarcopenia. The ratio for paravertebral muscle area-to-VB area at the L4 level was the only measure with good interobserver reliability, whereas the other three of the four ratios were moderate. All measurements had excellent correlations for intraobserver reliability.

Conclusions: We postulate that a patient with a T-score <–1 for total paravertebral muscle area-to-VB area ratio at the L4 level is the most reliable method of all our measurements that can be used to diagnose a patient undergoing lumbar spine surgery with sarcopenia.

Keywords: Sarcopenia; Lumbar; Spine; Surgery; Computed tomography

Introduction

Sarcopenia is qualitatively defined as a progressive impairment of muscle function secondary to a loss of skeletal muscle mass [1]. However, unlike osteoporosis, there is no consensus on the objective quantification of sarcopenia. The accurate identification of patients with sarcopenia is important as it is a known risk factor for mobility impairments, falls, disabilities, loss of independence, hospitalization, and death [1-6]. It has been estimated that a mere 10.5% decrease in the prevalence of sarcopenia could reduce healthcare costs by 1.1 billion dollars per year in the United States alone [7].

There is a paucity of data on the association between sarcopenia and surgical complications following major lumbar spinal procedures; however, frailty has been linked with postoperative mortality and complications in many surgical specialties [8-11]. Given the lack of a consensus on a quantitative definition of sarcopenia, we sought to first develop a reliable and intuitive method of defining sarcopenia in the lumbar spine patient population using commonly obtained imaging studies to accurately assess surgical risks and complications in future clinical research. We sought to compare objective and reproducible cross-sectional areal measures of the psoas and paraspinal muscles in the lumbar spine between spine patients and young, healthy controls. As with densiometric scoring in osteoporosis, we established a T-score equivalency in which the T-score was determined by comparing sarcopenic morphometric measures between lumbar spine patients (>50 years old) and young, healthy controls. This study aimed to compare previously reported and new morphometric measurements of sarcopenia based on axial computed tomography (CT) images between a cohort of lumbar spine surgery patients and a cohort of young (20- to 30-year-old), healthy patients. This study aimed to establish normative threshold values for these measures that would allow for the preoperative determination of diminished muscle mass in lumbar spine surgery patients over the age of 50 years old. Given the introduction of new measures and the novel normative scoring methodology, a secondary purpose of our study was to test the interobserver and intraobserver reliability of our methods.

Materials and Methods

1. Participants and measurement

The Institutional Review Board approval of Mayo Clinic was obtained for this study, which included reviewer approved waiver of the requirement to obtain informed consent as a low risk retrospective review (IRB approval no., 17-000648-01). We first developed two primary cohorts—spine patients and young, healthy controls—fo comparative analysis. An institutionally maintained database of lumbar spine surgery patients over the age of 50 years old was examined to find 90 consecutive lumbar spine surgery patients (average age, 67.9 years old; 60% female, 40% male) with available lumbar CT scans. To establish a young, healthy control cohort, our institution's trauma registry was used to identify patients (N=295) with nonspine trauma (American Spinal Injury Association Impairment Scale=1 for any region and Injury Severity Score <10 for no spine fractures or injuries) between the ages of 20 and 30 years who had undergone lumbar spine CT scan as part of the initial trauma assessment. Patients with diagnosed spine injuries or anomalies were excluded. Cross-sectional muscular areas were measured on routinely obtained preoperative CT imaging (i.e., CT L-spine) in spine patients and on trauma CT scans for the young, healthy controls. Axial CT cuts through mid-point of both L3 and L4 pedicles were selected, and the cross-sectional area of the following anatomic structures was measured: vertebral body (VB, excluding obvious osteophytes), left
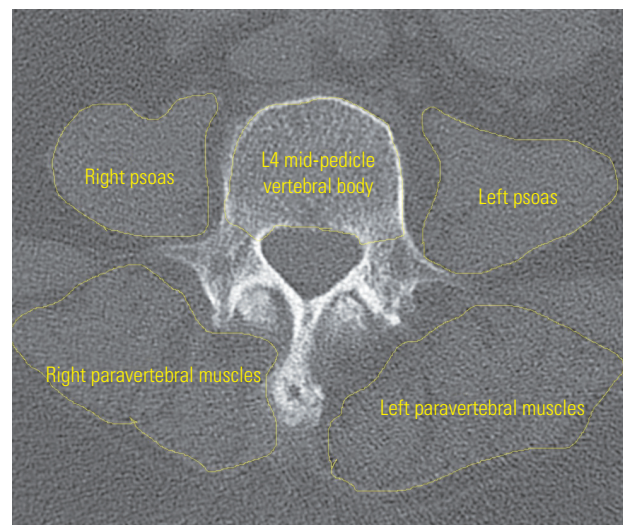

Fig. 1. Morphometric measurements on axial computed tomography.

and right psoas muscles, and left and right paravertebral muscles. Using the L4 level as an example, Fig. 1 demonstrates these area measurements. All measurements were made using our institution's proprietary picture archiving and communication system radiology software known as QReads (Mayo Clinic, Rochester, MN, USA). After compiling all morphometric measurements, both the left and right sides of a given muscle group—psoas or paravertebral muscles—were summed; additionally, the ratio of total muscle group surface area-to-VB area was calculated. To normalize measurements between patients of different sizes and genders, the ratio of total muscle group surface area-to-VB area was calculated. The reliability data were used to determine which level (L3 or L4) and which muscle group (psoas muscle or paravertebral muscles) would be most reproducible. Three independent researchers performed the morphometric measurements during the first pass to determine interobserver reliability; the same process was repeated for a subset of randomly selected patients (n=29) at least 2 weeks later to determine intraobserver reliability data. Intraclass correlation coefficients (ICC) were calculated to assess inter-rater reliability of the various morphometric measurements; ICC was rated poor for values <0.5, moderate for values between 0.5 and 0.75, good for values between 0.75 and 0.9, and excellent for values >0.9 [12]. Statistical analysis was conducted with R Foundation for Statistical Computing (Vienna, Austria).

1. Analysis of morphometric measures

To compute a method for a standardized report on morphometric sarcopenia measures, we calculated T-scores for various measurements as follows. We first calculated the mean and standard deviation of each measurement and ratio for young, healthy controls. For each patient in our cohort of interest, a T-score was defined as the distance between standard deviations of each sarcopenia measurement in spine patients and the mean scores of young, healthy controls. Last, these methods were repeated with spine patients and young, healthy control cohorts divided on the basis of gender to establish gender-based thresholds and evaluate whether reducing absolute areal measures to ratios mitigated the impact of genders on this method. Statistical significance was determined with chi-square or Fisher's exact test, depending on the number of occurrences.

Results

Using our database of 140 surgical spine patients, we analyzed 90 spine patients: 36 male patients and 54 female patients; six were excluded for age (<50 years) at the time of surgery; another 44 patients were excluded because the axial CT scans did not include the entire paravertebral muscle area. The young, healthy cohort with which comparisons were made to calculate T-scores was made up of 295 young, healthy patients from the trauma registry. Prior to calculating T-scores, interobserver and intraobserver reliability was calculated to determine which morphometric measurements were the most reproducible. The ICCs were calculated for both psoas and paravertebral muscles to VB ratios at L3 and L4. The ICC calculation results, along with 95% confidence intervals, are illustrated in Tables 1 and 2 for interobserver and intraobserver reliability, respectively. Three of the four sarcopenia scores were rated moderate correlations (ICC, 0.5–0.75), whereas the L4 paravertebral muscle-to-VB ratio was rated good (ICC >0.75) for interobserver reliability. All measurements had excellent reliability for intraobserver reliability (ICC >0.9). The confidence intervals would suggest that ICC is moderate-to-good for interobserver reliability and good-to-excellent for intraobserver reliability for three of the four measurements. Tables 3 and 4 show the comparison results of morphometric data between males and

Table 1. Interobserver reliability analysis

Variable	Intraclass correlation coefficient (95% confidence interval)
L3 psoas to VB ratio	0.66 (0.48–0.81)
L3 paravertebral to VB ratio	0.72 (0.55–0.85)
L4 psoas to VB ratio	0.71 (0.54–0.84)
L4 paravertebral to VB ratio	0.77 (0.62–0.87)

VB, vertebral body.

Table 2. Intraobserver reliability analysis

Variable	Intraclass correlation coefficient (95% confidence interval)
L3 psoas to VB ratio	0.98 (0.93–0.99)
L3 paravertebral to VB ratio	0.94 (0.83–0.98)
L4 psoas to VB ratio	0.91 (0.78–0.98)
L4 paravertebral to VB ratio	0.92 (0.78–0.98)

VB, vertebral body.

Table 3. Male to female comparison of morphometric data at L3

Variable	Female (N=109)		Male (N=158)		p-value
	Mean±SD	Range	Mean±SD	Range	
Right psoas muscle (mm²)	795±231	374–1,521	1,378±1,009	427–2,813	<0.001
Left psoas muscle (mm²)	792±228	343–1,484	1,325±349	181–2,745	<0.001
Total psoas muscle (mm²)	1,587±435	743–2,625	2,703±1,149	1,064–4,818	<0.001
VB area (mm²)	989±165	655–1,528	1,257±189	791–2,076	<0.001
Total psoas muscle to VB ratio	1.62±0.409	0.83–2.99	2.18±0.884	0.751–11.02	<0.001
Right paravertebral muscle (mm²)	2,026±383	1,129–3,124	2,926±534	1,302–4,698	<0.001
Left paravertebral muscle (mm²)	2,057±434	1,144–3,499	2,926±532	1,363–4,298	<0.001
Total paravertebral muscle (mm²)	4,064±817	1,652–6,623	5,852±1,013	2,837–8,780	<0.001
Total paravertebral muscle to VB ratio	4.12±0.87	1.77–6.45	4.68±0.97	1.78–7.83	<0.001

SD, standard deviation; VB, vertebral body.

Table 4. Male to female comparison of morphometric data at L4

Variable	Females (N=109)		Males (N=158)		p-value
	Mean±SD	Range	Mean±SD	Range	
Right psoas muscle (mm²)	1,047±249	460–1,745	1,673±341	528–2,507	<0.001
Left psoas muscle (mm²)	1,056±262	393–1,844	1,678±333	744–3,043	<0.001
Total psoas muscle (mm²)	2,099±489	853–358	3,344±661	1,210–5,401	<0.001
VB area (mm²)	1,037±154	760–1,577	1,309±191	906–1,954	<0.001
Total psoas muscle to VB ratio	2.03±0.446	0.78–3.29	2.59±0.550	0.80–4.43	<0.001
Right paravertebral muscle (mm²)	2,174±456	1,054–3,498	2,974±470	1,077–4,131	<0.001
Left paravertebral muscle (mm²)	2,174±408	1,197–3,311	2,991±457	1,504–4,090	<0.001
Total paravertebral muscle (mm²)	4,349±851	2,454–6,810	5,982±914	3,008–8,181	<0.001
Total paravertebral muscle to VB ratio	4.220±0.73	2.64–6.22	4.638±0.80	2.43–7.21	<0.001

SD, standard deviation; VB, vertebral body.

females for our control group. This comparison aimed to assess whether normalizing muscle area as a ratio to VB area would account for differences between genders. Both muscle mass and VB areas were remarkably larger in males compared with those in females, and what was evident is that the ratio of muscle mass to VB area was not enough to account for the differences between genders. Thus, this prompted a further division of the data and gender-specific T-score calculations between spine patients and control cohorts (Tables 5 and 6 for males and females, respectively).

Discussion

The ratio measurement of level and muscle to the VB area with the best interobserver reliability was determined to

be the paravertebral muscle group at L4. All measurements had excellent intraobserver reliability, and thus, this factor did not affect the choice for the best measure of sarcopenia. However, one important consideration noted by all three researchers was that, in some instances, the lumbar spine CT scans did not always include the entire paravertebral muscle area. Thus, a measurement often extended to the borders of the axial cut and likely underrepresented the true total area of paravertebral muscles. This was never an issue with psoas muscle measurements, mainly because the entire cross-sectional area was always within the captured area of the CT scans. Hence, we know that we prefer measuring the total paravertebral muscle area based on higher ICC. However, if the total area cannot be captured on the CT scan in a specific case, a measure of total psoas muscle area would be very reason-

Table 5. Male case and control comparison of morphometric data at L3 and L4

Variable	Males (N=36)	p-value
Total psoas to VB ratio at L3		<0.001
Case	1.63±0.47 (0.94–3.49)	
Control	2.18±0.88 (0.75–11.02)	
T-score groups		
<-2	0	
-2 to -1	7	
-1 to 0	25	
0 to 1	3	
>1	1	
Total paravertebral muscle to VB ratio at L3		<0.001
Case	3.67±0.96 (1.92–6.69)	
Control	4.68±0.97 (1.78–7.83)	
T-score groups		
<-2	6	
-2 to -1	13	
-1 to 0	12	
0 to 1	4	
>1	1	
Total psoas to VB ratio at L4		<0.001
Case	1.93±0.48 (1.08–3.57)	
Control	2.5±0.55 (0.79–4.42)	
T-score groups		
<-2	5	
-2 to -1	14	
-1 to 0	13	
0 to 1	3	
>1	1	
Total paravertebral muscle to VB ratio at L4		<0.001
Case	3.24±0.89 (1.34–5.91)	
Control	4.63±0.85 (2.44–7.2)	
T-score groups		
<-2	12	
-2 to -1	17	
-1 to 0	4	
0 to 1	2	
>1	1	

Values are presented as mean±standard deviation (range) or number.
VB, vertebral body.

Table 6. Female case and control comparison of morphometric data at L3 and L4

Variable	Females (n=54)	p-value
Total psoas to VB ratio at L3		<0.001
Case	1.30±0.30 (0.87–1.97)	
Control	1.62±0.41 (0.83–2.99)	
T-score groups		
<-2	0	
-2 to -1	22	
-1 to 0	22	
0 to 1	10	
>1	0	
Total paravertebral muscle to VB ratio at L3		<0.001
Case	3.33±0.90 (1.02–4.93)	
Control	4.12±0.87 (1.77–6.45)	
T-score groups		
<-2	7	
-2 to -1	21	
-1 to 0	14	
0 to 1	12	
>1	0	
Total psoas to VB ratio at L4		<0.001
Case	1.58±0.41 (0.84–2.87)	
Control	2.02±0.44 (0.77–3.3)	
T-score groups		
<-2	7	
-2 to -1	22	
-1 to 0	15	
0 to 1	9	
>1	1	
Total paravertebral muscle to VB ratio at L4		<0.001
Case	3.05±0.88 (1.15–4.84)	
Control	4.22±0.7 (2.6–6.2)	
T-score groups		
<-2	21	
-2 to -1	17	
-1 to 0	9	
0 to 1	7	
>1	0	

Values are presented as mean±standard deviation (range) or number.
VB, vertebral body.

able and reliable. Although the interobserver reliability was not excellent for any given measure, it was still considered very appropriate for use, given the moderate-to-good spectrum of ICC seen in the analysis of confidence interval. We speculate that if radiology software can be used to automate this process in the future, the reliability would be even higher with the elimination of human er-

ror in these measurements. Given the positive correlation between osteoporosis and sarcopenia from previous studies, we similarly modeled osteopenia/osteoporosis on our definition of sarcopenia [9,13-15]. We postulate that a patient with a T-score <−1 for total paravertebral muscle (or total psoas muscle in those scans where the total area might not be included)-to-VB area ratio measured at L4 can be diagnosed with sarcopenia.

To date, suggested screening methods for sarcopenia have ranged from anthropometric measurements, grip strength, and questionnaires [16]. Steihaug et al. [16] admit that these methods are less precise than dual-energy X-ray absorptiometry (DXA) measurements, despite DXA being an expensive and time-consuming study. Since preoperative advanced imaging is obtained prior to any operative spine procedure, we believe our method can draw on readily available information to quickly screen patients with sarcopenia and develop a treatment plan to maximize their success postoperatively. A plethora of evidence suggests that adequate physical and nutritional interventions to prehabilitate sarcopenic patients prior to surgery can improve outcomes and mortality [16,17]. We recognize that prehabilitation may not always be possible for spine trauma or emergent spine procedures, although identifying the patients who are sarcopenic with our method may help inform their postoperative course with more vigorous exercises and nutritional therapies. In a randomized controlled trial of elderly patients, frailty scores, physical activities, gait speed, and energy all increased considerably as early as 3 months after the onset of nutritional supplementation and physical training [18].

Extensive research has documented the opportunistic use of readily available CT imaging for applications such as the screening of patients for osteoporosis prior to obtaining a formal DXA scan. Anderson et al. [19] reviewed a method of using such opportunistic imaging in osteoporosis screening based on the Hounsfield unit (HU). The method was reproducible, with ICC falling between 0.95 and 0.98; it was also highly correlated to bone mineral density and T-scores obtained from DXA [19]. The possibility of analyzing the HU from available CT has further been analyzed in the assessment of sarcopenia, citing a decrease in the HU as fatty infiltration occurs [19,20]. Keller et al. [21] concluded that both the measurement of cross-sectional area and density of paraspinal muscles based on axial CT scans are reliable techniques. Thus, the use of the HU to assess muscle density is a technique that we would

look to combine with our current cross-sectional measurements of muscle-to-VB ratio in our attempt to most accurately quantify sarcopenia.

It is important to recognize the potential limitations of our method for identifying sarcopenia. First, our data were derived from a single institution; thus, the impact of generalizing them to patient populations in other institutions is unknown. Second, it is unknown how a different CT scanner or poorer image quality would affect the sensitivity of identifying patients with sarcopenia. Although we announced that our method was adequately reproducible between clinicians, we assume a more accurate screening of sarcopenia is possible with a computer algorithm. If this method of defining sarcopenia gains currency, we hypothesize that digital computation would be more precise and time-efficient. Finally, a shortcoming of CT is radiation associated with this imaging modality; thus, it is recommended that the reproducibility of this method be evaluated using magnetic resonance imaging.

We have developed a method to use readily available preoperative imaging with high interobserver reliability to help identify patients with sarcopenia who may benefit from more vigorous preoperative and postoperative exercises and nutrition therapies. Future studies may draw on this method to determine the effect of sarcopenia on postoperative outcomes in both the field of spine surgery and other disciplines; this is a worthwhile evaluation to minimize the risks of postoperative complications. We acknowledge that this is a relatively novel technique for assessing sarcopenia; therefore, future studies are advised to validate our method externally. In the end, we hope we can stress the importance of this underappreciated measure of frailty and try to objectively improve the identification of patients who may benefit from more vigorous nutritional and physical therapies.

Conclusions

In the end, we hope we can stress the importance of this underappreciated measure of frailty and try to objectively improve the identification of patients who may benefit from more vigorous nutritional and physical therapies.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

References

1. Landi F, Calvani R, Ortolani E, et al. The association between sarcopenia and functional outcomes among older patients with hip fracture undergoing in-hospital rehabilitation. Osteoporos Int 2017;28:1569-76.

2. Baumgartner RN, Koehler KM, Gallagher D, et al. Epidemiology of sarcopenia among the elderly in New Mexico. Am J Epidemiol 1998;147:755-63.

3. Janssen I, Heymsfield SB, Ross R. Low relative skeletal muscle mass (sarcopenia) in older persons is associated with functional impairment and physical disability. J Am Geriatr Soc 2002;50:889-96.

4. Newman AB, Kupelian V, Visser M, et al. Sarcopenia: alternative definitions and associations with lower extremity function. J Am Geriatr Soc 2003;51:1602-9.

5. Bijlsma AY, Meskers CG, Westendorp RG, Maier AB. Chronology of age-related disease definitions: osteoporosis and sarcopenia. Ageing Res Rev 2012;11:320-4.

6. Edwards MH, Dennison EM, Aihie Sayer A, Fielding R, Cooper C. Osteoporosis and sarcopenia in older age. Bone 2015;80:126-30.

7. Janssen I, Shepard DS, Katzmarzyk PT, Roubenoff R. The healthcare costs of sarcopenia in the United States. J Am Geriatr Soc 2004;52:80-5.

8. Yoo T, Lo WD, Evans DC. Computed tomography measured psoas density predicts outcomes in trauma. Surgery 2017;162:377-84.

9. Di Monaco M, Vallero F, Di Monaco R, Tappero R. Prevalence of sarcopenia and its association with osteoporosis in 313 older women following a hip fracture. Arch Gerontol Geriatr 2011;52:71-4.

10. Lin HS, Watts JN, Peel NM, Hubbard RE. Frailty and post-operative outcomes in older surgical patients: a systematic review. BMC Geriatr 2016;16:157.

11. Jones K, Gordon-Weeks A, Coleman C, Silva M. Radiologically determined sarcopenia predicts morbidity and mortality following abdominal surgery: a systematic review and meta-analysis. World J Surg 2017;41:2266-79.

12. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15:155-63.

13. El Maghraoui A, Ebo'o FB, Sadni S, Majjad A, Hamza T, Mounach A. Is there a relation between pre-sarcopenia, sarcopenia, cachexia and osteoporosis in patients with ankylosing spondylitis? BMC Musculoskelet Disord 2016;17:268.

14. Ji HM, Han J, Won YY. Sarcopenia and osteoporosis. Hip Pelvis 2015;27:72-6.

15. He H, Liu Y, Tian Q, Papasian CJ, Hu T, Deng HW. Relationship of sarcopenia and body composition with osteoporosis. Osteoporos Int 2016;27:473-82.

16. Steihaug OM, Gjesdal CG, Bogen B, Kristoffersen MH, Lien G, Ranhoff AH. Sarcopenia in patients with hip fracture: a multicenter cross-sectional study. PLoS One 2017;12:e0184780.

17. Levett DZ, Edwards M, Grocott M, Mythen M. Preparing the patient for surgery to improve outcomes. Best Pract Res Clin Anaesthesiol 2016;30:145-57.

18. Ng TP, Feng L, Nyunt MS, et al. Nutritional, physical, cognitive, and combination interventions and frailty reversal among older adults: a randomized controlled trial. Am J Med 2015;128:1225-36.

19. Anderson PA, Polly DW, Binkley NC, Pickhardt PJ. Clinical use of opportunistic computed tomography screening for osteoporosis. J Bone Joint Surg Am 2018;100:2073-81.

20. Lee SH, Park SW, Kim YB, Nam TK, Lee YS. The fatty degeneration of lumbar paraspinal muscles on computed tomography scan according to age and disc level. Spine J 2017;17:81-7.

21. 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;28:1455-60.