Quantitative Magnetic Resonance Imaging Analysis of the Cervical Spine Extensor Muscles: Intrarater and Interrater Reliability of a Novice and an Experienced Rater

Maryse Fortin[1,2], Octavian Dobrescu[3], Peter Jarzem[1,3], Jean Ouellet[1,3], Michael H. Weber[1,3]

[1]Division of Orthopedics, Department of Surgery, McGill University, Montreal, QC, Canada
[2]Perform Centre, Concordia University, Montreal, QC, Canada
[3]Faculty of Medicine, McGill University, Montreal, QC, Canada

Study Design: Reliability study.

Purpose: To examine the reliability of novice and experienced raters for measurements of the size and composition of the cervical extensor muscles using a thresholding technique.

Overview of Literature: Although some authors have reported on the dependability of magnetic resonance imaging (MRI) measurements of the cervical muscles, there remains some variability regarding intrarater and interrater reliabilities, and few studies have examined the associated measurement error. Whether the rater's experience noticeably influences the reliability and precision of such measurements has also not been examined.

Methods: A sample of 10 patients with cervical pathologies was selected. Muscle cross-sectional area (CSA), functional cross-sectional area (FCSA), and signal intensity of the cervical extensor muscles were acquired from axial T2-weighted MRIs by a novice and an experienced rater. All measurements were obtained twice, at least 5 days apart, while the raters were blinded to all earlier measurements.

Results: Interrater reliability estimates (intraclass correlation coefficients) varied between 0.84 and 0.99 for the novice rater and between 0.94 and 0.99 for the experienced rater, indicating excellent reliability. The standard error of measurement for the novice rater was, however, noticeably higher for all cervical muscle measurements. Most of the interrater estimates showed excellent agreement with the exception of CSA measurement of the semispinalis cervicis at C4–C7 and FCSA measurement of the multifidus and semispinalis cervicis at C4–C7, which showed poor interrater reliability.

Conclusions: The proposed method of investigating cervical muscle measurements was highly reliable; however, novice raters should receive adequate training before using this method for diagnostic, research, and clinical purposes.

Keywords: Paraspinal muscles; Musculoskeletal abnormalities; Magnetic resonance imaging; Methodological study

Introduction

The paraspinal muscles are deep back and neck muscles that run in parallel on both sides of the spine and attach directly onto the vertebrae, providing mobility of individual segments and stability of the spine [1,2]. Morphometric alterations in the paraspinal muscles (e.g., atrophy, fatty infiltration, asymmetry) have been reported in patients with low back pain [3-5], and recent imaging studies have detected similar changes in the cervical muscles of patients with chronic neck pain [6,7]. Patients with persistent whiplash-associated disorders have been found to have more fatty infiltration in the multifidus and semispinalis cervicis muscles than healthy controls [7]. Significant multifidus atrophy has also been reported in patients with unilateral cervical radiculopathy [6]. Similarly, magnetic resonance imaging (MRI) signal changes in the ipsilateral multifidus extending from one level superior to two levels inferior to the injured level was reported with cervical root avulsion injury [8]. Moreover, reduced deep cervical flexor muscle activation during craniocervical flexion and delayed activation during postural perturbations were observed in patients with neck pain [9,10]. Such evidence highlights the clinical importance of assessing muscle dysfunction in patients with chronic neck pain.

Quantitative cervical muscle measurements can be obtained by real-time ultrasonography [11,12] and MRI [6,13,14]. MRI remains the gold standard modality for muscle imaging, as the high resolution allows for the assessment of muscle size and composition (e.g., fatty infiltration). However, most MRI studies have only examined the association between the cross-sectional area (CSA) of the cervical muscles and symptoms [14-17]. Although Elliott et al. [18] developed a method of calculating an index of fat within the cervical muscles (i.e., muscle total signal intensity/signal intensity of a standardized intermuscular fat region), this technique does not allow for the exact determination of muscle functional cross-sectional area (FCSA, fat-free area). Muscle FCSA has been suggested to be a good indicator of muscle atrophy and contractibility and can be quantified using a thresholding technique based on the difference in signal intensity between muscle and fat tissue. However, we are aware of only one study investigating cervical muscles that used such a technique, and the reliability was not examined [6]. Although some authors have reported on the dependability of similar MRI muscle measurements, there remains some variability regarding intrarater and interrater reliabilities, and few studies have examined the associated measurement error. Whether the rater's experience noticeably influences the reliability and precision of cervical muscle MRI measurements has also not been examined.

The purpose of this study was to determine the reliability and standard error of measurements (SEMs) of cervical extensor muscle CSA and composition using a thresholding technique, and to compare intrarater reliability and examine interrater reliability between a novice and an experienced rater.

Materials and Methods

1. Patient sample

A sample of 10 patients (six women and four men) was selected from an internal research database, which included patients with commonly diagnosed spine pathologies. The subjects included in this study were diagnosed with either cervical spondylotic myelopathy or cervical stenosis and had symptoms severe enough to be referred to a spine surgeon. Patients were excluded if they were younger than 18 years, were not able to undergo MRI acquisition, had previous cervical spine surgery, or were pregnant. Informed consent was obtained from all participants. This study was approved by the Research Ethics Board of the McGill University Health Centre (13-436-GEN).

2. Muscle measurements

All muscle measurements were performed by a novice rater (O.D., a medical student) and an experienced rater (M.F., with more than 5 years of experience in MRI paraspinal muscle measurements). In preparation for the measurements, the novice rater received training from the experienced rater on how to segment the different cervical extensor muscles and determine muscle FCSA using a thresholding technique. For practice, the novice rater analyzed a sample of six patients before the start of the measurement study. Each rater obtained the muscle measurements twice, a minimum of 5 days apart, while blinded to each other's measurements and patients' clinical information. After the first set of measurements was completed, the images were reordered and blinded to be similarly assessed again.

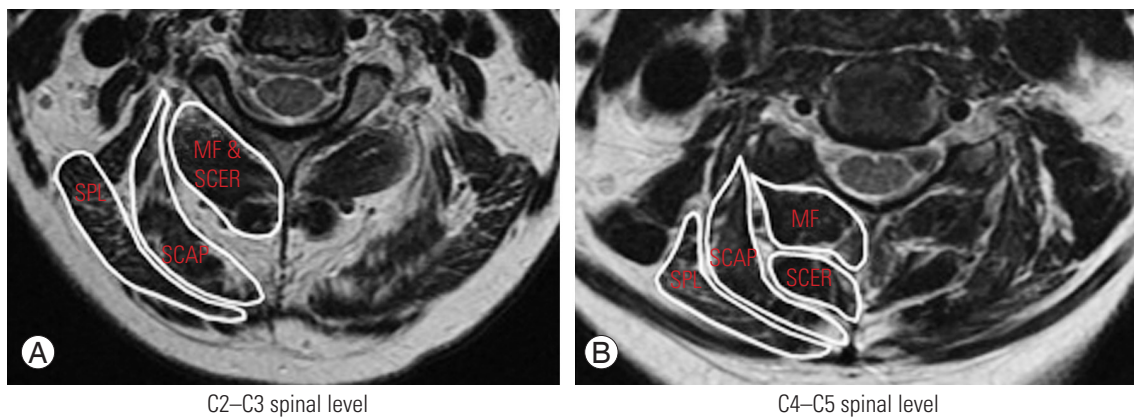Quantitative measurements of the cervical multifidus,

C2–C3 spinal level	C4–C5 spinal level

Fig. 1. Measurement of total cross-sectional area for the MF, SCER, SCAP, and SPL muscles at C2–C3 **(A)** and C4–C5 **(B)**. The MF and SCER muscles were measured together at the C2–C3 spinal level. MF, multifidus; SCER, semispinalis cervicis; SCAP, semispinalis capitis; SPL, splenius capitis.
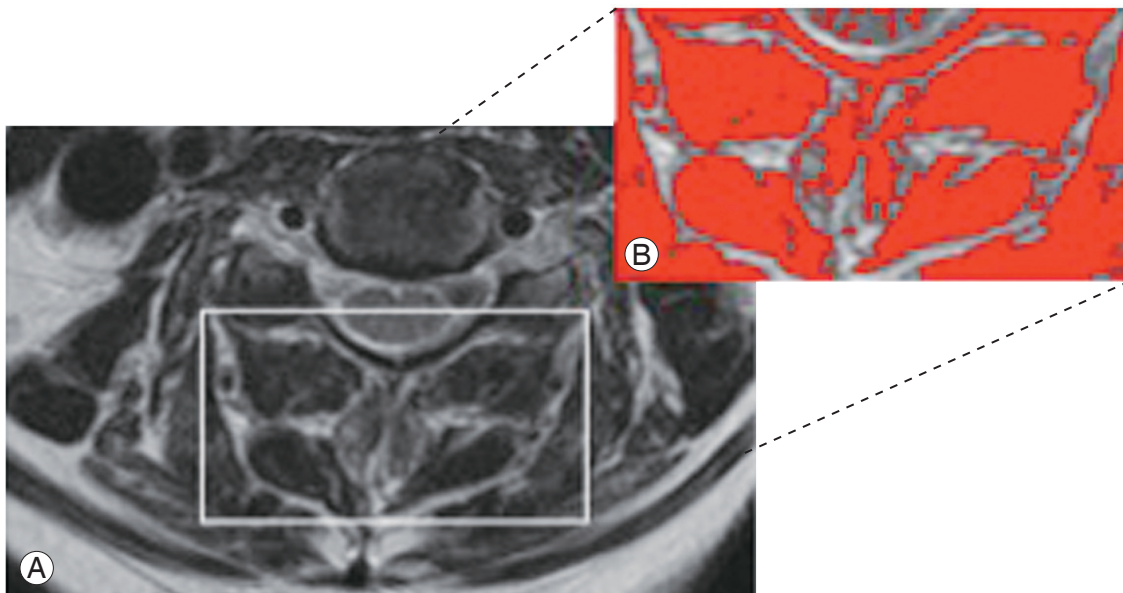

Fig. 2. (A) Axial T2-weighted magnetic resonance imaging image used to acquire cervical muscle measurements. **(B)** The image shows the application of a signal threshold filter (ImageJ analysis software, ver. 1.43; National Institutes of Health, Bethesda, Maryland, USA) to highlight the fat-free muscle area and obtain the functional cross-sectional area measurement (in red).

semispinalis cervicis, semispinalis capitis, and splenius capitis extensor muscles were taken from T2-weighted axial MRI images using ImageJ imaging software (ver. 1.43; National Institutes of Health, Bethesda, MD, USA; downloadable at http://rsbweb.nih.gov/ij/download.html) after 3-demensional multiplanar reconstruction using the 32-bit OsiriX software program (ver. 3.8.1; Pixmeo, Geneva, Switzerland) to position the image slices perpendicular to the muscle mass. The muscle measurements of interest included the following: total CSA (muscle size) (Fig. 1), FCSA representing the fat-free area (Fig. 2), and

signal intensity (directly obtained from the CSA measurement), which was used as an indicator of fatty infiltration (high signal intensity indicates more fatty infiltration). All muscle measurements were obtained at the C2–C3, C3–C4, C4–C5, C5–C6, and C6–C7 levels, through the center of each intervertebral disc. Exceptionally, the multifidus and semispinalis cervicis muscles were measured together at C2–C3 due to the large amount of periarticular fat and lack of identifiable muscle boundaries at this level. Muscle FCSA was measured by selecting a threshold signal within the total muscle CSA to include only pixels within the

lean muscle tissue range. Because the signal intensity of homogeneous tissue may vary between subjects and on one scan slide within subjects [19], the gray scale range for lean muscle tissue was established for every subject and scan slice. This paraspinal muscle thresholding technique has been described in detail elsewhere [20].

3. Statistical analysis

Statistical analysis was performed with the IBM SPSS ver. 20.0 (IBM Corp., Armonk, NY, USA). The intrarater reliability for each rater and the interrater reliability between the raters were determined by computing the intraclass correlation coefficient (ICC) for each measurement variable and every muscle of interest. The ICC (2,1) was calculated using a two-way random-effects model and absolute agreement; with this form of ICC, each subject is measured by each rater, and the raters are considered to be representative of a larger population of similar raters (i.e., reliability is calculated from a single measurement). The first set of measurements of each rater was used to calculate the interrater reliability ICCs. The reliabilities for the upper (C2–C3, C3–C4) and lower (C4–C5, C5–C6, C6–C7) cervical levels were assessed separately. The ICCs were interpreted using the following classification, as suggested by Portney and Watkins [21]: 0.00–0.49 poor,

Table 1. Mean (standard deviation) of cervical extensor muscle measurements by the novice and the experienced rater

Muscle	CSA (mm^2)		FCSA (mm^2)		Signal intensity	
	Novice	Experienced	Novice	Experienced	Novice	Experienced
C2–C3						
MF & SCER	231.77±83.37	242.65±66.86	146.20±76.97	86.79±55.14	182.39±71.24	213.25±99.92
SCAP	426.28±163.67	441.55±153.03	317.68±146.11	241.71±117.24	157.17±87.34	149.78±78.38
SPL	404.15±109.81	384.40±118.81	362.99±115.50	286.92±107.12	115.68±60.18	110.70±56.53
C3–C4						
MF	175.94±47.70	183.40±47.39	99.73±26.64	67.62±26.22	207.86±103.29	225.21±121.98
SCER	128.35±37.65	123.41±28.35	106.41±29.40	88.28±27.34	122.16±63.80	117.49±65.43
SCAP	459.72±149.03	371.95±137.91	317.57±137.24	231.39±103.31	172.86±66.75	126.22±53.02
SPL	312.95±111.03	294.64±119.10	277.75±108.21	215.57±94.44	114.84±63.93	102.85±55.27
C4–C5						
MF	210.40±36.44	217.72±50.54	146.42±32.26	103.89±37.15	153.04±73.60	165.41±81.43
SCER	178.74±57.23	171.73±45.87	138.96±42.53	113.16±35.92	134.33±70.54	127.62±67.40
SCAP	358.08±104.23	366.99±113.60	273.09±98.29	220.32±108.35	142.47±65.47	137.12±63.35
SPL	242.06±55.89	221.91±58.71	206.51±53.99	149.57±61.19	117.71±65.73	109.53±60.83
C5–C6						
MF	218.43±69.78	228.71±68.82	127.79±53.21	78.79±30.71	185.46±96.34	191.73±92.63
SCER	210.20±57.70	183.56±53.08	153.99±38.04	119.48±34.47	150.33±73.64	135.03±72.29
SCAP	319.30±82.26	322.61±104.80	254.32±80.73	187.05±93.65	139.60±71.23	135.74±70.44
SPL	251.25±77.76	208.43±70.02	217.41±71.91	141.69±63.84	115.46±56.63	103.44±46.66
C6–C7						
MF	245.14±103.92	209.03±68.84	124.07±57.48	68.90±34.73	209.19±104.96	210.52±110.13
SCER	215.35±78.92	173.15±51.15	164.87±59.31	118.70±33.99	140.50±75.74	128.26±75.44
SCAP	282.80±60.48	293.70±83.87	218.82±55.74	186.79±95.68	143.26±83.37	131.36±70.22
SPL	211.69±80.02	195.30±71.00	182.70±76.72	133.13±70.50	118.32±63.64	107.50±56.13

Values are presented as mean±standard deviation.
CSA, cross-sectional area; FCSA, functional cross-sectional area; MF, multifidus muscle; SCER, semispinalis cervicis muscle; SCAP, semispinalis capitis muscle; SPL, splenius capitis muscle.

0.50–0.74 moderate, and 0.75–1.00 excellent. The SEM was also calculated to provide an estimate of the expected error related to each muscle measurement.

Results

Descriptive data (mean and standard deviation) of the cervical muscle measurements for the novice and experienced rater are shown in Table 1.

1. Intrarater reliability

The intrarater reliability results for measurements of the right side for the novice and the experienced rater are presented in Table 2. The results for the left side were virtually equivalent and are not presented. Intrarater reliability was excellent for all measurements of CSA, FCSA, and mean signal intensity and varied between 0.84 and 0.99 for the novice rater and between 0.94 and 0.99 for the experienced rater. The ICCs for all the different muscle measurements were comparable across muscles and spinal levels (upper versus lower cervical levels). The SEM values were relatively small and helped to confirm the accuracy of the different cervical muscle measurements, although the SEM values for the experienced rater were markedly smaller than those for the novice rater.

2. Interrater reliability

The interrater reliability results for measurements of the right side for the novice and the experienced rater are presented in Table 3. Again, the results for the left side were virtually equivalent and are not presented. The interrater ICCs for the different measurements of CSA and signal intensity were all excellent and varied between 0.75 and 0.98, with the exception of the measurement of CSA for the semispinalis cervicis at C4–C7, which showed moderate interrater reliability (ICC 0.58). Most of the interrater ICCs for the FCSA measurements showed moderate interrater agreement and varied between 0.60 and 0.77, except for the measurements of the multifidus and semispinalis cervicis at C4–C7, which showed poor interrater reliability. The lower agreement between the two raters for the FCSA measurements was also confirmed by the wider 95% confidence interval (CI) and the larger SEMs.

Table 2. Intrarater reliability of cervical extensor muscle measurements by the novice and the experienced rater

Muscle	CSA				FCSA				Signal intensity			
	Novice		Experienced		Novice		Experienced		Novice		Experienced	
	ICC (95% CI)	SEM	ICC (95% CI)	SEM	ICC (95% CI)	SEM	ICC (95% CI)	SEM	ICC (95% CI)	SEM	ICC (95% CI)	SEM
C2–C4												
MF	0.96 (0.91–0.98)	13.28	0.99 (0.98–0.99)	4.53	0.94 (0.85–0.97)	14.92	0.98 (0.96–0.99)	5.11	0.97 (0.943–0.99)	13.25	0.99 (0.98–0.99)	7.69
SCER	0.96 (0.92–0.98)	14.50	0.99 (0.98–0.99)	5.59	0.94 (0.85–0.97)	14.76	0.97 (0.94–0.99)	6.28	0.96 (0.904–0.98)	14.17	0.99 (0.98–0.99)	8.01
SCAP	0.94 (0.86–0.97)	36.28	0.99 (0.99–1.00)	8.01	0.95 (0.89–0.98)	29.59	0.98 (0.95–0.99)	14.84	0.93 (0.842–0.97)	19.69	0.99 (0.98–0.99)	5.13
SPL	0.92 (0.80–0.96)	32.95	0.99 (0.98–0.99)	8.81	0.94 (0.85–0.97)	28.51	0.98 (0.95–0.99)	12.85	0.99 (0.985–0.99)	4.27	0.99 (0.99–1.00)	2.44
C4–C7												
MF	0.91 (0.82–0.95)	21.63	0.99 (0.98–0.99)	5.85	0.91 (0.83–0.96)	13.81	0.96 (0.93–0.98)	6.70	0.98 (0.971–0.99)	10.53	0.99 (0.99–0.99)	5.95
SCER	0.90 (0.80–0.95)	20.31	0.98 (0.96–0.99)	6.53	0.90 (0.81–0.95)	14.37	0.93 (0.87–0.97)	8.32	0.97 (0.948–0.98)	11.24	0.99 (0.98–0.99)	6.20
SCAP	0.83 (0.66–0.92)	35.16	0.98 (0.97–0.99)	10.77	0.87 (0.67–0.94)	28.98	0.99 (0.97–0.99)	9.23	0.99 (0.987–0.99)	5.51	0.99 (0.98–0.99)	4.65
SPL	0.84 (0.69–0.92)	28.56	0.98 (0.97–0.99)	7.46	0.89 (0.78–0.94)	22.28	0.98 (0.97–0.99)	7.50	0.91 (0.827–0.95)	17.58	0.96 (0.93–0.98)	9.48

CSA, cross-sectional area; FCSA, functional cross-sectional area; ICC, intraclass correlation coefficient; CI, confidence interval; SEM, standard error of measurement; MF, multifidus muscle; SCER, semispinalis cervicis muscle; SCAP, semispinalis capitis muscle; SPL, splenius capitis muscle.

Table 3. Interrater reliability of cervical extensor muscle measurements between the novice and the experienced rater

Muscle	C2–C4		C4–C7	
	ICC (95% CI)	SEM	ICC (95% CI)	SEM
CSA				
MF	0.86 (0.69–0.94)	23.71	0.75 (0.54–0.87)	30.58
SCER	0.88 (0.74–0.95)	26.32	0.58 (0.26–0.78)	31.42
SCAP	0.87 (0.67–0.95)	51.90	0.86 (0.73–0.93)	38.27
SPL	0.93 (0.83–0.97)	31.03	0.80 (0.44–0.92)	28.88
FCSA				
MF	0.60 (–0.09 to 0.87)	27.09	0.38 (–0.10 to 0.71)	28.44
SCER	0.64 (–0.04 to 0.88)	25.35	0.45 (–0.03 to 0.74)	24.81
SCAP	0.72 (–0.02 to 0.91)	56.88	0.73 (0.10–0.90)	49.79
SPL	0.77 (–0.05 to 0.94)	49.30	0.62 (–0.07 to 0.88)	38.65
Signal intensity				
MF	0.92 (0.69–0.96)	34.03	0.98 (0.96–0.99)	11.12
SCER	0.87 (0.72–0.95)	33.45	0.96 (0.85–0.98)	13.52
SCAP	0.87 (0.30–0.96)	23.23	0.97 (0.94–0.99)	9.75
SPL	0.97 (0.85–0.99)	8.45	0.94 (0.82–0.97)	12.08

ICC, intraclass correlation coefficient; CI, confidence interval; SEM, standard error of measurement; CSA, cross-sectional area; MF, multifidus muscle; SCER, semispinalis cervicis muscle; SCAP, semispinalis capitis muscle; SPL, splenius capitis muscle. FCSA, functional cross-sectional area.

Discussion

The primary purpose of this study was to assess the intrarater and interrater reliabilities in a novice and an experienced rater to obtain CSA and composition measurements of the cervical extensor muscles using a thresholding technique. The intrarater ICC estimates for CSA, FCSA, and signal intensity measurements were all greater than 0.84, indicating excellent reliability for both the novice and the experienced rater. The intrarater reliability estimates were comparable and varied between 0.84 and 0.99 for the novice rater and between 0.94 and 0.99 for the experienced rater. As intrarater reliability estimates were also consistent across muscles and upper versus lower cervical levels, our results suggest that muscle size, within the range studied and the spinal level (C2–C7), do not influence the interrater reliability. However, the SEM for the novice rater was noticeably higher for all three cervical muscle measurements, suggesting that the level of experience does play a role in the reliability and precision of measurements. The interrater ICC estimates also showed good to excellent reliability and agreement between the novice and the experienced rater, except for the FCSA

measurements, which revealed lower ICCs and a wider 95% CI.

Our findings related to intrarater reliability are similar to those of other studies examining measurements of total CSA of the cervical spine extensor muscles. Ulbrich et al. [22] reported intrarater ICCs varying between 0.78 and 0.98 for CSA measurement of the deep extensor muscles at the C2–C5 spinal levels, whereas other authors reported ICCs for intrarater reliability that were slightly higher (0.82 to 0.99) [16,23,24]. Previous studies measuring total CSA reported interrater ICCs varying between 0.52 and 0.85 [22,24,25], which also corroborates our findings. However, most of these studies were limited by failure to report the ICCs for individual muscles and cervical levels, failure to report the numbers of patients or slices measured for the reliability analysis, lack of assessment of muscle composition, and failure to report the associated measurement error. In addition, whether the rater's experience influences the reliability and precision of such cervical muscle measurements has been neglected. A few studies have investigated measures of cervical muscle composition (e.g., fatty infiltration) [7,16,18] and reported interrater reliability indices for cervical muscle fat index or ratio

varying between 0.82 and 0.98 [7,16,18] and interrater ICCs between 0.75 and 0.97 [7,18]. Only Elliott et al. [23] investigated the measurement error associated with MRI cervical extensor muscle measurements; a SEM of 0.96 to 4.87 was reported for the interrater reliability of cervical extensor muscle CSA measurements at the C3–C7 levels, which is somewhat similar to the SEM obtained by the experienced rater in our study. However, our results add to the existing literature by demonstrating that although intrarater reliability estimates were similar between a novice and an experienced rater, the SEM for measurements by the novice rater was markedly higher, suggesting lower precision of measurement.

A threshold technique was used to calculate FCSA based on differences in pixel intensity between muscle (low intensity) and fat tissues (high intensity) on T2-weighted axial images. This technique has been described in detail previously and was found to be highly reliable to assess lumbar paraspinal muscle composition [20]. We are not aware of any other studies using a thresholding technique for the assessment of cervical spine muscle morphology, with the exception of Chae et al. [6]. However, the technique used was not clearly described, and reliability indices were not reported. Although Elliott et al. [18] developed a method of calculating an index of fat within the cervical muscles, this technique does not allow for the exact determination of muscle FCSA or the absolute concentration of fat tissue. Muscle atrophy can occur without a reduction in the total CSA and may be characterized by the replacement of muscle with fat or fibrous tissue. Consequently, FCSA (area of lean muscle tissue) is a better indicator of muscle atrophy and contractibility. Our results suggest that the described thresholding technique was highly reliable to determine cervical muscle FCSA in a clinically relevant population. However, although both the novice and the experienced rater had excellent intrarater ICCs, some of the interrater ICCs for FCSA were lower, suggesting a lack of agreement between the two raters. A thorough examination of the results revealed that this discrepancy was due to the novice rater's setting higher threshold values for the gray scale range of lean muscle tissue, which led to larger FCSA, as shown in Table 1. Although care was taken to avoid the inclusion of any visible pixel of fat when tracing the sample regions of interest for determination of the gray scale range, variations in signal intensity in atrophied cervical muscle can be subtle, and the lack of experience of the novice rater in reading MRI

images may explain the difference in threshold values. Since the novice rater only practiced this technique on a series of six patients prior to the beginning of this study, one can assume that supplementary training would have addressed this issue and improved the agreement between the two raters. In addition to FCSA, we also examined the reliability of measurements of the mean signal intensity of total CSA, which can be used as an indicator of muscle fatty infiltration. Our results showed excellent intrarater ICCs for both the novice and the experienced rater, with excellent agreement between the raters.

Cervical muscle alterations, such as muscle atrophy, fatty infiltration, asymmetry, and delayed muscular activation, have been reported in patients with chronic neck pain [6-10]. However, different techniques and measurements (e.g., signal intensity, fat-index ratio, and qualitative scale) have been used to assess cervical muscle composition. The thresholding technique presented in this study was highly reliable to assess cervical muscle FCSA, a better indicator of atrophy and muscle composition. Using this technique to identify variations in cervical muscle morphometry among different symptomatic and asymptomatic populations would be beneficial to facilitate comparison among studies and provide basic data for future cross-sectional and longitudinal investigations. Preliminary evidence suggests that recognition of changes in cervical muscle morphometry may be helpful diagnostically and clinically, but further studies using an accurate and quantitative measurement approach are needed to clarify their relationship with neck pain, cervical pathology, functional status, and muscular strength.

The limitations of this study include the small sample size and its restriction to only two raters, which somewhat limits the generalizability of our results to other examiners. However, taking into account that replicate measurements by the novice and the experienced rater were obtained on the same image to remove any extraneous potential source of measurement, it is very likely that examiners with comparable levels of experience would have produced similar results. We also accounted for this limitation in our statistical analysis by using the ICC (2,1), which maximizes the generalizability of our results.

Conclusions

Measurements of the CSA and composition of the cervical extensor muscles were obtained by a novice and an

experienced rater using a thresholding technique. Overall, good to excellent interrater reliability and intrarater reliability were demonstrated for both raters. However, the findings of the novice rater had larger SEM values, suggesting lower measurement precision, and a few FCSA measures showed low agreement between the two raters. In general, our results suggest that the described thresholding technique was highly reliable to determine cervical muscle FCSA in a clinically relevant population. Future studies would benefit from using a similar standard protocol to investigate measurements of cervical muscle composition. However, novice raters should receive adequate training, and their reliability and measurement precision should be assessed before this method is used for diagnostic, research, and clinical purposes.

Conflict of Interest

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

References

1. Wilke HJ, Wolf S, Claes LE, Arand M, Wiesend A. Stability increase of the lumbar spine with different muscle groups: a biomechanical in vitro study. Spine (Phila Pa 1976) 1995;20:192-8.
2. Solomonow M, Zhou BH, Harris M, Lu Y, Baratta RV. The ligamento-muscular stabilizing system of the spine. Spine (Phila Pa 1976) 1998;23:2552-62.
3. Danneels LA, Vanderstraeten GG, Cambier DC, Witvrouw EE, De Cuyper HJ. CT imaging of trunk muscles in chronic low back pain patients and healthy control subjects. Eur Spine J 2000;9:266-72.
4. Kulig K, Scheid AR, Beauregard R, Popovich JM Jr, Beneck GJ, Colletti PM. Multifidus morphology in persons scheduled for single-level lumbar microdiscectomy: qualitative and quantitative assessment with anatomical correlates. Am J Phys Med Rehabil 2009;88:355-61.
5. Mengiardi B, Schmid MR, Boos N, et al. Fat content of lumbar paraspinal muscles in patients with chronic low back pain and in asymptomatic volunteers: quantification with MR spectroscopy. Radiology 2006;240:786-92.
6. Chae SH, Lee SJ, Kim MS, Kim TU, Hyun JK. Cervical multifidus muscle atrophy in patients with uni-

7. lateral cervical radiculopathy. J Korean Acad Rehabil Med 2010;34:743-51.
7. Elliott J, Jull G, Noteboom JT, Darnell R, Galloway G, Gibbon WW. Fatty infiltration in the cervical extensor muscles in persistent whiplash-associated disorders: a magnetic resonance imaging analysis. Spine (Phila Pa 1976) 2006;31:E847-55.
8. Hayashi N, Masumoto T, Abe O, Aoki S, Ohtomo K, Tajiri Y. Accuracy of abnormal paraspinal muscle findings on contrast-enhanced MR images as indirect signs of unilateral cervical root-avulsion injury. Radiology 2002;223:397-402.
9. Falla D, Jull G, Hodges PW. Feedforward activity of the cervical flexor muscles during voluntary arm movements is delayed in chronic neck pain. Exp Brain Res 2004;157:43-8.
10. Falla DL, Jull GA, Hodges PW. Patients with neck pain demonstrate reduced electromyographic activity of the deep cervical flexor muscles during performance of the craniocervical flexion test. Spine (Phila Pa 1976) 2004;29:2108-14.
11. Kristjansson E. Reliability of ultrasonography for the cervical multifidus muscle in asymptomatic and symptomatic subjects. Man Ther 2004;9:83-8.
12. Rezasoltani A, Ali-Reza A, Khosro KK, Abbass R. Preliminary study of neck muscle size and strength measurements in females with chronic non-specific neck pain and healthy control subjects. Man Ther 2010;15:400-3.
13. Elliott JM, Jull GA, Noteboom JT, Durbridge GL, Gibbon WW. Magnetic resonance imaging study of cross-sectional area of the cervical extensor musculature in an asymptomatic cohort. Clin Anat 2007;20:35-40.
14. Matsumoto M, Ichihara D, Okada E, et al. Cross-sectional area of the posterior extensor muscles of the cervical spine in whiplash injury patients versus healthy volunteers: 10 year follow-up MR study. Injury 2012;43:912-6.
15. Fernandez-de-las-Penas C, Albert-Sanchis JC, Buil M, Benitez JC, Alburquerque-Sendin F. Cross-sectional area of cervical multifidus muscle in females with chronic bilateral neck pain compared to controls. J Orthop Sports Phys Ther 2008;38:175-80.
16. De Loose V, van den Oord M, Keser I, et al. MRI study of the morphometry of the cervical musculature in F-16 pilots. Aviat Space Environ Med

2009;80:727-31.

17. Airi Oksanen, Erkintalo M, Metsahonkala L, et al. Neck muscles' cross-sectional area in adolescents with and without headache: MRI study. Eur J Pain 2008;12:952-9.

18. Elliott JM, Galloway GJ, Jull GA, Noteboom JT, Centeno CJ, Gibbon WW. Magnetic resonance imaging analysis of the upper cervical spine extensor musculature in an asymptomatic cohort: an index of fat within muscle. Clin Radiol 2005;60:355-63.

19. Ranson CA, Burnett AF, Kerslake R, Batt ME, O'Sullivan PB. An investigation into the use of MR imaging to determine the functional cross sectional area of lumbar paraspinal muscles. Eur Spine J 2006;15:764-73.

20. Fortin M, Battié MC. Quantitative paraspinal muscle measurements: inter-software reliability and agreement using OsiriX and ImageJ. Phys Ther 2012;92:853-64.

21. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. 2nd ed. Upper Saddle River: Prentice Hall; 2000.

22. Ulbrich EJ, Aeberhard R, Wetli S, et al. Cervical muscle area measurements in whiplash patients: acute, 3, and 6 months of follow-up. J Magn Reson Imaging 2012;36:1413-20.

23. Elliott J, Jull G, Noteboom JT, Galloway G. MRI study of the cross-sectional area for the cervical extensor musculature in patients with persistent whiplash associated disorders (WAD). Man Ther 2008;13:258-65.

24. Okada E, Matsumoto M, Ichihara D, et al. Cross-sectional area of posterior extensor muscles of the cervical spine in asymptomatic subjects: a 10-year longitudinal magnetic resonance imaging study. Eur Spine J 2011;20:1567-73.

25. Thakar S, Mohan D, Furtado SV, et al. Paraspinal muscle morphometry in cervical spondylotic myelopathy and its implications in clinicoradiological outcomes following central corpectomy: clinical article. J Neurosurg Spine 2014;21:223-30.