Inter-Rater Reliability of Ultrasound Imaging of the Trunk Musculature Among Novice Raters

Deydre S. Teyhen, PT, PhD, Steven Z. George, PT, PhD, Jessica L. Dugan, PT, Jared Williamson, DPT, Brett D. Neilson, DPT, John D. Childs, PT, PhD

Objective—The purpose of this study was to determine the inter-rater reliability of ultrasound imaging for assessing trunk muscle morphologic characteristics at rest and while contracted among different pairs of novice raters. The secondary purpose was to compare 3 different measurement techniques for assessing lateral abdominal muscle thickness.

Methods—A single-group repeated measures reliability study was conducted on 21 healthy participants (mean ± SD, 21.5 ± 4.4 years; 5 female and 16 male) without low back pain. Ultrasound images of the transversus abdominis, internal oblique, rectus abdominis, and lumbar multifidus muscles were obtained by different pairs of novice raters in a counterbalanced order. All raters received a standardized training program before obtaining measurements.

Results—The intraclass correlation coefficient (1, 3) point estimates ranged from 0.86 to 0.94; the standard error of the measurement ranged from 0.04 to 0.16 cm for the thickness values and 0.67 cm² for the cross-sectional area of the rectus abdominis muscle. There was no meaningful difference between the different measurement techniques used to analyze the lateral abdominal muscles.

Conclusions—Good to excellent reliability was obtained for all measures by novice raters. Minimal differences in reliability were noted between the different measurement techniques to assess lateral abdominal muscle thickness.

Key Words—internal oblique; lumbar multifidus; rectus abdominis; transversus abdominis; ultrasound

Motor control exercise therapy (ie, lumbar stabilization) has been demonstrated to be an effective management strategy for improving pain and function for those with nonspecific low back pain.1–3 Moreover, researchers have identified a subgroup of patients with low back pain who appear to preferentially benefit from motor control exercises.4–6 In general, this subgroup comprises individuals with deficits in the deep trunk musculature, such as the transversus abdominis7–11 and the lumbar multifidus.12–15 Although researchers have increasingly identified associations between low back pain and underlying neuromuscular control deficits,16 reliable and valid measurement tools that are noninvasive, provide real-time feedback, and are useful for research and clinical decision making have been scarce.
Recently, researchers have been assessing the role of rehabilitative ultrasound imaging\(^17\) to aid in the evaluation and management of those with low back pain. In a recent systematic review, Koppenhaver et al\(^18\) concluded that ultrasound imaging is a valid measure of trunk muscle size, provides an indirect assessment of muscle activation, and is sensitive to change. From a construct validity perspective, researchers have been able to demonstrate deficits in muscle size (atrophy) and altered activity of the lateral abdominal muscles\(^9,19\) and the lumbar multifidus.\(^12,14,15\) From a criterion validity perspective, researchers have been able to demonstrate an association between ultrasound imaging measurements of muscle morphologic characteristics with both magnetic resonance imaging and electromyography.\(^20–23\) Additionally, ultrasound imaging can be used to provide real-time feedback to both the provider and the patient about exercise performance and motor learning to aid in rehabilitation.\(^24–27\) Clinically, the role of ultrasound imaging to aid rehabilitation can be appreciated in a recent prospective cohort study by Hides et al.\(^28\) In this study, the researchers were able to demonstrate a decreased cross-sectional area and asymmetry of the lumbar multifidus in elite cricketers with low back pain compared to those without low back pain. Furthermore, they were able to demonstrate that motor control training that included ultrasound guided biofeedback training was able to improve the cross-sectional area and symmetry of the lumbar multifidus muscle at the end of the 13-week training program. These improvements in muscle morphologic characteristics were associated with clinical improvements in the athletes’ symptoms.

Although the use of ultrasound imaging in the evaluation of muscle morphologic characteristics and function for low back pain is promising, psychometric properties such as reliability and precision (standard error of the measurement and minimal detectable change) need to be established. In one of the larger reliability studies, Koppenhaver et al\(^29\) demonstrated inter-rater reliability [intraclass correlation coefficient (ICC) \((2,2)\)] point estimates of 0.80 to 0.94 for assessing lateral abdominal and lumbar multifidus muscle thickness. Additionally, researchers have been able to demonstrate that an average of 3 measures of muscle thickness at rest and while contracted resulted in better reliability and precision compared to a single measurement.\(^30,31\) Although 2 systematic reviews concluded that most studies indicated good reliability, further research is needed.\(^32,33\) One identified limitation has been the same sample size, with most studies having fewer than 10 participants, which has resulted in excessively wide confidence intervals.\(^32,33\) Some of the previous studies also used only a single set of images or failed to repeat the entire measurement procedures between raters, thereby not eliminating other potential sources of error.\(^32,33\) Another current limitation in the existing literature is that reliability assessments have only been completed with the muscle at rest and not while contracted.\(^33\) Mixed results have been reported for measures of reliability and precision when assessed by novice raters.\(^29,30,34,35\) Additionally, most inter-rater reliability studies have compared the same set of assessors.\(^35\) Finally, comparisons of different measurement techniques to assess the lateral abdominal muscles have been scarce.\(^9,10,29,30,36\)

On the basis of the limitations in the existing research regarding the psychometric properties of ultrasound imaging, we attempted to address some of the current gaps for using ultrasound imaging to assess the trunk musculature. The primary purpose of this study was to assess the inter-rater reliability of ultrasound imaging for assessing trunk muscle morphologic characteristics at rest and while contracted among different pairs of novice raters. The evaluation included the transversus abdominis, internal oblique, rectus abdominis, and lumbar multifidus muscles. The secondary purpose was to compare 3 different measurement techniques for assessing lateral abdominal muscle thickness. We hypothesized that measurements of muscle morphologic characteristics obtained by ultrasound imaging would be adequately reliable (ICC > 0.75) and have good measurement precision.

**Materials and Methods**

**Participants**

Participants consisted of a subset of US Army soldiers enrolled in a cluster randomized trial on prevention of low back pain in the military.\(^37\) Over a 12-month period, 4325 soldiers undergoing a 16-week training program at Fort Sam Houston, Texas, to become combat medics volunteered to participate in the trial. Soldiers were eligible to participate if they were between 18 and 35 years of age (or emancipated minors), fluent in English, and enrolled in combat medic training. Soldiers were excluded if they had a history of low back pain that resulted in limited work or physical activity greater than 48 hours, seeking medical care, or prior surgery in the lumbar spine region. Soldiers were also excluded if they were unable to participate in unit physical training because of other musculoskeletal injuries, had a history of fracture (stress or traumatic) in the hip and/or pelvis, or were pregnant. Of the 4325 soldiers, a random sample of 200 participants was identified to participate in a physical examination, which included assessment of the abdominal and lumbar multifidus musculature.
This analysis represented a convenience sample of the first 21 participants, who completed a second examination for the purposes of establishing reliability of the study procedures. All participants signed consent forms approved by Brooke Army Medical Center’s Institutional Review Board.

**Examiners**
The examiners who participated in this study were all novice raters and included 2 research physical therapists and 4 physical therapy students enrolled in a doctor of physical therapy training program. Before testing, all examiners underwent training that consisted of 20 hours of hands-on training led by 1 of the coinvestigators (D.S.T.) experienced with the specific ultrasound imaging protocol used in the study. A proficiency evaluation was completed on each rater before data collection.

To minimize bias, investigators worked in pairs. One investigator in each team was designated as the recorder, and the other investigator was designated as the imager. The imager was responsible for positioning the transducer for optimal visualization of the musculature. Both the imager and the recorder had to agree on image quality and placement. If on-screen measurements were obtained, the recorder would annotate measurements and image information. Throughout the process, the imager was blinded to all measurements. Two pairs of raters evaluated each participant; selection of rated pairs was counterbalanced.

**Procedures**
This study was a single-group repeated measures design. Images of the abdominal and lumbar multifidus muscles were acquired in B-mode using a portable ultrasound unit (Titan; SonoSite, Inc, Bothell, WA) with a 5-MHz, 60-mm curvilinear array using techniques previously outlined.38,39 Image acquisitions for each muscle (rectus abdominis, transversus abdominis, internal oblique, and lumbar multifidus) and for each testing condition (rest and contract) were performed 3 times.31 Two pairs of raters were randomly selected to evaluate each participant. Participants changed plinths and walked between evaluations from each pair of raters. Images were obtained bilaterally; however, the results described are related to the right-sided musculature. A total of 60 images were assessed for each participant, resulting in a total of 1260 images analyzed. To help avoid an order effect, the images outlined below were obtained in a counterbalanced order.

Images of the lumbar multifidus at the L4-5 level were obtained with the participants in the prone position. Pillows were placed under the pelvis to minimize the lumbar lordosis. An inclinometer was used to ensure that the lumbar spine was within 10° of the horizontal.23,39,40 The transducer was placed longitudinally along the spine, allowing visualization of the sacrum and the causal lumbar zygapophyseal joints. The transducer was slightly lateral to the spinous process and angled slightly medial until the L4-5 zygapophyseal joint could be identified. Lumbar multifidus thickness measurements were made between the posterior-most portion of the L4-5 zygapophyseal joint and the plane between the muscle and subcutaneous tissue (Figure 1). Images of the lumbar multifidus were obtained at rest and during a submaximal contraction that consisted of a contralateral arm lift.23 The contralateral arm lift was performed by lifting the arm approximately 5 cm off the plinth with the elbows flexed 90° and shoulders abducted 120° while the participant held a hand weight based on body mass.23

Images of the rectus abdominis were obtained with the participant supine. The inferior border of the transducer was placed immediately above the umbilicus and moved laterally from the midline until the muscle cross section was centered on the image. If required, a standoff pad was used to ensure that the entire cross-sectional area of the rectus abdominis was visualized within the field of view. Measurements of muscle thickness and cross-sectional area were obtained at rest (Figure 2). Electronic on-screen calipers were aligned at the muscle belly’s center and measured from the inferior hyperechoic fascial line of the superficial border of the rectus abdominis to the superior hyperechoic fascial line of the inferior border of the rectus abdominis muscle; the cross-sectional area was obtained by tracing the inferior border of the hyperechoic fascial line.41

**Figure 1.** Longitudinal view of the lumbar multifidus using a split-screen facility. The lumbar multifidus is imaged at rest (left) and during a contralateral arm raise (right). Thickness measurements were made between the posterior-most portion of the L4-5 facet joint (F) and the plane between the muscle and subcutaneous tissue.
Images of the transversus abdominis and internal oblique muscles were acquired at rest and during the active straight leg raise maneuver. Ultrasound images were obtained with the transducer positioned on the anterolateral aspect of the abdominal wall, superior to the iliac crest and perpendicular to the midaxillary line. All images were obtained with the middle of the muscle belly centered within the field of view and at the end of a normal exhalation to control for the influence of respiration.

The active straight leg raise maneuver was selected to assess automatic changes in the thickness of the lateral abdominal muscles during a lower extremity task while allowing for comparison with prior publications. The active straight leg raise maneuver was performed with the participants positioned supine with arms resting across their chest, hips and knees extended, and heels placed 20 cm apart. During the active straight leg raise maneuver, the participant is asked to lift the lower extremity 5 cm off the plinth without bending the knee. Images were obtained before lifting the lower extremity while the participant was at rest and after the lower extremity was lifted. Three trials were obtained.

All images were saved, and measurements of lateral abdominal muscle thickness were performed using Image-Pro Plus version 4.5 software (Image Processing Solutions, Inc, Silver Spring, MD). For a sample of 10 participants, 3 measurement techniques were used to assess muscle thickness (Table 1). Technique A consisted of using a single measurement between the superficial and deep borders of the muscle, as visualized by the hyperechoic fascial lines (Figure 3). Technique B consisted of an average of 3 measurements of muscle thickness: the middle of the image and 1 cm to each side from midline (Figure 4). The final measurement technique (technique C) consisted of outlining the superior and deep borders of the muscles across the 2-cm segment of the muscle belly, as outlined in technique B (Figure 5). The mean horizontal difference between those 2 lines represented the muscle thickness value. On the basis of these results, the entire sample of 21 participants was analyzed using technique B (see “Results” and “Discussion” sections).

Data Analysis
The dependent measures for the transversus abdominis, internal oblique, and lumbar multifidus muscles consisted of thickness values at rest and while contracted. The dependent measures for the rectus abdominis muscle consisted of thickness and cross-sectional area values at rest. An average of 3 measurements was used as the measurement of interest. Inter-rater reliability was calculated using the ICC (1, 3) with 95% confidence intervals. Reliability values above 0.75 were considered good, and those below 0.75 were considered poor to moderate. To assess measurement precision, standard error of the measurement and minimal detectable change values were calculated. Data management and statistical analysis were performed using SPSS version 17.0 (SPSS Inc, Chicago, IL).

Figure 2. Cross-sectional scans of the rectus abdominis muscle. A. Measurements of muscle thickness were obtained between the deep and superficial borders of the rectus abdominis muscle. B. Measurement of the cross-sectional area was obtained by tracing the interior border of the rectus abdominis muscle.
Results

Twenty-one participants (mean ± SD, 21.5 ± 4.4 years; 5 female and 16 male) completed the reliability study. Demographic and baseline characteristics of the participants are provided in Table 1. Inter-rater reliability of the 3 measurement techniques used to assess the lateral abdominal muscles is outlined in Table 2 (n = 10). All ICC values ranged from 0.80 to 0.92, while the standard error of the measurement ranged from 0.03 to 0.08 cm. The inter-rater reliability for each dependent measure is outlined in Table 3 (n = 21). All ICC values ranged from 0.86 to 0.94, and the standard error of the measurement ranged from 0.04 to 0.16 cm for the thickness values and 0.67 cm² for the cross-sectional area of the rectus abdominis muscle.

Discussion

The primary purpose of this study was to assess inter-rater reliability among different pairs of novice raters to obtain muscle thickness measurements at rest and during a submaximal contraction of the transversus abdominis, internal oblique, and lumbar multifidus muscles and measurements of the thickness and cross-sectional area of the rectus abdominis at rest. Overall, the ICC point estimates were greater than 0.85, indicating good to excellent reliability. These findings are consistent with the 2 previously published systematic reviews32,33 and support our hypothesis that ultrasound imaging measurements are adequately reliable to assess muscle thickness. Previous researchers have demonstrated the importance of adequate rater training to obtain these measurements.34,35 Although we did not directly assess the benefits of training, our findings do support the notion that even novice raters who complete a standardized training program can reliably obtain these measurements, which is in agreement with previously published reliability studies.7,8,29,30,36,45 Although previous studies have assessed inter-rater reliability among novice raters, studies have been limited to 2 raters.29 Our results also add to the existing literature by demonstrating that inter-rater reliability remains consistent across a variety of rater pairs, which increases the generalizability of our findings.

Lumbar Multifidus

The inter-rater reliability of the lumbar multifidus at rest and during a submaximal test demonstrated ICC point estimates between 0.87 and 0.94, with standard error of the measurement values between 0.14 and 0.16 cm and minimal detectable change values between 0.38 and 0.46 cm. The ICC values reported are slightly lower than some previously published values27,34,48; however, they are similar to those reported by Kiesel et al (ICC = 0.85).23 The discrepancy in ICC point estimates may be related to methodological differences. Most of the previously published values assessed intra-rater reliability or allowed different raters to measure the same image. The standard error of the measurement values reported in this study are similar to those reported by Koppenhaver et al29 and Van et al.27 The minimal detectable change values only repre-

Table 1. Demographic and Baseline Characteristics of the Participants (n = 21)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>21.5 ± 4.4 (18–32)</td>
</tr>
<tr>
<td>Female, n (%)</td>
<td>5 (23.8)</td>
</tr>
<tr>
<td>Height, m</td>
<td>1.7 ± 0.1 (1.5–1.8)</td>
</tr>
<tr>
<td>Weight, kg</td>
<td>70.9 ± 8.7 (54.9–86.2)</td>
</tr>
<tr>
<td>Body mass index, kg/m²</td>
<td>24.1 ± 3.1 (19.3–31.2)</td>
</tr>
<tr>
<td>No. of days exercise/wk</td>
<td>2.7 ± 1.1 (1–5)</td>
</tr>
<tr>
<td>No. of y consistently exercised</td>
<td>2.7 ± 1.5 (1–6)</td>
</tr>
</tbody>
</table>

Values are mean ± SD as applicable.
presented 10% to 16% of the lumbar multifidus muscle thickness values, demonstrating better precision than the measurements obtained for the lateral abdominal muscles.

**Abdominal Muscles**

Although Rankin et al\(^4\) previously reported that of all the abdominal muscles the rectus abdominis is the thickest muscle, limited information is available regarding the reliability of measuring the thickness and cross-sectional area of the rectus abdominis at rest. As would be expected, the values reported in this study are slightly lower than the intra-rater reliability values reported by Rankin et al\(^4\) and are consistent with the other measures of reliability. As research continues to explore the changes in rectus abdominis musculature associated with pregnancy and its potential impact on low back pain, further research to help standardize the measurement technique is indicated.

The inter-rater reliability of the transversus abdominis and internal oblique muscles at rest and while during the active straight leg raise maneuver demonstrated ICC point estimates between 0.86 and 0.93, with standard error of the measurement values between 0.04 to 0.07 cm and minimal detectable change values between 0.10 and 0.19 cm. These values are consistent with previously published values.\(^{29,30,32,50}\) The standard error of the measurement for the transversus abdominis and internal oblique represents about 7% to 10% of the muscle thickness at rest or during the active straight leg raise maneuver. However, the minimal detectable change represents 20% to 25% of the mean muscle thickness values. Future research needs to investigate techniques to improve the precision of these measurements to adequately assess changes in muscle thickness values over time.

**Measurement Technique for the Lateral Abdominal Muscles**

Koppenhaver et al\(^2\) demonstrated that more measurement error was attributed to image acquisition when compared to digital measurements of muscle thickness. This is hypothesized to be largely influenced by the variability in

---

**Figure 4.** Transverse view of the lateral abdominal muscles: transversus abdominis (TrA) and internal oblique (IO). Measurements of muscle thickness were obtained as an average of 3 measures of muscle thickness: the middle of the muscle belly and 1 cm to each side from midline. All measurements were obtained between the superficial and deep borders of the transversus abdominis and internal oblique muscles.

**Figure 5.** Transverse view of the lateral abdominal muscles: transversus abdominis (TrA) and internal oblique (IO). The superficial and deep borders of the transversus abdominis and internal oblique muscles were outlined. Measurements of muscle thickness were obtained by calculating the average distance between the hyperechoic borders of these muscles.
transducer placement on the body, pressure used by the examiner to maintain skin contact, and the transducer’s angulation. However, scant evidence exists to determine whether different techniques used for digital measurement could influence inter-rater reliability. Therefore, the secondary purpose of this study was to determine the optimal measurement technique for assessing lateral abdominal muscle thickness. The most common technique is to use a single measurement recorded as the distance between the superficial and deep borders in the middle of the muscle belly (technique A), as visualized by the hyperechoic fascial lines (Figure 3). However, variations in placement of the transducer and the measurement line could negatively affect reliability when measurements are obtained over time. Therefore, techniques that average multiple measures of muscle thickness have been proposed to improve reliability. This study compared 2 additional techniques to measure muscle thickness. Technique B consisted of an average of 3 measures of muscle thickness across a 2-cm segment of the muscle (Figure 4), whereas technique C represented the mean distance between the inferior and superior fascial borders over the 2-cm segment outlined in technique B.

Table 2. Inter-Rater Reliability Comparing 3 Measurement Techniques for the Lateral Abdominal Muscle (n = 10)

<table>
<thead>
<tr>
<th>Muscle (State)</th>
<th>Mean ± SD, cm</th>
<th>ICC (1, 3) (95% CI)</th>
<th>SEM, cm</th>
<th>MDC, cm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transversus abdominis (rest)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technique A</td>
<td>0.42 ± 0.12</td>
<td>0.92 (0.71–0.98)</td>
<td>0.03</td>
<td>0.09</td>
</tr>
<tr>
<td>Technique B</td>
<td>0.42 ± 0.10</td>
<td>0.90 (0.63–0.98)</td>
<td>0.03</td>
<td>0.09</td>
</tr>
<tr>
<td>Technique C</td>
<td>0.42 ± 0.10</td>
<td>0.83 (0.33–0.96)</td>
<td>0.04</td>
<td>0.11</td>
</tr>
<tr>
<td>Transversus abdominis (ASLR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technique A</td>
<td>0.47 ± 0.13</td>
<td>0.82 (0.29–0.96)</td>
<td>0.06</td>
<td>0.15</td>
</tr>
<tr>
<td>Technique B</td>
<td>0.47 ± 0.11</td>
<td>0.83 (0.35–0.96)</td>
<td>0.05</td>
<td>0.13</td>
</tr>
<tr>
<td>Technique C</td>
<td>0.46 ± 0.10</td>
<td>0.87 (0.46–0.97)</td>
<td>0.04</td>
<td>0.10</td>
</tr>
<tr>
<td>Internal oblique (rest)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technique A</td>
<td>1.00 ± 0.14</td>
<td>0.89 (0.58–0.97)</td>
<td>0.05</td>
<td>0.13</td>
</tr>
<tr>
<td>Technique B</td>
<td>0.99 ± 0.13</td>
<td>0.91 (0.64–0.98)</td>
<td>0.04</td>
<td>0.11</td>
</tr>
<tr>
<td>Technique C</td>
<td>1.01 ± 0.14</td>
<td>0.90 (0.61–0.97)</td>
<td>0.05</td>
<td>0.13</td>
</tr>
<tr>
<td>Internal oblique (ASLR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technique A</td>
<td>1.08 ± 0.19</td>
<td>0.87 (0.48–0.97)</td>
<td>0.07</td>
<td>0.19</td>
</tr>
<tr>
<td>Technique B</td>
<td>1.08 ± 0.19</td>
<td>0.80 (0.16–0.95)</td>
<td>0.08</td>
<td>0.23</td>
</tr>
<tr>
<td>Technique C</td>
<td>1.10 ± 0.20</td>
<td>0.86 (0.43–0.97)</td>
<td>0.08</td>
<td>0.21</td>
</tr>
</tbody>
</table>

ASLR indicates active straight leg raise; CI, confidence interval; ICC, intraclass correlation coefficient; MDC, minimal detectable change; and SEM, standard error of the measurement. Technique A, single measure obtained at the center of the muscle belly; technique B, 3 measures of muscle thickness obtained at the center of the muscle belly and 1 cm medial and lateral to the center; and technique C, average distance between the inferior and superior fascial borders over the 2-cm segment outlined in technique B.

Table 3. Inter-Rater Reliability (n = 21)

<table>
<thead>
<tr>
<th>Muscle and State</th>
<th>Mean ± SD, cm</th>
<th>ICC (1, 3) (95% CI)</th>
<th>SEM</th>
<th>MDC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rectus abdominis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thickness, cm</td>
<td>1.03 ± 0.22</td>
<td>0.91 (0.79–0.96)</td>
<td>0.07</td>
<td>0.22</td>
</tr>
<tr>
<td>CSA, cm²</td>
<td>718 ± 179</td>
<td>0.86 (0.65–0.94)</td>
<td>0.67</td>
<td>1.85</td>
</tr>
<tr>
<td>Transversus abdominis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rest, cm</td>
<td>0.41 ± 0.10</td>
<td>0.86 (0.65–0.94)</td>
<td>0.04</td>
<td>0.10</td>
</tr>
<tr>
<td>ASLR, cm</td>
<td>0.45 ± 0.11</td>
<td>0.87 (0.67–0.95)</td>
<td>0.04</td>
<td>0.11</td>
</tr>
<tr>
<td>Internal oblique</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rest, cm</td>
<td>0.93 ± 0.23</td>
<td>0.91 (0.77–0.96)</td>
<td>0.07</td>
<td>0.19</td>
</tr>
<tr>
<td>ASLR, cm</td>
<td>1.00 ± 0.26</td>
<td>0.93 (0.82–0.97)</td>
<td>0.07</td>
<td>0.19</td>
</tr>
<tr>
<td>Lumbar multifidus</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rest, cm</td>
<td>2.83 ± 0.46</td>
<td>0.87 (0.68–0.95)</td>
<td>0.16</td>
<td>0.46</td>
</tr>
<tr>
<td>Contracted, cm</td>
<td>3.50 ± 0.56</td>
<td>0.94 (0.86–0.98)</td>
<td>0.14</td>
<td>0.38</td>
</tr>
</tbody>
</table>

ASLR indicates active straight leg raise; CI, confidence interval; CSA, cross-sectional area; ICC, intraclass correlation coefficient; MDC, minimal detectable change; and SEM, standard error of the measurement.
the 2-cm segment of the muscle belly (Figure 5). The differences between the measurement techniques were negligible, and no single technique demonstrated a consistent pattern of improved reliability compared to the other techniques. However, the technique requiring the rater to outline the superior and inferior borders of the fascial lines over the 2-cm segment of muscle (technique C) is more time-consuming and therefore less clinically feasible compared to techniques A and B. On the basis of these results, it appears that a single measure of muscle thickness in the center of the muscle belly is equivalent to obtaining the average of 3 measures of muscle thickness over a 2-cm region of the muscle belly. Theoretically, the latter technique in which an average measure of muscle thickness over a 2-cm segment of the muscle is obtained may be more robust for analyzing changes in muscle morphology over time. Further research is warranted to establish whether this is the case.

Several limitations existed within this study. In line with the ultimate goal of the prevention of low back pain in the military trial,37 this study was conducted in healthy adults without low back pain. Despite being conducted on those without abnormalities, the results are similar to those with low back pain.29 However, further research is needed to assess the generalizability of these findings to those with more chronic low back pain and geriatric populations. The inter-rater reliability assessed in this study was conducted within the same session. The influence of larger time intervals on reliability needs to be assessed.32,33 A larger sample may have resulted in smaller confidence intervals around the point estimates of the ICC values. Finally, the reliability calculated in this study was based on an average of 3 measures. This is based on variations in muscle recruitment during submaximal tasks and prior research demonstrating the improved reliability.29–31,34

In conclusion, thickness measures of the transversus abdominis, internal oblique, and lumbar multifidus muscles obtained with ultrasound imaging at rest and during submaximal contractions demonstrate good to excellent reliability. Moreover, the results appear to be generalizable to novice raters who receive a standardized training program before obtaining measurements. Given good to excellent reliability, these results support the potential use of ultrasound imaging measurements for management decisions in the diagnosis and treatment of low back pain. Further research is needed to determine whether the use of ultrasound imaging in the patient care process translates into improved outcomes.

References


