Reliability of Ultrasound Measurements of the Lumbar Multifidus and Transversus Abdominis Muscles during Lying and Unstable Sitting Positions in Individuals With and Without Chronic Low Back Pain


ABSTRACT

Background: The reliability studies are limited to support ultrasound usage during dynamic conditions; for example, unstable sitting position.

Objective: This study aims to examine the reliability of ultrasound measurements of the lumbar multifidus and transversus abdominis during lying and unstable sitting positions in individuals with chronic low back pain (CLBP) and asymptomatic individuals considering abnormal lumbar lordosis.

Material and Methods: In this observational study, intrarater within-day and between-day reliability of muscle thickness and contraction ratio of the lumbar multifidus and transversus abdominis muscles were assessed using ultrasound imaging. In total, 40 participants (27 with CLBP, 13 asymptomatic individuals) with abnormal lumbar lordosis were recruited. The degree of lumbar lordosis has been measured by a flexible ruler. The muscle thickness was assessed at lying and sitting on a gym ball for both muscles in three sessions.

Results: Both groups had well to high ICCs of thickness measurement and contraction ratio in the transversus abdominis and lumbar multifidus muscles during both static (ICC= 0.71-0.99) and semi-dynamic conditions (ICC= 0.73-0.98). The standard error of measurements and minimal detectable changes were rather small in both groups.

Conclusion: Ultrasound imaging is a highly reliable method to assess muscle thicknesses and contraction ratio of the transversus abdominis and lumbar multifidus during different conditions, even in patients with CLBP and abnormal lumbar lordosis.

Keywords
Reproducibility; Diagnostic Imaging; Back Muscles; Lumbar Lordosis; Transversus Abdominis; Low Back Pain

Introduction

Chronic low back pain (CLBP) is one of the leading causes of disability in the world [1]. Approximately 85% of the individuals with this condition experience low back pain (LBP) with no evident anatomical pathology, which is labeled “nonspecific LBP” [2]. The transversus abdominis (TrA) and lumbar multifidus (LM) muscles...
have been proposed to play an important role in spinal stability and shown to have functional deficits in individuals with LBP [3]. These muscles are commonly assessed by ultrasound imaging (US) to determine muscle morphology and function in both research and clinical practice [3-5].

Thickness change and electromyography (EMG) activity of these muscles are linearly correlated at low contraction levels [6]. Accordingly, the US assessment of muscle thickness changes reflects muscle activity at low contraction levels [7]. Also, contraction ratio (CR) defined as contracted thickness/resting of muscle has been suggested as a potential indicator of muscle-tissue status [8]. However, in addition to validity, it is necessary to establish the reliability of US measurements to ensure proper interpretation of results.

Reliability is considered as a psychometric value, indicating the degree that repeated measurements produce comparable results with a decrease in measurement errors. Previous studies have investigated the reliability of US measures of the TrA [9-13] and LM [10-12, 14] in individuals with LBP. However, they have mostly used simple tasks such as abdominal drawing-in maneuver (ADIM) [11, 12], active straight leg raise (ASLR) [11], or contralateral arm lifting (CLAT) [11, 12, 14] to activate the deep trunk muscles preferentially. Moreover, they have shown poor to excellent reliability results for US measurements of the abdominal muscles in people with LBP.

Another task frequently used in clinical practice to facilitate trunk muscle activity is sitting on an unstable surface (e.g., gym ball) that can automatically activate the TrA [15] and LM muscles in individuals with Chronic LBP (CLBP) [4]. Considering lower stability on a gym ball and instantaneously changing of the person’s position compared to stable positions (e.g., CLAT or ASLR), trunk muscle thicknesses may alter in different moments, and the reliability of US measurements may be reduced. The review of the literature showed that there are very few studies that investigated the reliability of US measurements of the LM and TrA muscles in sitting positions with various stability levels [4, 9]. Arab et al., found high within- and between-day reliability for US measurements of the abdominal muscle thickness in individuals with and without LBP in sitting positions with different stability levels [9]. Scott et al., [4] also reported that intraclass correlation coefficients (ICC) scores were greater than 0.9 for repeated measures of the US intrarater reliability for the LM muscle during sitting on a gym ball in individuals with CLBP. However, those studies have investigated the reliability of US measurements of the TrA [9-12, 16] and LM muscles [10-12, 14, 16] in individuals with CLBP without considering lumbar lordosis.

Among the causes of LBP, the alteration of lumbar lordosis plays a significant role in LBP [17]. Alteration of lumbar curvature is associated with various factors, including age, gender, lumbo-pelvic muscles, the orientation of the thoracic spine, and pelvis [18-20]. The abdominal and back musculatures affect pelvic inclination and lumbar lordosis [18]. Individuals with abnormal lumbar lordosis (ABL) commonly show signs of abdominal and back muscle inefficiency, and previous studies have demonstrated an association between the trunk muscle function and LBP [21, 22]. Considering the importance of TrA and LM muscle function in LBP and alteration of lumbar lordosis, the use of gym ball in spine stability exercises may be useful for individuals with CLBP with ABL to improve balance, posture, and the TrA and LM muscle activation [4, 15, 23]. Also, assessment of these muscle thickness and their CR can provide a better understanding of TrA and LM muscle function in individuals with CLBP with ABL during sitting on a gym ball.

To the best of our knowledge, there is no study investigating the reliability of US measurements of the LM and TrA muscles in individuals with CLBP with ABL. Therefore,
Ultrasound Measurement of Trunk Muscles

the current study aimed to investigate collectively intrarater within-day and between-day reliability of US thickness measurements and muscle CR of the LM and TrA muscles in lying and sitting on gym ball in individuals with nonspecific CLBP with ABLL and healthy individuals with ABLL.

Material and Methods

Participants

This observational study had a test-retest reliability design to measure intrarater within-day and between-day reliability in individuals with nonspecific CLBP with ABLL and asymptomatic individuals with ABLL.

A convenience sample of 40 individuals aged 25-55 years old participated in this study. Twenty-seven participants who had a diagnosis of CLBP by an orthopedic specialist were consecutively included. The inclusion criteria for the CLBP group with ABLL were localized back pain between the 12th rib and the gluteal folds lasting more than three months, and their lumbar lordosis angle was more or less than the normal range. In this study, the normal lumbar lordosis was considered ranging from 37º to 42º with a standard deviation of 15°, and outside of this range was identified as ABLL [24]. The lumbar lordosis angle was measured with a flexible ruler. Participants were excluded if they had a history of pain radiating beyond the buttock, sciatica or other radicular involvement, spinal surgery, nerve root compression, neurological deficits, rheumatic diseases, diabetes, pregnancy, lower extremity injuries, neuromuscular diseases or normal lumbar lordosis angle. Also, 13 asymptomatic individuals with ABLL were recruited from the staff and students at the School of Rehabilitation Science as the asymptomatic group with no history of LBP, pain, and dysfunction in the thoracic, pelvis, or lower extremities during the preceding six months. Asymptomatic individuals were excluded if they had a history of LBP during the preceding six months, normal lumbar lordosis angle, and a history of musculoskeletal, cardiopulmonary, or neuromuscular diseases.

The Persian version of the Oswestry Disability Index (ODI) was used to assess disability in participants with CLBP [25] and the Visual Analogue Scale (VAS) to assess pain intensity. In addition, all participants completed the Tegner Activity Rating Scale to assess the activity level [26]. The data were collected from July 2017 to September 2018. The characteristics of the participants are listed in Table 1. All participants received information about the study and signed a consent form before participation. Human Ethics Committee of the Iran University of Medical Sciences approved this

Table 1: Characteristics of the participants in each group.

<table>
<thead>
<tr>
<th>Variable</th>
<th>CLBP (n=27)</th>
<th>Asymptomatic (n=13)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>39.2 ± 9.6</td>
<td>32.1 ± 8.00</td>
<td>0.150</td>
</tr>
<tr>
<td>Gender (female)</td>
<td>10 (37%)</td>
<td>7 (53.8%)</td>
<td>0.314</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>25.6 ± 3.4</td>
<td>24.8 ± 3.6</td>
<td>0.422</td>
</tr>
<tr>
<td>Activity level (Tegner scale)</td>
<td>3.4 ± 0.7</td>
<td>3.1 ± 0.4</td>
<td>0.212</td>
</tr>
<tr>
<td>Lordosis (Degree)</td>
<td>40.8 ± 8.6</td>
<td>34.7 ± 7.3</td>
<td>0.520</td>
</tr>
<tr>
<td>VAS</td>
<td>3.5 ± 1.4</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Disability (ODI)</td>
<td>22.6 ± 12.1</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

CLBP: Chronic Low back pain, BMI: Body mass index, VAS: Visual Analogue Scale, ODI: Oswestry Disability Index. Data are presented as means and standard deviation (Mean ± SD). † Independent sample t-test, ‡ Chi-square test.
study, and all procedures were conducted according to the declaration of Helsinki.

Data acquisition
In this study, a diagnostic US imaging unit set in B-mode (Sonoace R7-Samsung Medison, South Korea) was used by a single examiner to record the images. Following previous studies [9, 11, 15], all the US measurements were performed from the right side of participants by a physical therapist specialized in musculoskeletal disorders with one-year US imaging practice. Before imaging, the examiner demonstrated all asked tasks for the participant.

Ultrasound measurement of the TrA muscle
The TrA thickness measurement was obtained at the end of expiration during 1) supine lying with 60° hip flexion as controlled by a goniometer and a pillow under head and hands resting on the chest (static at rest), 2) supine lying during ASLR test (static at contraction), 3) sitting comfortably on a 65cm diameter gym ball with a straight back, feet on the floor and arms resting on the opposite shoulders (semi-dynamic at rest) and 4) sitting on a 65 cm diameter gym ball and lifting the left foot off the floor (semi-dynamic at contraction) [15].

For the ASLR test, the participant was asked to raise and hold the lower extremity 5cm off the table without bending the knee [13], and the height of 5 cm was marked on the wall. For TrA thickness measurement, a linear 50 mm, 5-7.5 MHz probe was placed halfway along a line joining the anterior superior iliac spine to just below the ribcage in the mid-axillary line. This point appeared to represent best the range of thicknesses of abdominal muscles [27]. To ensure the placement of the probe at the same location, some landmarks such as ASIS, the point below the ribcage, the line between those points, and the halfway point of that line were marked on the skin before initial placement of the probe. Clear images of the muscle thickness were frozen and stored for analysis.

Ultrasound measurement of the LM muscle
The LM thickness measurement was performed at the L4-L5 level during 1) prone lying (static at rest) [7, 28], 2) prone lying during the CLAT (static at contraction), 3) sitting on a 65 cm diameter gym ball with feet on the floor (semi-dynamic at rest) and 4) sitting on a 65 cm diameter gym ball and lifting the left foot off the floor (semi-dynamic at contraction). For CLAT, the participants were instructed to lift the contralateral upper limb approximately 5 cm off the table while her/his upper limbs were repositioned overhead, elbows flexed to 90, and shoulders abducted to 120 as measured using a goniometer. The height of 5 cm was marked on the wall before the start [7]. The spinous process of L5 was marked to guide the placement of the probe before the assessments [14]. After applying the curvilinear 50 mm, 5 MHz probe longitudinally and centrally on the target spinous process, the probe was moved laterally to identify the relevant facet joint. The probe did not move during the testing procedure.

Image Analysis
All US images were processed offline using ImageJ software (Version1.52p; National Institutes. of Health, Bethesda, MD, USA) to calculate the muscle thickness. Linear measurements between the superficial and deep hyperechoic fasciae perpendicular to the muscle fibers in millimeter (mm) were taken for the TrA muscle thickness (Figure 1) [13]. The thickness of the LM muscle was taken as the linear distance between the tip of the targeted zygaphophyseal joint to the inside edge of the superior border of the multifidus muscle (Figure 1) [14]. The cursor points carefully measured the TrA and LM muscle thicknesses. A vertical straight line through the center of the US image was used to ensure the standardized placement of the measurement line [16]. The contraction thickness ratio of the TrA and LM muscles were also calculated as contracted.
Ultrasound Measurement of Trunk Muscles

The lumbar lordosis angle was measured using a flexible ruler as a noninvasive method. Two bony landmarks of the spinous process of L1 and S2 were marked on the skin [30]. Then, the flexible ruler was fixed on the lumbar region over the mentioned landmarks until it shaped lumbar lordosis curvature while the participant stood comfortably on both feet in normal position and fixed his/her eyes on the opposite wall. Then, without changing the shape, the ruler was carefully placed on a white paper, and the curve with the L1 and S1 markers was transferred to the paper. The L1 spot to the S2 spot was connected by a straight line on the paper and was drawn a line perpendicular to its center, passing the curve. These lines were named L and H, respectively. Lumbar lordosis can be calculated by replacing the lengths of these lines in the following equation: \( \theta = 4\text{ARC tag}[2H/L] \).

The lumbar lordosis angle was measured three times, and the average was calculated for further analysis. Based on previous studies, the lumbar lordosis angle is relatively low in sitting posture compared to standing [31, 32]. Thus, the lumbar lordosis angle was measured in the standing posture in the current study before the US measurements. High intra-rater reliability and validity for lumbar lordosis angle measurements have been reported in the standing posture using a flexible ruler [24, 30, 33].

**Procedure**

All experimental conditions were performed in a biomechanics laboratory, and all participants (with and without CLBP) were evaluated.
on three separate sessions with the same procedure. First, the examiner recorded all measurements and repeated the measurements after 1 h to calculate within-day reliability. After three days, the third session was completed for between-day reliability. The testing conditions and muscles were randomly selected to avoid order effects. The overview of participant flow and data collection are shown in Figure 2.

**Statistical Analysis**

Data were presented as means ± standard deviations (SDs) and analyzed using SPSS version 22.0 (IBM Corporation, Chicago, IL, USA). The normality of distribution for the demographic data and US measurements was assessed using the one-sample Kolmogorov-Smirnov test. The differences in quantitative and qualitative demographic data were assessed using independent sample t-test and chi-square tests, respectively. ICC (ICC$_{3,1}$; method: alpha, two-way mixed, consistency) was used to calculate intrarater within-day and between-day reliability for the average of three thickness measurements and thickness CR of the TrA and LM muscles. Model 3 (ICC$_{3,1}$) was employed because only one rater assessed all participants. The ICCs are classified as follow: <0.69, poor reliability; 0.70–0.79, fair reliability; 0.80–0.89 good reliability; 0.90–1.00 high reliability [34]. Standard error of measurement (SEM) [SEM = pooled SD $\sqrt{1 – ICC}$] and minimal detectable change (MDC) [MDC = $SEM \times z \times \sqrt{2}$] for a 95% confidence interval were also calculated. SEM value indicates the error of the instrument itself, i.e., the precision of the measurement, and MDC value reflects the smallest change in a score within an individual, which can be considered as a real change above measurement error with $p<0.05$. The statistical level of significance was set at 0.05.

**Results**

Normal distribution was observed for all variables. Independent sample t-test showed

---

**Figure 2:** Overview of participants and flow chart of data collection.
non-significant differences in the BMI or activity level between the CLBP and asymptomatic groups, but the LBP group was older and had higher lordosis than the asymptomatic group. Participants with CLBP also had mild to moderate pain intensity and disability levels (Table 1).

**Thickness measurements**

Table 2 presents US measurements of the TrA and LM muscles (Mean ± SD) in mm in both groups during different testing postures. The asymptomatic group generally had larger thicknesses of the TrA and LM muscles than the CLBP group. Moreover, the muscles showed higher thicknesses during contraction in the semi-dynamic posture (Table 2).

**Test-retest reliability**

Table 3 presents the ICC, SEM, and MDC values for within-day and between-day reliability of the US thickness measurements of the TrA and LM muscles during different testing postures in each group. As shown in Table 3, the US measurements were well to high for both within-day and between-day comparisons in the CLBP and asymptomatic groups. Overall, we had greater SEMs in the CLBP compared to the asymptomatic group, and the SEM values of the CLBP group were higher in the semi-dynamic posture than the static posture. Likewise, there were higher MDCs in the CLBP group than the asymptomatic individuals, and the MDC values of the CLBP group were higher in the semi-dynamic posture than the static posture (Table 3).

**Discussion**

In the current study, we aimed to collectively evaluate interrater within- and between-day reliability of US thickness measurements for the TrA and LM muscles in various lying and sitting positions in individuals with CLBP and ABLL and asymptomatic individuals with...
The findings showed well to high reliability US measurements for both muscle thickness and CR in both groups. These results confirmed previous findings of utilizing US to measure the thickness and CR of the TrA and LM muscles reliably [9, 11]. However, in those studies, the reliability of US thickness measurements was assessed in individuals with and without CLBP, without considering the lumbar lordosis angle as an influencing factor. Despite the effect of the TrA and LM muscle function on the lumbar lordosis angle, the intrarater within-day and between-day reliability in individuals with ABLL were high, and it was comparable with previous reliability studies that have evaluated individuals without considering ABLL. Therefore, US imaging is a reliable tool for evaluation in these patients or healthy people.

Some studies have evaluated the reliability of US for the TrA and LM muscle thickness in individuals with and without LBP during a simple task [10, 12, 35]. Although gym ball is commonly used to improve proprioception and to trigger trunk muscle activity in spinal rehabilitation programs, very few studies have assessed the reliability of these muscles during sitting on a gym ball. Moreover, to our knowledge, no studies have evaluated the reliability of CR of the TrA and LM muscles during sitting on a gym ball in individuals with and without CLBP.

Arab et al., (2013) found high ICC (0.85-0.95) and low SEM (0.19-0.78 mm) and MDC (0.52-2.15 mm) for the TrA muscle during lying and sitting on a gym ball in individuals with and without CLBP [9]. Similarly, small SEM and MDC scores were observed in the current study, but there were higher SEMs and MDCs in the CLBP group compared to the as-

Table 3: Within-day and between-day reliability of the ultrasound thickness measurements of the transversus abdominis (TrA) and lumbar multifidus (LM) muscles during different testing conditions in each group.

<table>
<thead>
<tr>
<th>Position</th>
<th>Muscle</th>
<th>Task</th>
<th>CLBP</th>
<th>Control</th>
<th>CLBP</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Within-day</td>
<td>Between-days</td>
<td>Within-day</td>
<td>Between-days</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ICC</td>
<td>SEM</td>
<td>MDC</td>
<td>ICC</td>
<td>SEM</td>
</tr>
<tr>
<td>Static</td>
<td>TrA</td>
<td>Rest</td>
<td>0.94</td>
<td>0.40</td>
<td>1.13</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Con</td>
<td>0.89</td>
<td>0.56</td>
<td>1.55</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CR</td>
<td>0.95</td>
<td>0.09</td>
<td>0.25</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>LM</td>
<td>Rest</td>
<td>0.93</td>
<td>1.70</td>
<td>4.72</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Con</td>
<td>0.92</td>
<td>1.70</td>
<td>4.72</td>
<td>0.83</td>
</tr>
<tr>
<td>Dynamic</td>
<td>TrA</td>
<td>Rest</td>
<td>0.95</td>
<td>0.04</td>
<td>0.12</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Con</td>
<td>0.91</td>
<td>0.57</td>
<td>1.59</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CR</td>
<td>0.93</td>
<td>0.11</td>
<td>0.31</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>LM</td>
<td>Rest</td>
<td>0.93</td>
<td>1.40</td>
<td>3.87</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Con</td>
<td>0.91</td>
<td>2.54</td>
<td>7.03</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CR</td>
<td>0.97</td>
<td>0.05</td>
<td>0.14</td>
<td>0.97</td>
</tr>
</tbody>
</table>

CLBP: Chronic Low back pain; TrA: Transversus abdominis; LM: Lumbar Multifidus; Con: Contraction; CR: contraction ratio; ICC: Intraclass correlation coefficient; MDC: Minimal detectable change; SEM: Standard error of measurement. All SEM and MDC values are in millimeters.
ymptomatic participants. In the CLBP group, the SEM and MDC values were higher during sitting on gym ball than the lying positions. Based on the SEM and MDC formula, a higher SD causes larger SEM and MDC. There are several possible explanations for higher SD and subsequent larger SEM and MDC in our findings. The large variability of the muscle thicknesses (observed in the current study) may result from multiple modulation strategies of the measured muscles that could increase the SDs. Besides, participants without CLBP may keep harder positions (e.g., sitting on a gym ball) with fewer displacements and less variability than individuals with CLBP. Also, the levels of contraction and thickness changes are different in each posture so that they may affect variability, SD, SEM, and MDC scores.

Based on the lower SEM and MDC scores of CR for the TrA and LM muscles compared to the resting and contracted thicknesses in the current study, it seems that the CR measures (contracted thickness/resting thickness) may be more useful than a single thickness measurement for monitoring of trunk muscle dysfunction in individuals with CLBP with ABLL in clinical practice and research. The reliability of the US in this context is rather scarce. Sarafraz et al., (2018) suggested using the US as a highly reliable method for measuring CR of LM at the L5 vertebral level in patients with CLBP with unilateral radiculopathy and healthy controls [36]. In total, both groups had higher SEMs for CR of TrA than the ones for LM. These differences may indicate greater changes in muscle thickness over time and/or muscle recruitment patterns of the TrA muscle than the LM muscle. Also, myofascial extensibility, intra-abdominal pressure, and forces from surrounding muscles may introduce variability that might affect TrA muscle thickness but not the LM muscle.

Good to high within-day and between-day reliability of US measurements in this study, even in the unstable sitting postures, may be related to several reasons. The US probe was not displaced as much as possible while recording muscle thickness, clear images were recorded, and cursor points precisely measured muscle thickness. Lower between-day reliability than within-day reliability in the CLBP group may be due to the impairment of deep trunk muscles or inconsistent motor performance due to pain [14] while in the asymptomatic group, within-day and between-day reliability were high.

The current study has several limitations. This study was not assessed interrater reliability. Moreover, muscle thickness and CR of other trunk muscles such as the oblique muscles or other LM levels were not evaluated. The US examiner was experienced, which could have resulted in high reliability.

**Conclusion**

The US thickness and CR measurements of the TrA and LM muscles have well to high intrarater within-day and between-day reliability as well as small SEMs and MDCs in static and semi-dynamic postures in individuals with and without CLBP with ABLL. The CR measures may be adequately reliable and helpful to assess the function of the TrA and LM muscles in asymptomatic and individuals with CLBP and ABLL. Therefore, this study suggests that real-time US imaging can be reliably used to assess the thickness and activity of the TrA and LM muscles in individuals with and without CLBP with ABLL.

**Acknowledgment**

This study was funded by Iran University of Medical Sciences. We thank Dr. Hedayat Abbastabar for Statistical Consulting.

**Conflict of Interest**

None

**References**

1. Hurwitz EL, Randhawa K, Yu H, Cote P, Halldeman S. The Global Spine Care Initiative: a summary of the global burden of low back and
Bozorgmehr A. et al


Ultrasound Measurement of Trunk Muscles


