Ultrasound Transducer Shape has no effect on Measurements of Lumbar Multifidus Muscle Size

Peter R Worsley* (PhD), Nicholas. Smith(BSc), Martin B Warner(MSc), Maria. Stokes(PhD)

.Faculty of Health Sciences, University of Southampton, UK, SO17 1BJ

*Corresponding author:

Peter Worsley MCSP PhD

Faculty of Health Sciences, Building 45

Highfield Campus, University of Southampton, Southampton

UK, SO17 1BJ

Tel. +44(0)2380 598957.

Email: P.R.Worsley@soton.ac.uk

Abstract

Objective: Evidence is currently lacking for guidance on ultrasound transducer configuration (shape) when imaging muscle to measure its size. This study investigated measurements of lumbar multifidus with curvilinear and linear ultrasound transducers.

Method: Fifteen asymptomatic, males (aged 21-32 years) had their right lumbar multifidus imaged at L3 by an experienced ultrasound imager. Two transverse images were taken with two transducers (5MHz curvilinear and 6MHz linear) and linear/CSA measurements were made off-line. Reliability of image interpretation between days was assessed using intra-class correlation coefficients (ICC). Muscle measurements using the two transducers were compared using Bland and Altman plots and paired t-tests. Relationships between CSA and linear measurements were examined using Pearson's Correlation Coefficients.

Results: There were no significant differences ($p>0.05$) in the measurements of the two transducers. Thickness and CSA measurements had small differences between transducers, with mean differences of 0.01cm (SDdiff=0.21cm) and 0.03cm^2 (SDdiff=0.58cm^2) respectively. Width measures had a mean difference of 0.14cm, with the linear transducer giving larger measures. Significant correlations (p <0.001) were found between all linear measures and CSA, with both transducers (r=0.78-0.89). Interpretation of images was generally reliable between days, with thickness measures being the most repeatable (ICC>0.93) and CSA the least repeatable using the linear transducer (ICC=0.78).

Conclusion: Measurements of multifidus at L3 were not influenced by the configuration of transducers of similar frequency. For the purposes of image interpretation, the curvilinear transducer gave better reflection of the lateral muscle border than the linear transducer, suggesting it as the preferable for imaging lumbar multifidus.

1. INTRODUCTION

In recent years ultrasound technology has developed to a stage where it has found widespread application, both in diagnostic use and in rehabilitation (Whittaker et al., 2007) . Rehabilitative ultrasound imaging (RUSI) can offer a safe, objective and relatively inexpensive means of examining the muscles and there is a growing body of evidence supporting the use of RUSI in physiotherapy research and practice (Whittaker et al., 2007). The two areas in which the modality is commonly applied are to evaluate muscle and related soft tissue morphology and to provide visual feedback to aid interventions in rehabilitation.   The RUSI technique has undergone extensive reliability studies, most of which has shown high reliability for both taking and interpreting the ultrasound images (Kiesel et al., 2007; Koppenhaver et al., 2009; Wallwork et al., 2007). The validity of RUSI against the gold standard of magnetic resonance imaging (MRI) has also been established for several muscles, such as lumbar multifidus cross-sectional area (Hides et al., 1995), abdominal muscle thickness (Hides et al., 2006), cervical multifidus muscle thickness (Lee et al., 2006), infraspinatus muscle thickness (Juul-Kristensen et al., 2000), supraspinatus muscle thickness and CSA (Juul-Kristensen et al., 2000, and lower trapezius (O'Sullivan et al., 2009).   An area that has received relatively little attention is determining appropriate scanner specifications for imaging muscles and the present study addresses the topic of transducer shape (or configuration).

RUSI has been used to study lumbar multifidus in several studies, with clinically important findings, principally muscle atrophy, in lower back pain patients (Hides et al., 2008; Wallwork et al., 2009). The type of transducer used when imaging lumbar multifidus has varied between studies, with some using a curvilinear transducer (Hides et al., 1992; Kiesel et al., 2007; Lee et al., 2006; Stokes et al., 2005) and others using a linear transducer (e.g. Hides et al., 1995), without justification of transducer choice. Some guidance based on anecdotal evidence in a clinical commentary paper suggested that more sound waves would reach the lateral border of multifidus with a curvilinear transducer than a linear transducer (Stokes et al., 2007). The frequency of the transducer can be justified by the depth at which a muscle lies; as a general

rule, the highest frequency transducer that can image an area of interest should be used (Whittaker et al., 2007). As highlighted in a recent review by Stokes et al (2007), for lumbar multifidus a 5MHz is often used (Stokes et al., 2007).

Transducer selection (curved or linear) is often carried out by the sonographer in order to obtain the best clarity of images. However differing shapes of transducer determine the nature and direction of the sounds waves and the subsequent attenuation and reflection from the underlying tissues. Soft and hard tissue borders of interest must be perpendicular to the transducer array in order to obtain the best reflection and therefore clarity of image, so the underlying shape of the muscle has a bearing on choice of transducer. It is unknown, however, if the shape of the transducer affects the measurements of muscle dimensions. In a recent study, phantom objects were imaged using both curvilinear and linear transducers (Warner et al., 2008). Small differences were found between linear and curvilinear array transducer configurations when measuring simulated muscle cross-section dimensions using a semi-solid phantom (Warner et al., 2008). The results, however, showed no clinically significant difference between the transducers (as judged against muscle data in the literature) when assessed with Bland and Altman plots (Warner et al., 2008).

The purpose of the present study was to build on the work with phantom images (Warner et al., 2008) and apply it to simulate a clinical situation by studying a commonly imaged muscle, the lumbar multifidus. This aim was to determine whether there was a difference between measurements made from images obtained using linear and curvilinear transducers. The null hypothesis was that there would be no difference between measurements of lumbar multifidus dimensions made using linear and curvilinear ultrasound transducers.

2. METHODS

2.1 Participants

Fifteen males (aged 21-32 years; mean height 180cm, weight 81.7 kg and body mass index, BMI, 25.5) were recruited from staff and students at the University of Southampton. Exclusion

criteria were: extended history of severe lower back pain, any wound or skin condition in the region to be scanned, any previous spinal or pelvic fractures that may interfere with positioning or any known spinal abnormality (such as scoliosis). Ethical approval for the project was obtained from the School of Health Professions & Rehabilitation Sciences Ethics Committee and all participants gave their written informed consent prior to participation.

2.2 Ultrasound Imaging Technique

Participants lay prone on a plinth, with a pillow under the pelvis to reduce the lumbar lordosis. The third lumbar spinous process (L3) was identified by palpation, using the posterior superior iliac crests as a reference point for the fifth lumbar spinous process (L5). L3 could then be identified by palpating in a cranial direction and was marked with a water-soluble chinograph pencil. Although L4 has been studied widely and may be more relevant in back pain, the entire cross section of the muscle could not be seen in the present participants, possibly because their muscles were hypertrophied due to regular sporting activity. Therefore, L3 was studied. An Esaote Pie Medical 'Aquila' ultrasound scanner (ESAOTE S.p.A. Genova, Italy), with a 6 MHz linear transducer array (60mm footprint) and 5 MHz curvilinear transducer array (80mm footprint), was used to obtain images. The imaging procedure was undertaken by an experienced operator (MS) and based on that described by Stokes et al (2005), where the transducer was first placed longitudinally over the mid-line of the lumbar spine at the allocated landmark over L3 to aid orientation (Figure 1a). The transducer was then rotated 90 degrees to produce a transverse image of the bilateral multifidus muscles (Figure 1b).

[insert FIGURE 1 near here]

The transducer was then moved to the left of the spinous process over multifidus and perpendicular to the lamina, and then angled to find the sharpest image of the muscle's borders (Figure 2). Changes in transducer angle were kept as small as possible and it has been shown

that these should be below 10 degrees to avoid altering the muscle's dimensions (Whittaker et al., 2009). Contraction of multifidus was induced by the participant lifting the ipsilateral leg from the plinth to confirm identification of the muscle borders. When multifidus was at rest and clearly defined, an image was taken and stored on a memory card using an assigned coding system for later analysis by another investigator (PW). The process of obtaining an image was repeated with the second ultrasound transducer and the order of transducers was randomised between participants.

Measurements were taken from the images off-line, using Image J software on a computer screen (http://rsbweb.nih.gov/ij/). Thickness of lumbar multifidus was defined as the maximal distance from the lamina to the superficial fascia of the muscle belly. Width of the muscle was defined as being from the lateral edge of the spinous process to the inside of the lateral fascial border, with the measurement taken at the mid-point of the thickness measurement (half way up the muscle belly). CSA of the muscle belly was measured by tracing the inside border of the muscle fascia and the outer line of the spinous process. Interpretation of the images was completed by the same investigator (PW) twice on two separate days for the same images. The mean of the measures taken from the first session were then used for the final analysis. The measures taken on both days were used to assess the between day reliability of the measurement process.

[Insert FIGURE 2 near here]

2.3 Analysis of Data

Reliability of image interpretation between days for one rater (PW) was assessed using intra-class correlation coefficients ICC(1,2) (Shrout et al., 1979). Muscle measurements using the two transducers were compared using Bland and Altman plots and paired t-tests. Pearson's correlation coefficients were calculated to examine the relationship between the linear measurements and CSA for both the linear and curvilinear transducers, to see if transducer type

had an effect on this relationship. The linear measurements were squared for this analysis to give them comparable dimensions to CSA.

3. RESULTS

3.1 Reliability of Image Interpretation

Prior to the main investigation the intra-rater reliability of linear and CSA measurements of image interpretation was assessed. Reliability of between day repeated measurements made on the same images were good to excellent, with ICC(1,2) ranging from 0.78-0.99 Table 1).

[Insert TABLE 1 near here]

Linear measurements of thickness and width showed the highest reliability in both transducers with ICC(1,2) of >0.93 and >0.88 for thickness and width respectively. The curved transducer also showed high reliability for CSA (ICC=0.89) but reliability using the linear transducer was lower (ICC=0.78).

3.2 Comparison of Measurements from the Linear and Curvilinear Transducers

The results from the Bland and Altman plots and T-test analysis showed that there were no significant differences between the transducer measurements (Table 2).

[Insert TABLE 2 near here]

The Bland and Altman plots (Figure 3) showed minimal between-transducer differences. Confidence intervals for the mean difference ($\overline{d}$) between transducers were close to zero (Table 2). For thickness (Fig 3a), $\overline{d}$ of 0.01cm showed a clear similarity between transducers. The width

measurement (Figure 3b) showed the largest mean difference between the transducers (0.14cm±0.3cm), indicating a bias towards a larger measurement with the linear transducer. This highlighted the potential for error between transducers but was not large enough to result in a significant difference on the paired t-tests (Table 2). CSA measures between the transducers showed a small mean difference (0.03cm^2), but the standard deviation in difference was large (0.6cm^2) (Table 2).

[Insert FIGURE 3 near here]

The shape ratio of the linear dimensions (thickness/width) were similar for both transducers, with a linear transducer ratio of 1.03 (SD=0.13) and curvilinear transducer ratio of 1.09 (SD=0.11).

3.3 Correlation between Linear and Cross-Sectional Area Measurements

The correlation between the linear dimensions and muscle CSA were all statistically significant for each transducer ($p<0.05$). The strength of the relationships (r) varied, with a similar order of strength for both transducers; thickness^2 being weakest, then width^2, and thickness x width being strongest (Table 3). The correlation coefficient (r) values were identical for the two transducers, except for thickness*width, which was higher for the curvilinear (r =0.89) than the linear (r =0.85) transducer.

[Insert TABLE 3 near here]

4. DISCUSSION

The main finding indicates that measurements of all three dimensions of multifidus muscle size (thickness, width and CSA) at L3 were not influenced by whether a linear or curvilinear transducer was used, therefore confirming the null hypothesis that measurements would not differ. Intra-rater reliability of measurements on the same images on different days was good to excellent. Linear measures of the muscle were highly correlated with CSA but thickness, which is most commonly used in the literature , had the lowest correlation.

Both linear measurements (thickness and width) were similar between the transducers, with low mean differences which were not significantly different (Table 3). However, the slight bias observed for width measurement in the Bland and Altman analysis, with the linear transducer showing larger measures, may have been influenced by difficulty in defining the lateral muscle border. Studies in the literature report multifidus thickness more commonly than width or CSA (Hides et al., 2008; Wallwork et al., 2009) and thickness is more rapid and easy to use in clinical practice. The mean differences in muscle thickness between the two transducers (0.01cm) was well within reported values for minimally detectable change from between-day reliability studies, e.g. 0.07 cm (Wallwork et al., 2007). This provides evidence that multifidus thickness measurements are valid using either type of transducer. Multifidus is commonly imaged in the parasaggital plane with the transducer in a longitudinal orientation (Kiesel et al., 2007). Although the present findings on transverse images were not confirmed for longitudinal images, it is unlikely that thickness measurements would differ by simply rotating the transducer by 90^{0}.

The present findings are consistent with the previous research conducted on a semi-solid phantom object containing wires at known distances apart (Warner et al., 2008). In both studies, high levels of repeatability for image interpretation and small between transducer differences were observed. Warner et al. (2008) found mean differences between linear and curvilinear transducers of 0.02cm (SD=0.04cm) for thickness and -0.09-0.1cm (SD=0.03cm) for width. They also found larger differences in the CSA estimation between linear and curvilinear

transducers (-0.33-0.24cm^2). These errors are larger than those found in the present study, although this reflects the fact that Warner et al. (2008) were measuring much larger objects (13-18.86cm^2). Only one previous study has compared linear and curvilinear ultrasound transducers on muscle thickness and found no significant difference for the transversus abdominis (TrA) muscle with the two transducers (McMeeken et al., 2004). The TrA is relatively thin compared with multifidus (approximately 0.5cm compared with 2.7cm), so there was less scope for finding differences. However, the present study also found no difference with thickness measurements. Measurements of CSA or width had therefore not been studied previously.

Current evidence of the association of pathology and muscle atrophy using ultrasound imaging have found significant differences in lumbar multifidus between lower back pain (LBP) patients and healthy controls (Hides et al., 2008; Wallwork et al., 2009). Hides et al. (2008) showed large differences in CSA at the fourth and fifth vertebral levels (mean differences 1.35-3.26cm^2), of which many of these were significant. These differences are well outside the range of uncertainty that the present study has shown between the transducers.

The findings from the present study support the physical principles behind the two different ultrasound transducers. On some of the images it was difficult to determine the borders for the CSA, particularly with the linear transducer. The anterior (superficial) and posterior borders are well defined by the fascia and echogenic lamina respectively. Figure 2 illustrates how difficult it can be to identify the lateral border, particularly with the linear transducer. The interpretation of this border will have an effect on the width measurement but more significantly on the CSA, and misinterpretation will therefore contribute to discrepancies. The curvilinear transducer has the advantage of exposing the lateral border to more sound waves and thus better attenuation, than the linear transducer. This fact was reflected in the stronger relationship between linear and CSA measurements with the curvilinear transducer and the lower reliability of CSA interpretation using the linear transducer.   Another consideration in the relationship between the muscle dimensions is the irregular shape of the

multifidus in the present group, possibly due to hypertrophy, compared with sedentary controls, who are reported to have round or slightly ovoid muscles (Stokes et al, 2005). The present group tended to have a trapezoid or almost triangular multifidus shape, which would give poorer correlation values between linear and CSA measures than for more regularly shaped muscles (Stokes et al., 2005). The habitual activity levels of the present subpopulation could therefore account for both the greater size and different shape of the multifidus muscle to previously studies sedentary populations (Stokes et al., 2007). In a review of posterior paraspinal ultrasound imaging, it was highlighted that multifidus at L3 was measured to be 4.9cm^2 (Hides et al., 1995; Stokes et al., 2007). However, the CSA of lumbar multifidus in the present study was much larger (6.89cm^2), reflecting the young athletic males included in the study.

Multifidus thickness appears to be a useful measure for assessing contractile ability during voluntary contractions or automatic contraction during functional tasks (Kiesel et al., 2007). Thickness may not be adequate for evaluating the effects of interventions for increasing multifidus muscle size, particularly when the shape becomes irregular, as thickness does not reflect CSA strongly enough. A curvilinear transducer would therefore be recommended for such studies, despite not having a significant effect on measurements, because it enables the lateral border to be seen more clearly, making interpretation of the image easier.

4.1 Limitations of the study

The study lacked external validity against a gold standard for muscle size, such as magnetic resonance imaging, but this has been established previously for multifidus by Hides et al (1995). The study on a phantom, however, used wires at known distances (Warner et al., 2008) and similar bias towards more accurate measures with the curvilinear transducer were found. Reliability testing was restricted to image interpretation and did not include reliability of the technique itself (within or between days), since the person performing imaging was experienced. There would still be a degree of error involved in using the different transducers

but the similarity of findings with those from the study of a phantom (Warner et al., 2008) suggests that the differences could, at least partly, be explained by the transducers. Another limitation was the time lapse from image collection to interpretation. Real-time imaging enables dynamic assessment of muscle borders by asking the subject to contract the muscles, e.g. by lifting the ipsilateral leg, Movements of the fascicles in multifidus have a swirling appearance, which does not occur in longissimus, making it easier to define the border between them. Immediate measurements after these observations may help make measurements more accurate than later interpretation of a still image. In addition to image interpretation for making measurements, the imaging technique itself can introduce errors. It was reported that a 30 degree angle variance in the transducer head relative to the skin could result in an error of 15% in non-tapering muscle (Dupont et al., 2001). More recently, Whittaker et al. (2009) established that changes in transducer angle below 10 degrees did not result in significant changes in muscle dimensions. Finally, with the small sample of sporting males of a similar body type, the present findings cannot be generalised to sedentary healthy subjects or patients with smaller muscles.

4.2 Recommendations

On the basis of the present study showing that the curvilinear transducer produced the clearest images of the lateral border and more repeatable measures of CSA , it is therefore recommended for imaging lumbar multifidus. Choice of transducer may not always be an option if only one is available. Accurate muscle thickness measurements can be made with a linear transducer but width and CSA may be less accurate and less reliable for monitoring changes in multifidus over time. Certainly when purchasing a scanner, if only one transducer is affordable, then a curvilinear transducer is recommended.

5. CONCLUSIONS

Measurements of multifidus width, thickness and CSA made using linear and curvilinear transducers were not significantly different. High levels of between day intra-rater reliability were shown for image interpretation using both curved and linear transducers. There were also no differences between transducers regarding correlation between linear measures and CSA. However, the curvilinear transducer gave better clarity of the lateral border of the muscle and is therefore recommended as the transducer of choice for ease if interpreting images of the lumbar multifidus. This study highlights the need for research to determine appropriate scanner specifications for RUSI in different muscles and study populations.

References

Dupont, A-C, Sauerbrei, EE, Fenton, PV, Shragge, PC, Loeb, GE, Richmond, FJR. Real-time sonography to estimate muscle thickness: Comparison with MRI and CT, 2001; pp. 230-236.

Hides, J, Gilmore, C, Stanton, W, Bohlscheid, E. Multifidus size and symmetry among chronic LBP and healthy asymptomatic subjects. Manual Therapy 2008; 13(1): 43-49.

Hides, JA, Cooper, DH, Stokes, MJ. Diagnostic Ultrasound Imaging for Measurement of the Lumbar Multifidus Muscle in Normal Young Adults. Physiotherapy Theory and Practice 1992; 8(1): 19 - 26.

Hides, JA, Richardson, C, Jull, G. Magnetic Resonance Imaging and Ultrasonography of the Lumbar Multifidus Muscle: Comparison of Two Different Modalities. Spine 1995; 20(1): 54-58.

Kiesel, KB, Uhl, TL, Underwood, FB, Rodd, DW, Nitz, AJ. Measurement of lumbar multifidus muscle contraction with rehabilitative ultrasound imaging. Manual Therapy 2007; 12(2): 161-166.

Koppenhaver, SL, Hebert, JJ, Fritz, JM, Parent, EC, Teyhen, DS, Magel, JS. Reliability of Rehabilitative Ultrasound Imaging of the Transversus Abdominis and Lumbar Multifidus Muscles. Archives of Physical Medicine and Rehabilitation 2009; 90(1): 87-94.

Lee, S, Chan, C, Lam, T, Lam, C, Lau, N, Lau, R, Chan, S. Relationship Between Low Back Pain and Lumbar Multifidus Size at Different Postures. Spine 2006; 31(19): 2258.

McMeeken, JM, Beith, ID, Newham, DJ, Milligan, P, Critchley, DJ. The relationship between EMG and change in thickness of transversus abdominis. Clinical Biomechanics 2004; 19(4): 337-342.

Shrout, P, Fleiss, J. Intraclass correlations: Uses in assessing rater reliability. Psychological Bulletin 1979; 86(2): 420-428.

Stokes, M, Hides, J, Elliott, J, Kiesel, K, Hodges, P. Rehabilitative ultrasound imaging of the posterior paraspinal muscles. Journal of Orthopaedic & Sports Physical Therapy 2007; 37: 581-595.

Stokes, M, Rankin, G, Newham, DJ. Ultrasound imaging of lumbar multifidus muscle: normal reference ranges for measurements and practical guidance on the technique. Manual Therapy 2005; 10(2): 116-126.

Wallwork, TL, Hides, JA, Stanton, WR. Intrarater and interrater reliability of assessment of lumbar multifidus muscle thickness using rehabilitative ultrasound imaging. Journal of Orthopaedic & Sports Physical Therapy 2007; 37: 608-612.

Wallwork, TL, Stanton, WR, Freke, M, Hides, JA. The effect of chronic low back pain on size and contraction of the lumbar multifidus muscle. Manual Therapy 2009; 14(5): 496-500.

Warner, MB, Cotton, AM, Stokes, MJ. Comparison of curvilinear and linear ultrasound imaging probes for measuring cross-sectional area and linear dimensions. Journal of Medical Engineering & Technology 2008; 32(6): 498 - 504.

Whittaker, VJ, Teyhen, D, Elliot, J, Cook, K, Langevin, H, Dahl, H, Stokes, M. Rehabilitative Ultrasound Imaging: Understanding the Technology and Its Applications. Journal of Orthopaedic and Sports Physical Therapy 2007; 37(8): 434-449.

FIGURE LEGENDS

Figure 1

Ultrasound images taken with a 6MHz linear transducer.   **(a)** Sagittal image along the midline of the spine showing L3/L4 spinous processes (SP). **(b)** Bilateral transverse image at L3 showing the multifidus muscle (M), SP and lamina (L)

Figure 2

Ultrasound scans at L3 taken with different transducers and showing the multifidus muscle (M), spinous process (SP), lamina (L), and the thickness (antero-posterior) measurement indicated by a white line. **(a)** Linear array image **(b)** Curvilinear array image).

Figure 3

Bland and Altman plot for measurements of multifidus at L3 using the linear and curvilinear transducers: **(a)** thickness **(b)** width (**c**) cross-sectional area (CSA). Mean difference (solid line) and 95% limits of agreement (dashed lines).
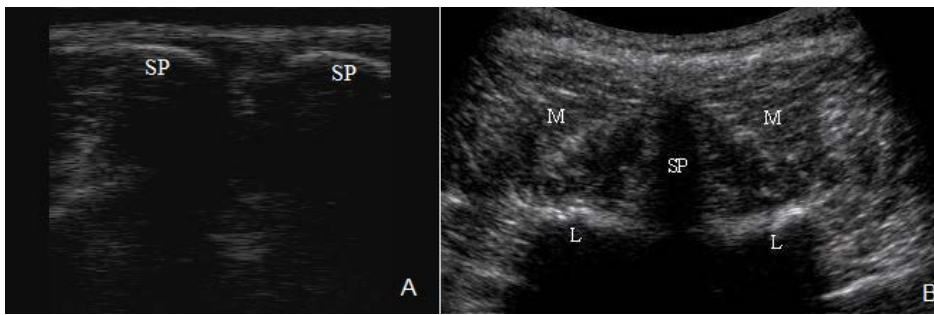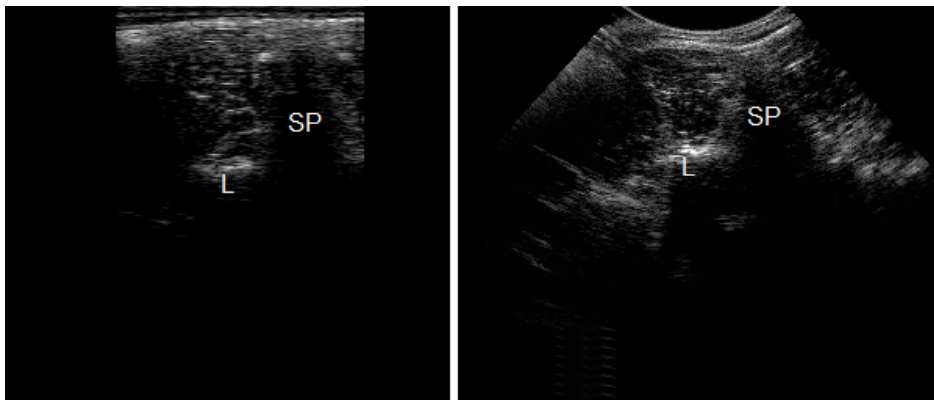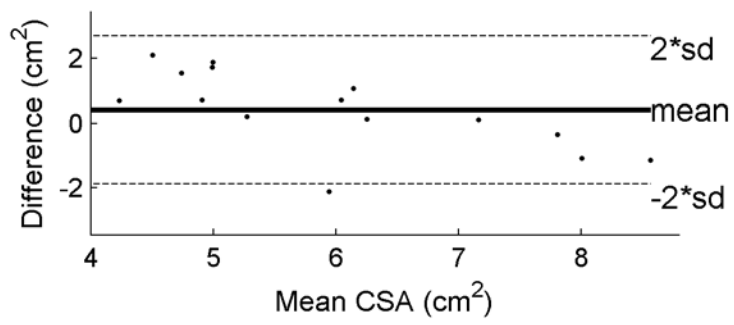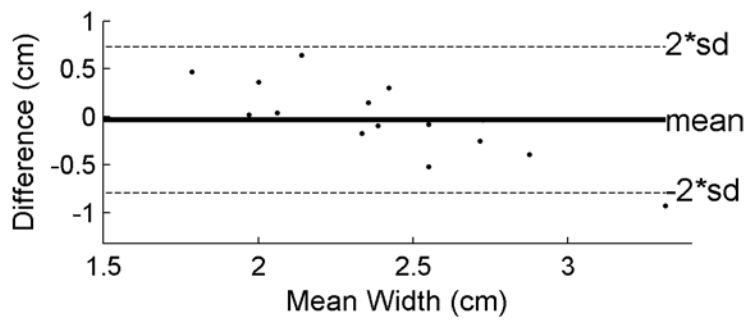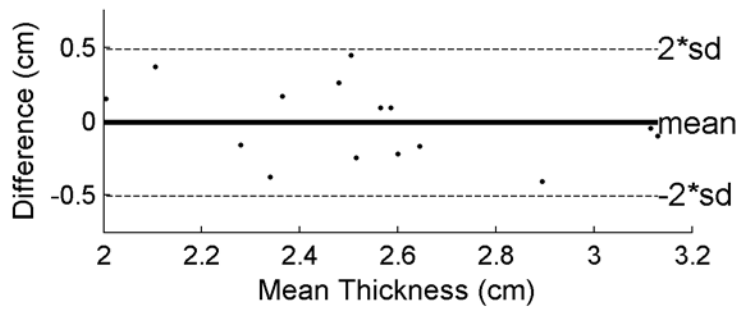
Figure 1

Figure 2

Figure 3

Table 1. Intra-class correlation coefficient (ICC) analysis of between day interpretation of the same images taken with curvilinear and linear transducers

		ICC(1,2)	95% Confidence Interval	
			Lower Bound	Upper Bound
Curvilinear	Thickness	0.93	0.80	0.98
	Width	0.88	0.64	0.96
	CSA	0.89	0.68	0.96
Linear	Thickness	0.99	0.98	0.99
	Width	0.92	0.75	0.97
	CSA	0.78	0.35	0.93

CSA, cross-sectional area

Table 2. Comparison of the two transducers when measuring lumbar multifidus at L3; mean difference $(\overline{d})$ and 95% limits of agreement (LoA) for $(\overline{d})$. No significant differences between transducers were found with T-test (2-tailed) (p>0.05).

	Transducer		Bland and Altman			
Measure	Curvilinear	Linear	$\overline{d}$	95% LoA	SD_d	T-test (p)
Thickness	2.73cm ± 0.31	2.72cm ± 0.39	0.01cm	-0.11-0.13	0.22	0.90
Width	2.52cm ± 0.22	2.66cm ± 0.44	-0.14cm	-0.31-0.23	0.30	0.87
CSA	6.90cm²± 1.1	6.87cm²± 1.1	0.03cm²	-0.31-0.36	0.60	0.87

Mean ± 1 standard deviation

Table 3. Correlations between the linear measurements (squared) and cross sectional area (CSA) of multifidus at L3 for both curvilinear and linear transducers. Pearson's correlation coefficients (r); significance levels (p) indicated by ** for p <0.001.

Transducer	Linear Dimension	r	p
	Thickness^2	0.78	>0.001**
Curvilinear	Width^2	0.83	>0.001**
	Thickness*Width	0.89	>0.001**
	Thickness^2	0.78	>0.001**
Linear	Width^2	0.82	>0.001**
	Thickness*Width	0.85	>0.001**