Manual Therapy 16 (2011) 463–469

Contents lists available at ScienceDirect

Manual Therapy

journal homepage: www.elsevier.com/math

Original article

Discriminative and reliability analyses of ultrasound measurement of abdominal muscles recruitment

Paulo H. Ferreira [a,*], Manuela L. Ferreira [b], Dafne P. Nascimento [a,b], Rafael Z. Pinto [b], Marcia R. Franco [b], Paul W. Hodges [c]

[a] *Discipline of Physiotherapy, Faculty of Health Sciences, University of Sydney, Sydney, Australia*
[b] *The George Institute for Global Health, Sydney Medical School, University of Sydney, Sydney, Australia*
[c] *Centre for Clinical Research Excellence in Spinal Pain, Injury and Health, School of Health and Rehabilitation Sciences, University of Queensland, Brisbane, Australia*

ARTICLE INFO

Article history:
Received 20 October 2010
Received in revised form
9 February 2011
Accepted 11 February 2011

Keywords:
Low back pain
Transversus abdominis
Rehabilitative ultrasound imaging

ABSTRACT

Rehabilitative ultrasound imaging has a great potential to be used as a tool in the assessment of trunk muscle function in patients with low back pain (LBP). However, a further investigation of the discriminative ability of this tool as well as the effect of operators' levels of training on reliability is warranted.

Discriminative analysis of ultrasound and electromyography (EMG) measurements of transversus abdominus (TrA), obliquus internus (OI), and obliquus externus (OE) muscles function between people with and without LBP and the effect of operator's training on reliability of TrA muscle function of chronic LBP patients were conducted. For the discriminative study, measurements were collected from 10 subjects with LBP and 10 matched controls during isometric low load tasks with their limbs suspended. For the reliability study, in stage 1 the reliability of single ultrasonographic static images involved 4 operators (1 trained and 3 non-trained), whereas, in stage 2, two operators (1 trained and 1 non-trained) were used to determine the reliability of TrA thickness change. Methods used in the statistical analysis were pearson correlation and receiver operating characteristic curve for the discriminative study and intraclass correlation coefficient (ICC) for the reliability study.

While ultrasound measures of OE muscle function showed poor association with EMG ($r = 0.28$, $p = 0.22$), TrA and OI function showed moderate to excellent association (TrA: $r = 0.74$, $p < 0.000$; OI: $r = 0.85$, $p < 0.000$). Ultrasound and EMG measures of TrA and OI function discriminated LBP patients from controls. Reliability of the assessment of TrA function with a trained operator (ICC = 0.92; 95% CI: 0.81–0.97) was substantially higher than a non-trained one (ICC = 0.44; 95% CI: −0.41–0.78).

In conclusion, ultrasound measures of deep trunk function is a valid discriminative tool in LBP but highly dependent on operator's level of training.

1. Introduction

Although the evidence for the clinical efficacy of motor control exercises targeting deep trunk muscles appears to be conflicting for acute low back pain (LBP), recent systematic reviews have been shown that this form of exercise is effective in reducing pain and disability in chronic LBP (Ferreira et al., 2006; Macedo et al., 2009). The use of rehabilitative ultrasound imaging (RUSI) for the assessment of deep trunk muscle function, particularly transversus abdominis (TrA), and as a biofeedback tool has become increasingly popular with the implementation of motor control exercises (Hodges, 2005; Teyhen et al., 2007; Whittaker et al., 2007). RUSI is an alternative method of measurement of TrA activation to the reference gold standard, finewire electromyography (EMG), is costly, invasive and offers potential risks for patients such as infection (McMeeken et al., 2004; Whittaker et al., 2007).

RUSI assessment of TrA activation has been used in studies investigating alterations in muscle activation associated with LBP (Critchley and Coutts, 2002; Ferreira et al., 2004; Teyhen et al., 2005; Norasteh et al., 2007; Mannion et al., 2008; Tsao and Hodges, 2008), and as an outcome measure of exercise interventions aimed at restoring these altered patterns of muscle activation (Critchley and Coutts, 2002; Teyhen et al., 2005; Mannion et al., 2008; Tsao and Hodges, 2008). Recent data (Ferreira et al., 2010) has shown that LBP patients with a poor baseline ability to recruit TrA measured using RUSI, had greater reductions in pain when

* Corresponding author. Tel.: +61 2 93519397; fax: +61 2 93519601.
E-mail address: paulo.ferreira@sydney.edu.au (P.H. Ferreira).

1356-689X/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.math.2011.02.010

treated with motor control exercise compared to patients with better TrA recruitment at baseline who underwent the same intervention, highlighting the use of RUSI to identify patients who may benefit from motor control exercise.

Different measures of muscle activation using RUSI have been proposed. These include the degree of sliding of fascicles, diameter, cross sectional area, volume, muscle fascicle length, pennation angle and thickness (Chhem et al., 1994; Hodges, 2005; Whittaker et al., 2007). Measures of TrA thickness can be implemented with single static ultrasonographic recordings of resting or contracted muscle parameters or as the degree of change in thickness expressed as a proportion of a baseline resting value (Costa et al., 2009).

In a previously published study (Ferreira et al., 2004), ultrasound measures of abdominal muscles recruitment TrA, obliquus internus (OI), and obliquus externus (OE) were indirectly compared with EMG in healthy subjects and LBP patients. Results showed that patients with LBP had significantly less activation of TrA compared to healthy subjects. No significant difference was found between groups for OI or OE. Although these results provide preliminary evidence of the ability of RUSI to discriminate people with LBP from healthy subjects, no direct estimate or parameter of the discriminative ability of ultrasound and the reference EMG protocol was provided. In addition, although ultrasound and EMG had similar results, no direct relationship between these two measures was reported.

Another important issue related to the use of RUSI in practice is the magnitude of the reliability of ultrasound measures of abdominal muscles recruitment. This fact remains under frequent scrutiny (Krebs, 1986; de Vet et al., 2006). Results from a recent systematic review have shown that the reliability for RUSI in the assessment of TrA activation ranges from ‘good to excellent’ for single measures of thickness and ‘poor to good’ for measures of thickness change as a proportion of a resting baseline value (Costa et al., 2009). Multiple sources of error can affect the reliability of RUSI in TrA activation, particularly when changes in thickness from resting are analyzed. These sources include i) the impact of associated structures (fascia and organs such as the bladder) on TrA thickness, ii) inaccurate identification of landmarks, iii) position of patient/transducer, iv) trial variation in performance of the activation tasks, v) presence of a competing force on the muscle such as contraction of an adjacent muscle, and vi) training of the operator (Teyhen et al., 2007; Whittaker et al., 2007; Costa et al., 2009).

The need to establish appropriate training programs for operators of RUSI is recognized as a priority (Hodges, 2005; Whittaker et al., 2007). In musculoskeletal diagnostic ultrasonography, structured training of operators is crucial, with formal training being offered by the British Society for Rheumatology, the American College of Rheumatology and the European League Against Rheumatism (Kane et al., 2004). Studies have shown that the level of training and experience achieved by the ultrasound operator in the diagnosis of musculoskeletal conditions affects reliability and lower training levels leads to misdiagnosis and increased complications (Gibbon, 1998; Grebenik et al., 2004). The use of a correct ultrasound scanning technique is considered just as important as image interpretation skills. The need for training guidelines in ultrasonography and competency evaluation in decision-making has also been recognized in other areas such as cardiac diagnosis (DeCara et al., 2003) and prenatal (Maul et al., 2004).

The purpose of the present study was to investigate the discriminative ability and reliability of a protocol previously published in the assessment of abdominal muscle recruitment using ultrasonography in LBP (Ferreira et al., 2004). Specifically, we aim to investigate the relationship between ultrasound and EMG, compare the relative ability of these two measures to discriminate LBP patients from healthy subjects, and to investigate the effect of operators’ level of previous training on the reliability of TrA muscle recruitment.

2. Methods

2.1. Study 1: discriminative analysis of ultrasound measurement of abdominal muscle recruitment

To investigate the discriminative ability of ultrasound measurement of abdominal muscle recruitment we analyzed the data collected from a sample of participants enrolled in a previously published study (Ferreira et al., 2004). In this previous study, ultrasound measures of muscle thickness of TrA, OI, and OE were indirectly compared with EMG. However, no measurement of the ability of these two measurements to discriminate LBP patients from healthy subjects and no direct correlation between them was reported.

2.1.1. Participants

We analyzed data from twenty study participants (10 with a history of LBP and 10 controls). Study participants in the control group had a mean (SD) age, height, and weight of 32.7 (10.6) years, 159.5 (37.8) cm, and 68.2 (12.6) kg, respectively. Study participants were excluded from this group if they had a history of LBP that had restricted function or caused them to have time off work or a history of major illness, or pregnancy in the past 2 years. Study participants in the LBP group had a mean age, height, and weight of 27.8 (5.1) years, 171.5 (10.3) cm, and 68.6 (13.1) kg, respectively. Participants were included in the LBP group if they had a history of at least one episode of LBP that had limited functional activities (work, sports) in the past 18 months and had an episode of LBP within the past 6 months.

2.1.2. Procedure

A more detailed description of the procedures used for the ultrasound measurement of abdominal muscles’ recruitment is provided in the original study (Ferreira et al., 2004). In brief, participants were positioned in supine on a plinth, the hips flexed to 50° and knees flexed to 90°. The legs were suspended by slings wrapped around the knees and ankles and attached to a metal bar placed over the participant. The transducer was positioned transversely across the participant’s abdominal wall along a line midway between the inferior angle of the rib cage and the iliac crest. The medial edge of the transducer was placed approximately 10 cm from the midline. In the test position, study participants were instructed to remain relaxed before testing and then perform isometric knee flexion or extension efforts to target forces based on 7.5% of their body weight measured with a train gauge attached to the ankle. The order of directions of movement was randomized. Three repetitions of each task were performed and static ultrasound images were made at rest and once the target isometric knee flexion or extension force had been reached. All measures were made at the end of expiration, measured with an inductance plethysmograph (Respitrace, NIMMS) placed around the chest (Fig. 1).

2.1.3. Ultrasonography

Ultrasound images were made with a 5.5 cm, 5-MHz curved array transducer (Synergy, GE-Diasonics). The transducer was placed in a dense foam cube to minimize changes in angulation or pressure and was placed transversely across the abdominal wall along a line midway between the inferior angle of the rib cage and the iliac crest. The medial edge of the transducer was placed approximately 10 cm from the midline. The location of the transducer was marked so that identical placement would be used for all measurements. Images were frozen and measured with custom

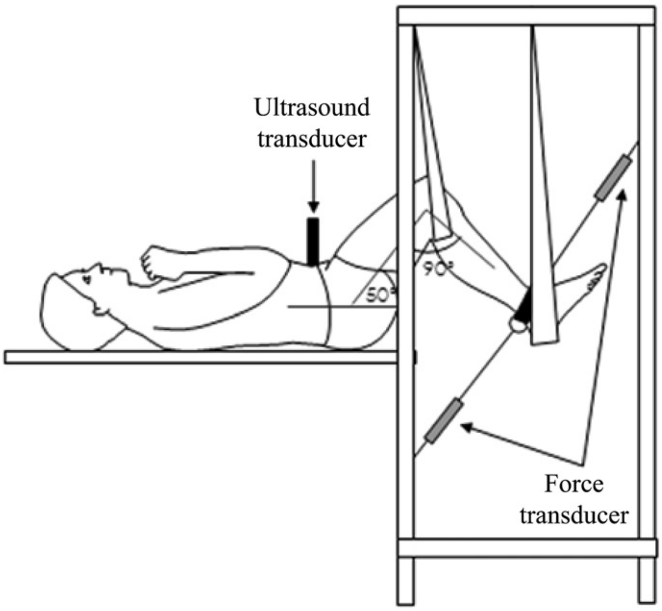

Fig. 1. Experimental setup. The mass of the legs was supported in slings and, in this position, participants performed knee flexion or extension efforts to target force.

designed software using Lab View (National Instruments). A grid was placed over the image and measures of muscle thickness of TrA were made at three sites: the middle of the image and sites 1 cm (calibrated to the image scale) to either side of the midline. Cursors were placed on the superficial and deep boundaries of the muscles at the edge of the hypoechoic region which represents the location of the fascial separation between the muscles. The average of the three measures was recorded for later analysis and the change in thickness was expressed as a proportion of the thickness at rest (Ferreira et al., 2004). Analyses were performed by an assessor blind to group allocation.

2.1.4. EMG

EMG recordings were made using intramuscular finewire electrodes threaded into a hypodermic needle (0.6 × 32 mm) and inserted under the guidance of ultrasound imaging into the right ventrolateral abdominal wall muscles TrA, OI, and OE. Electrodes were inserted midway between the inferior angle of the rib cage and the iliac crest, approximately in the anterior axillary line. EMG data were amplified 2000 times, band-pass filtered between 20 and 1 kHz (Neurolog, Digitimer, UK) and sampled at 2 kHz using a Power1401 and Spike2 software (CED, UK). For purposes of EMG normalization, participants performed two maximal voluntary contractions (MVC) for each muscle. Participants were positioned supine with the hips and knees flexed to 45° and contracted maximally against manual resistance. The tasks were a forced expiratory maneuver for TrA, rotation of the trunk to the left for OE and to the right for OI.

2.1.5. Statistical analysis

For the discriminative study, Pearson Product Moment Correlation was used to investigate the relationship between EMG and ultrasound measures of TrA, OI, and OE recruitment. For the correlation analysis data for cases and controls were pooled. Receiver Operating Characteristic (ROC) curves (DeLong et al., 1988) were constructed to investigate the ability of the EMG and ultrasound measures to discriminate cases from controls. The proportion of the area under the ROC curve (AUC) provides a measure of the test's ability to discriminate people with and without the condition of interest (DeLong et al., 1988). An AUC of 1.0 represents perfect discrimination, whereas an AUC of 0.5 represents discrimination no better than chance (DeLong et al., 1988). ROC analyses were conducted for TrA, OI, and OE ultrasound and EMG measures. Data were analyzed with SPSS Version 17.0 (SPSS Inc., Chicago, IL) and Hong Kong prediction program version 3.0 (Department of Obstetrics and Gynecology – The Chinese University of Hong Kong). Paired AUC values were compared using Delong's test (DeLong et al., 1988).

2.2. Study 2: reliability of ultrasound measurement of TrA muscle recruitment and effect of operators' training

Given the current interest in the reliability of ultrasonographic assessment of TrA recruitment (Costa et al., 2009) and the fact that the patterns of recruitment of this muscle is associated with clinical outcomes of pain and disability in LBP patients treated by motor control exercises (Ferreira et al., 2010), we attempted to investigate the reliability of ultrasound measurement of TrA muscle recruitment and the effect of operators' training using a sample of 10 LBP patients recruited exclusively for study 2. Our aim was two-fold: to investigate the interrater reliability of single measures of TrA thickness obtained from ultrasonographic images (stage 1) and to investigate the intrarater and interrater reliability measures of TrA thickness change, determined by the degree of change in TrA thickness between resting and contracted states (stage 2). We were also interested in investigating the effect of ultrasound operators' previous levels of training (Fig. 2).

2.2.1. Stage 1: interrater reliability of single images of TrA thickness

For the interrater reliability of single images, one previously trained operator (operator 1) and 3 non-trained operators (operator 2; operator 3; operator 4) measured the thickness of 10 randomly selected ultrasonographic images of TrA (single measures of thickness at rest or during contraction) from our sample of 10 chronic LBP participants recruited exclusively for study 2. The trained operator received training in assessing TrA thickness during 3 months preceding data collection. Non-trained operators received basic information on how to measure TrA thickness using the grid and software provided but no previous training was offered. The aim of this stage was to investigate whether operators' level of training affect the assessment of single measures of TrA thickness from ultrasonographic images.

2.2.2. Stage 2: intrarater and interrater reliability of measures of TrA thickness change

In this stage the measure of interest was the intrarater and interrater reliability of RUSI measures of thickness change, and was determined by the degree of change in TrA thickness between the rest and contracted states according to a previous published protocol (Ferreira et al., 2004). This stage included two operators (operator 1; operator 5) who measured RUSI of TrA thickness change of 10 chronic LBP participants during isometric leg contractions at two time points with an interval of 1 week. While operator 1 received the 3-month period of training in the ultrasonography protocol, the operator 5 had no previous training. The aim of this study stage was to investigate whether operators' previous level of training affect the assessment of measures of TrA thickness change as a proportion of a resting baseline. Ultrasound measures of TrA thickness change were performed using the protocol described in study 1.

2.2.3. Participants

Inclusion and exclusion criteria for LBP patients participating in the second stage of the study assessing reliability of TrA thickness change were the same as for study 1. Twenty patients with a mean

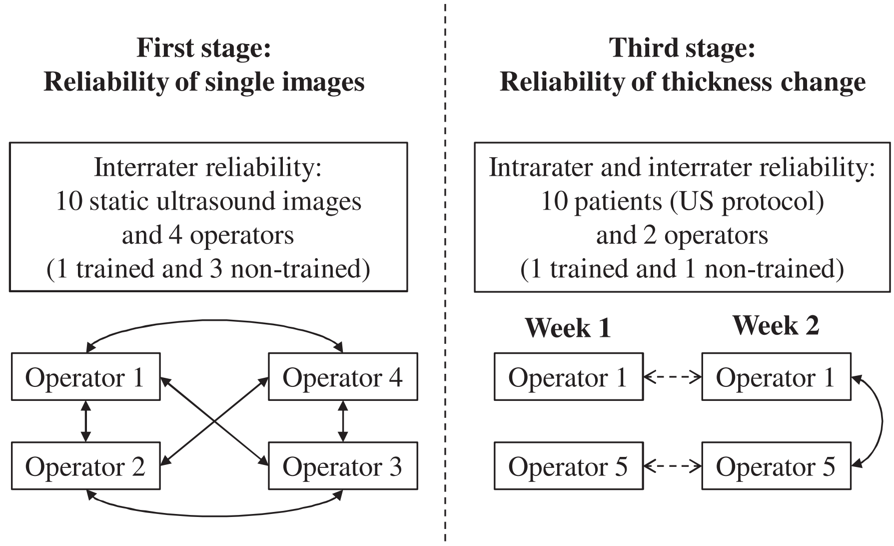

Fig. 2. Representing the two stages of the reliability analysis including type of reliability and number of operators in each stage. Interrater reliability is represented by a straight/curved line and intrarater reliability by a dashed line.

(SD) age of 51.6 (15.5) and mean (SD) levels of pain (0–10 numerical scale) (Ross, 1997) and disability (0–24 Roland Morris Disability Questionnaire) (Roland and Morris, 1983) of 5.7 (2.2) and 10.5 (5.9) respectively.

2.2.4. Statistical analysis

In order to analyze reliability of single images of TrA thickness and TrA thickness change, reliability and agreement measures were calculated for all comparisons. Intraclass correlation coefficients (ICC) were used in the analyses, but the ICC type varied according to the purpose of each stage. For the first stage, an ICC (2,1) was used to compare interrater reliability thickness between pairs of all four operators involved in this stage (Fig. 2). In the second stage that investigated reliability of thickness change, intrarater and interrater reliability were calculated using ICC (3,1) and ICC (2,1) respectively. SPSS Version 17.0 (SPSS Inc., Chicago, IL) was used to calculate reliability coefficients. We also calculated the standard error of the measurement ($SEM = S\sqrt{1 - ICC}$) (Portney and Watkins, 2000) and the minimal detectable change ($MDC = 1.96 \times \sqrt{2} \times SEM$) (Stratford and Binkley, 2000) as additional measures of reliability.

Studies were approved by the Ethics Committees of the University of Sydney and the South Western and Western Sydney Area Health Services.

3. Results

3.1. Study 1: discriminative analysis of ultrasound measurement of abdominal muscle recruitment

Pearson Product Moment Correlation coefficients for the association between ultrasound measures of abdominal muscles' recruitment and EMG varied for different muscles (Table 1; Fig. 3). Strong correlations between EMG and ultrasonography were found for TrA and OI while poor correlation was found for OE.

Results of the ROC curve analysis showed that the AUC was smaller for OE compared to TrA and OI for both ultrasound and EMG measures (Table 2). The AUC for OE was less than 0.5 for both ultrasound and EMG measures. For TrA, the area under the curve was similar for ultrasound and EMG, whereas for OI it was greater for EMG than for ultrasound. Analysis using Delong's test showed no significant difference between EMG and ultrasound for TrA ($z = 0.087$; $p = 0.465$), OI ($z = 0.851$; $p = 0.197$), or OE ($z = 0.414$; $p = 0.339$).

3.2. Study 2: reliability of ultrasound measurement of TrA muscle recruitment and effect of operators' training

ICCs for the analysis of single images ranged from 0.68 (good) to 0.99 (excellent) for interrater reliability. The SEM for single images ranged from 0.60 to 2.71 and the average MDC was 5.22 mm (Table 3).

Intrarater reliability of thickness change in the second stage was excellent ($ICC_{(3,1)} = 0.92$) for the previously trained operator (operator 1) and poor ($ICC_{(3,1)} = 0.44$) for the untrained operator (operator 5). The ICC for the interrater reliability between these two operators was poor ($ICC_{(2,1)} = 0.42$). The average MDC was 14.87% (Table 3).

4. Discussion

The use of RUSI in the assessment of the trunk muscles and as a feedback tool in the treatment of LBP using motor control exercises is becoming popular particularly among physiotherapists. The appeal of RUSI is the fact that it is not an invasive tool and apparently more user friendly than other biomechanical diagnostic tools such as EMG. This apparent easy accessibility of RUSI is likely to haze the need for extensive training of ultrasound operators if the appropriate use of ultrasound as a measure of muscle recruitment is to be achieved.

The interest in RUSI as a measure of muscle recruitment in LBP has been reflected on the number of studies investigating the reliability parameters (Bunce et al., 2002; Kidd et al., 2002; Norasteh et al., 2007; Mannion et al., 2008). However, the effect

Table 1
Correlation coefficients between electromyography and ultrasonography measures of abdominal muscle activation.

Muscles	Correlation coefficients (95% CI)	*p* values
TrA	0.74 (0.37–0.87)	$p < 0.000$
OI	0.85 (0.76–0.96)	$p < 0.000$
OE	0.28 (−0.16 to 0.66)	$p < 0.226$

TrA, transversus abdominis; OI, obliquus internus; OE, obliquus externus.

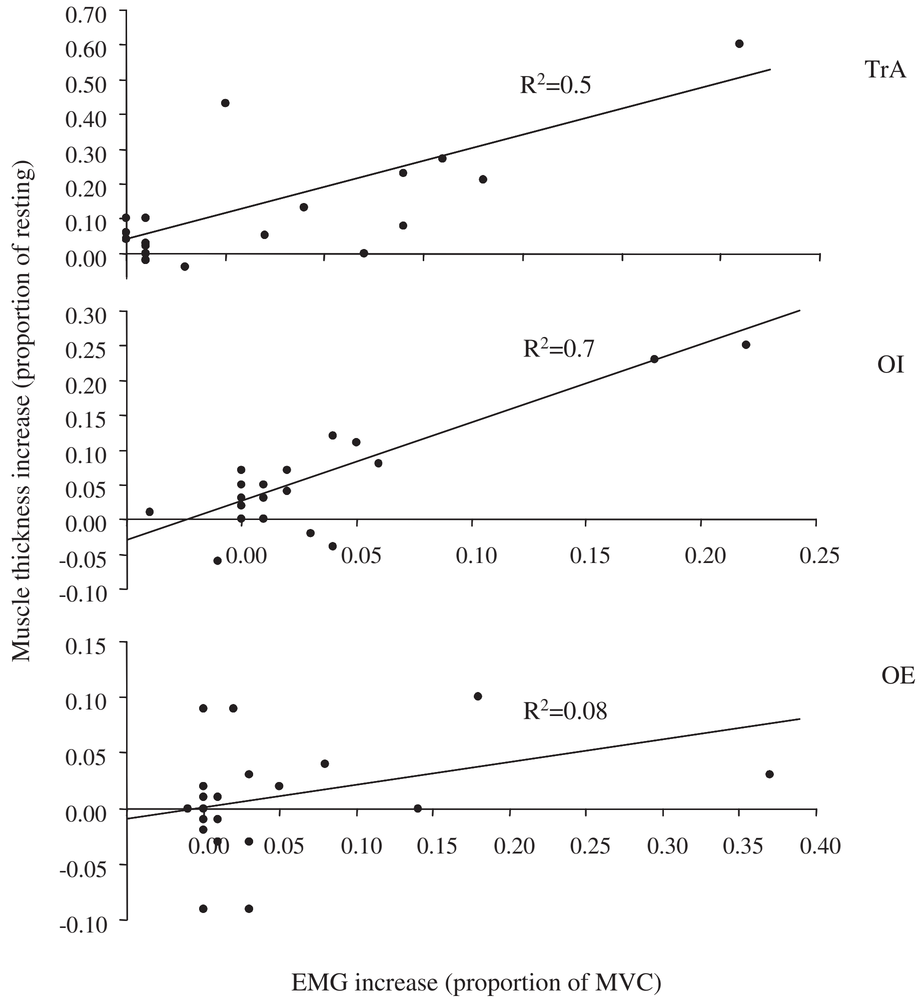

Fig. 3. Correlation between muscle thickness increase (proportion of resting thickness) measured by ultrasound and EMG increase (proportion of MVC) for TrA, OI and OE. Lines represent R squared best fit.

of operators' level of training on reliability estimates has not been extensively investigated in LBP. Additionally, although the ultrasound test protocol investigated in this study has been shown to be a valid tool in the assessment of trunk muscles such as TrA, OI, and OE in LBP (Ferreira et al., 2004), there is current no published data on important issues such as the reliability of the measurement, direct indexes of the association between ultrasound and the gold standard EMG, and parameters of the ability of the measurement to discriminate LBP patients from healthy subjects. The ultrasound protocol investigated in this current study has been found to be useful in predicting LBP patients that will respond more favorably to motor control exercises, with those patients showing smaller increases in TrA thickness at baseline demonstrating significantly better improvements in pain and disability compared to those patients with greater increases in TrA thickness at baseline, who underwent the same intervention (Ferreira et al., 2010). Because of the potential of this measure to be used not only as a diagnostic tool but also as a tool to identify subgroups of LBP patients that respond better to a specific treatment, the investigation and report of reliability and discriminative ability of the measure is warranted.

Table 2
ROC analyses for ultrasound and EMG measures.

Muscles		AUC values (95%CI)
TrA	Ultrasound	0.79 (0.58–1.00)
	EMG	0.77 (0.56–0.98)
OI	Ultrasound	0.66 (0.39–0.93)
	EMG	0.75 (0.52–0.98)
OE	Ultrasound	0.49 (0.22–0.76)
	EMG	0.46 (0.19–0.73)

TrA, transversus abdominis; OI, obliquus internus; OE, obliquus externus. AUC, area under the curve.

Table 3
Reliability of Transversus Abdominus Measures: intraclass correlation coefficient (ICC), standard error of the measurement (SEM) and minimal detectable change (MDC) for all pairs of comparisons calculated for each stage of the study.

Reliability comparisons pairs	ICC (95%CI)	SEM	MDC
First Stage: Reliability of single images $ICC_{(2,1)}$			
Operators: 1 × 2	0.68 (0.17–0.91)	2.71 mm	7.52 mm
Operators: 1 × 3	0.91 (0.58–0.98)	1.10 mm	3.06 mm
Operators: 1 × 4	0.68 (0.17–0.91)	2.66 mm	7.37 mm
Operators: 2 × 3	0.82 (0.46–0.95)	2.10 mm	5.83 mm
Operators: 2 × 4	0.99 (0.96–1.00)	0.60 mm	1.66 mm
Operators: 3 × 4	0.81 (0.44–0.95)	2.12 mm	5.88 mm
Second Stage: Reliability of thickness change $ICC_{(3,1\ \text{intrarater};\ 2,1\ \text{interrater})}$			
Operators: 1 × 1	0.92 (0.81–0.97)	3.38%	9.38%
Operators: 5 × 5	0.44 (−0.41–0.78)	6.01%	16.65%
Operators: 1 × 5	0.42 (−0.48–0.78)	6.71%	18.59%

ICC, intraclass correlation coefficient; SEM, standard error of the measurement; MDC, minimal detectable change.

4.1. Study 1: discriminative analysis of ultrasound measurement of abdominal muscle recruitment

Results of the discriminative analysis showed no statistically significant differences between ultrasound and EMG for TrA, OI, and OE. The proportion of the area under the ROC curve (AUC) reflects the test's ability to discriminate people with and without the condition of interest with an AUC of 1.0 representing perfect discrimination, whereas an AUC of 0.5 represents discrimination no better than chance (DeLong et al., 1988). The greatest AUC was found for ultrasound measure of TrA recruitment (AUC = 0.79) while the smallest AUC was found for EMG activity of OE (AUC = 0.46). These results are in line with the previous findings that while patients with LBP have smaller values of TrA recruitment, no differences are found for OE when compared with healthy subjects (Ferreira et al., 2004). For TrA, the area under the curve was very similar for ultrasound and EMG, showing that ultrasound is likely to be a valid and non-invasive alternative in the assessment of TrA recruitment. Interestingly, results of the present study showed that ultrasound measure of OI recruitment could also discriminate LBP patients from healthy subjects, although the AUC for EMG (0.75) was greater than ultrasound (0.66). Although clinicians can discriminate LBP patients from healthy people by means of an interview or administering questionnaires, the results of this study demonstrate that ultrasound is a valid discriminatory tool in LBP and has the potential to be further used as outcome tool with the implementation of motor control exercises targeting deep trunk muscles.

Substantial correlations between EMG and ultrasonography were found for TrA and OI while only slight correlation coefficients were found for OE. These results are in line with previous research showing that ultrasound is probably not a valid measurement of OE muscle recruitment (Hodges et al., 2003b). Previous research has shown that OE activity is increased with postural tasks in patients with LBP, (Radebold et al., 2000) during shoulder movements with experimentally induced pain (Hodges et al., 2003a) or when pain is anticipated (Moseley et al., 2004). However, no difference in ultrasound measure of OE recruitment was found in the present study between patients with LBP and healthy subjects, and given the findings that OE does not show acceptable estimates of concurrent or discriminatory validity, the use of ultrasound to measure OE recruitment or activity in LBP is not advocated. For TrA and OI the association between ultrasound and EMG found in our study confirms previous findings that for low load tasks increases in TrA and OI thickness are associated with increases in EMG activity (Hodges et al., 2003b). It should be noted though, that confidence intervals for TrA were large (0.37–0.87), and future studies should investigate whether this finding is a function of a small sample or true variability in the population. Interestingly, Pearson's correlation for OI (0.85) was greater than for TrA (0.74), a finding that supports OI recruitment as an acceptable discriminative measure in LBP.

4.2. Study 2: reliability of ultrasound measurement of TrA muscle recruitment and effect of operators' training

One of the aims of this study was to analyze separately the reliability of assessing TrA thickness in single static ultrasound images and the whole ultrasound protocol, including data acquisition and image measurement. We further aimed to compare intrarater reliability estimates of a previous trained ultrasound operator and a non-trained operator as well as the interrater reliability between them. Results indicate that measures of interrater reliability of single static images were highly reliable irrespective of whether a previous trained operator is compared with non-trained operators (ICCs ranging from 0.68 to 0.91) or non-trained operators are mutually compared (ICCs ranging from 0.82 to 0.99). Therefore, training does not appear to be a significant factor for analysis of single static measures and the reliability estimates are in accordance with a previous systematic review showing that measurement of TrA muscle thickness in single static images are acceptable (Costa et al., 2009). However, when the reliability of the whole ultrasound protocol was analyzed, the intrarater reliability for a previous trained operator was substantially higher than the non-trained operator (ICCs = 0.92 and 0.44 respectively). The interrater reliability was low (0.42) and probably not acceptable for a biological measure. The estimates of SEM and MDC associated with reliability were also higher for the non-trained operator with a substantial MDC of 16.65% for the non-trained operator. With this magnitude of MDC, measures of TrA thickness performed by an ultrasound operator with no or low levels of experience would not be able to detect the range of previous clinical within patient change in TrA thickness associated with the implementation of motor control exercises in patients with LBP (95%CI: 5–10%) (Ferreira et al., 2010). The MDC found for a trained ultrasound operator was lower (9.38%). An experienced operator would be able to measure changes in TrA thickness associated with treatment in those patients that improved significantly but probably not in those with smaller increments in TrA recruitment. Future studies should investigate whether longer periods of operators training or adaptations in the current ultrasound protocol could increase the sensitivity of the measurement.

These results have important implications for the use of RUSI in research and in clinical practice. The measurement error associated with the use of RUSI appears to be part of the process of data acquisition and not image analysis. Non-experienced raters appear to be just as reliable as experienced ones in measuring TrA thickness of static ultrasound images. Errors during the process of data acquisition can arise from a variety of sources including patient positioning, placement of ultrasound transducer, pressure applied in the transducer, control of patients' breathing patterns, and monitoring of machine parameters such as resolution and focus. This study offers preliminary evidence that previous training affect the use of RUSI in the assessment of TrA recruitment. However, a small sample and methodological confinements particularly in regards to training regimens are part of the limitations of the study. Further studies investigating the effect of the implementation of structured training programs and the effect on the sensitivity of the measurement is warranted.

5. Conclusion

Ultrasound measurement of abdominal muscle recruitment is a valid tool to discriminate LBP patients and has acceptable association with the gold standard EMG, for TrA and OI muscles. The level of ultrasound operators previous training affects the reliability of TrA assessment during the process of ultrasound data acquisition and should be considered in future research and in the clinical use of ultrasound in LBP.

Acknowledgements

Rafael Zambelli Pinto is a PhD student supported by Capes Foundation, Ministry of Education of Brazil. Paul Hodges was supported by a fellowship from the National Health and Medical Research Council of Australia.

References

Bunce SM, Moore AP, Hough AD. M-mode ultrasound: a reliable measure of transversus abdominis thickness? Clinical Biomechanics 2002;17(4):315–7.

Chhem RK, Kaplan PA, Dussault RG. Ultrasonography of the musculoskeletal system. Radiologic Clinics of North America 1994;32(2):275–89.

Costa LO, Maher CG, Latimer J, Smeets RJ. Reproducibility of rehabilitative ultrasound imaging for the measurement of abdominal muscle activity: a systematic review. Physical Therapy 2009;89(8):756–69.

Critchley DJ, Coutts FJ. Abdominal muscle function in chronic low back pain: measurement with real-time ultrasound scanning. Physiotherapy 2002;88(6):322–32.

de Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. Journal of Clinical Epidemiology 2006;59(10):1033–9.

DeCara JM, Lang RM, Koch R, Bala R, Penzotti J, Spencer KT. The use of small personal ultrasound devices by internists without formal training in echocardiography. European Journal of Echocardiography 2003;4(2):141–7.

DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44(3):837–45.

Ferreira P, Ferreira M, Maher C, Refshauge K, Herbert RD, Hodges PW. Changes in recruitment of transversus abdominis correlate with disability in people with chronic low back pain. British Journal of Sports Medicine 2010;44:1166–72.

Ferreira PH, Ferreira ML, Hodges PW. Changes in recruitment of the abdominal muscles in people with low back pain: ultrasound measurement of muscle activity. Spine 2004;29(22):2560–6.

Ferreira PH, Ferreira ML, Maher CG, Herbert RD, Refshauge K. Specific stabilisation exercise for spinal and pelvic pain: a systematic review. Australian Journal of Physiotherapy 2006;52(2):79–88.

Gibbon WW. Diagnostic ultrasound in sports medicine. British Journal of Sports Medicine 1998;32(1):3.

Grebenik CR, Boyce A, Sinclair ME, Evans RD, Mason DG, Martin B. NICE guidelines for central venous catheterization in children. Is the evidence base sufficient? British Journal of Anaesthesia 2004;92(6):827–30.

Hodges PW. Ultrasound imaging in rehabilitation: just a fad? Journal of Orthopaedic and Sports Physical Therapy 2005;35(6):333–7.

Hodges PW, Moseley GL, Gabrielsson A, Gandevia SC. Experimental muscle pain changes feedforward postural responses of the trunk muscles. Experimental Brain Research 2003a;151(2):262–71.

Hodges PW, Pengel LH, Herbert RD, Gandevia SC. Measurement of muscle contraction with ultrasound imaging. Muscle & Nerve 2003b;27(6):682–92.

Kane D, Balint PV, Sturrock R, Grassi W. Musculoskeletal ultrasound—a state of the art review in rheumatology. Part 1: current controversies and issues in the development of musculoskeletal ultrasound in rheumatology. Rheumatology 2004;43(7):823–8.

Kidd AW, Magee S, Richardson CA. Reliability of real-time ultrasound for the assessment of transversus abdominis function. Journal of Gravitational Physiology 2002;9(1):P131–2.

Krebs DE. Declare your ICC type. Physical Therapy 1986;66(9):1431.

Macedo LG, Maher CG, Latimer J, McAuley JH. Motor control exercise for persistent, nonspecific low back pain: a systematic review. Physical Therapy 2009;89(1):9–25.

Mannion AF, Pulkovski N, Gubler D, Gorelick M, O'Riordan D, Loupas T, et al. Muscle thickness changes during abdominal hollowing: an assessment of between-day measurement error in controls and patients with chronic low back pain. European Spine Journal 2008;17(4):494–501.

Maul H, Scharf A, Baier P, Wüstemann M, Günter HH, Gebauer G, et al. Ultrasound simulators: experience with the SonoTrainer and comparative review of other training systems. Ultrasound in Obstetrics and Gynecology 2004;24(5):581–5.

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):337–42.

Moseley GL, Nicholas MK, Hodges PW. Does anticipation of back pain predispose to back trouble? Brain 2004;127(10):2339–47.

Norasteh A, Ebrahimi E, Salavati M, Rafiei J, Abbasnejad E. Reliability of B-mode ultrasonography for abdominal muscles in asymptomatic and patients with acute low back pain. Journal of Bodywork and Movement Therapies 2007;11(1):17–20.

Portney LG, Watkins MP. Foundations of clinical research applications to practice. 2nd ed. New Jersey: Prentice-Hall; 2000.

Radebold A, Cholewicki J, Panjabi MM, Patel TC. Muscle response pattern to sudden trunk loading in healthy individuals and in patients with chronic low back pain. Spine 2000;25(8):947–54.

Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. Spine 1983;8(2):141–4.

Ross RLP. Clinical assessment of pain. In: van Dieen JH, editor. Assessment in occupational therapy and physical therapy. Philadelphia: WB Saunders; 1997. p. 123–33.

Stratford PW, Binkley JM. A comparison study of the back pain functional scale and Roland Morris questionnaire. North American orthopaedic rehabilitation research network. Journal of Rheumatology 2000;27(8):1928–36.

Teyhen DS, Gill NW, Whittaker JL, Henry SM, Hides JA, Hodges P. Rehabilitative ultrasound imaging of the abdominal muscles. Journal of Orthopaedic and Sports Physical Therapy 2007;37(8):450–66.

Teyhen DS, Miltenberger CE, Deiters HM, Del Toro YM, Pulliam JN, Childs JD. The use of ultrasound imaging of the abdominal drawing-in maneuver in subjects with low back pain. Journal of Orthopaedic and Sports Physical Therapy 2005;35(6):346–55.

Tsao H, Hodges PW. Persistence of improvements in postural strategies following motor control training in people with recurrent low back pain. Journal of Electromyography and Kinesiology 2008;18(4):559–67.

Whittaker JL, Teyhen DS, Elliott JM, Cook K, Langevin HM, Dahl HH. Rehabilitative ultrasound imaging: understanding the technology and its applications. Journal of Orthopaedic and Sports Physical Therapy 2007;37(8):434–49.