ORIGINAL ARTICLE

An investigation of the reproducibility of ultrasound measures of abdominal muscle activation in patients with chronic non-specific low back pain

**Leonardo Oliveira Pena Costa · Chris G. Maher ·
Jane Latimer · Paul W. Hodges · Debra Shirley**

Abstract Ultrasound (US) measures are used by clinicians and researchers to evaluate improvements in activity of the abdominal muscles in patients with low back pain. Studies evaluating the reproducibility of these US measures provide some information; however, little is known about the reproducibility of these US measures over time in patients with low back pain. The objectives of this study were to estimate the reproducibility of ultrasound measurements of automatic activation of the lateral abdominal wall muscles using a leg force task in patients with chronic low back pain. Thirty-five participants from an existing randomised, blinded, placebo-controlled trial participated in the study. A reproducibility analysis was undertaken from all patients using data collected at baseline and after treatment. The reproducibility of measurements of thickness, muscle activation (thickness changes) and muscle improvement/deterioration after intervention (differences in thickness changes from single images made before and after treatment) was analysed. The reproducibility of static images (thickness) was excellent (ICC$_{2,1}$ = 0.97, 95% CI = 0.96–0.97, standard error of the measurement (SEM) = 0.04 cm, smallest detectable change (SDC) = 0.11 cm), the reproducibility of thickness changes was moderate (ICC$_{2,1}$ = 0.72, 95% CI 0.65–0.76, SEM = 15%, SDC 41%), while the reproducibility of differences in thickness changes from single images with statistical adjustment for duplicate measures was poor (ICC$_{2,1}$ = 0.44, 95% CI 0.33–0.58, SEM = 21%, SDC = 66.5%). Improvements in the testing protocol must be performed in order to enhance reproducibility of US as an outcome measure for abdominal muscle activation.

Keywords Ultrasound · Abdominal muscles · Reproducibility · Agreement · Reliability

Introduction

The use of motor control exercise (also known as specific stabilisation exercise) in the treatment of low back pain has become widespread [10, 12, 21, 22]. The rationale for the use of motor control exercises is that the deep abdominal and paraspinal muscles have a critical role in the dynamic control of the lumbar spine [17, 18]. For example, delayed onset of activity of the transversus abdominis muscle (TrA) has been reported in patients with recurrent low back pain [16, 17] compared to asymptomatic subjects. Most of the studies that have measured the activity of the deep spinal muscles use fine-wire electromyography (EMG), which is costly, time-consuming, potentially uncomfortable, and includes risks such as infection. An alternative approach is to measure the recruitment of the muscles indirectly by assessment of morphologic changes of the muscles (i.e. thickness changes) using real-time ultrasound imaging (US) [15].

If ultrasound measures of abdominal muscle activation are to be useful they need to have acceptable reproducibility.

L. O. P. Costa (✉) · C. G. Maher · J. Latimer
The George Institute for International Health,
The University of Sydney, PO Box M201,
Missenden Road, Sydney, NSW 2050, Australia
e-mail: lcos3060@gmail.com; lcosta@george.org.au

P. W. Hodges
NHMRC Centre of Clinical Research Excellence in Spinal Pain,
Injury and Health, School of Health of Rehabilitation Sciences,
The University of Queensland, Brisbane, QLD, Australia

D. Shirley
Faculty of Health Sciences, The University of Sydney,
Sydney, NSW, Australia

Reproducibility is defined as the degree to which repeated measurements provide similar results and will be used in this article as an umbrella term for reliability and agreement. Agreement assesses how close the results of repeated measurements are, by estimating the measurement error in repeated measurements. Reliability assesses whether study subjects could be distinguished from each other, despite measurement errors [8, 31].

The available studies that have evaluated the reproducibility of US measures of abdominal muscle activation answer some, but not all, of the questions about the use of this test to guide the clinical management of low back pain. The reproducibility of US measures for abdominal wall muscles was extensively discussed in a recent systematic review [6]. Most of the 21 eligible studies recruited healthy, young participants, and only three studies recruited participants with low back pain [13, 26, 32].These three studies provided good evidence that US can provide reproducible measures of thickness of the abdominal muscles, which was a consistent finding in the remaining studies of subjects without low back pain [2, 3]. Another important difference among the studies was the task performed by the participants during the US measures. In most of the available studies participants were asked to voluntary contract the abdominal muscles (i.e. abdominal "draw-in" manoeuvre) [26, 30, 32, 33]. Only a few studies used simple tasks to automatically activate the abdominal muscles (e.g. asking patients to move their legs or contract the leg muscles isometrically while the images were made from the abdominal wall) [4, 13].

Physiotherapists commonly use US for measuring either the activation of the muscle (i.e. the difference between the thickness from an image with the muscle activated and the thickness from an image with the muscle at rest—referred to as thickness changes) or the improvement in activation of the muscle after intervention (i.e. as an outcome measure of improvement—referred to as difference in thickness changes over time). Only two previous studies have investigated the reproducibility of measures of thickness changes of the abdominal muscles in patients with low back pain [13, 26]. One study that tested reproducibility of US using a leg force task to produce automatic changes of the abdominal muscles, concluded that the thickness changes measures will be only reproducible if the examiner was highly experienced [13]. The second study testing the reproducibility of US measures in patients performing the abdominal "draw-in" manoeuver found a wide variety of results ranging from poor to excellent reproducibility [26]. To date no study has investigated the reproducibility of the difference in thickness changes over time. Hence, it is still unclear whether US measures are reproducible for the most important measures in a clinical population, that is, measurement of muscle activation using thickness changes and also measurement of improvement/deterioration of muscle activation using differences in thickness changes over time.

The objectives of this study were to estimate the reproducibility of ultrasound measures of automatic activation of the lateral abdominal wall muscles during a leg force task in patients with chronic non-specific low back pain.

Methods

This study was nested within an existing randomised, blinded, placebo-controlled trial that compared the efficacy of motor control exercise (MCE) versus placebo in patients with chronic non-specific low back pain [25]. From the main study sample (n = 154), a sub-sample of the last 35 participants was selected in order to test the automatic recruitment of the abdominal wall muscles by real-time ultrasound imaging. Eleven of the 35 patients refused to be tested after the intervention period leaving 24 patients for the post-intervention follow up (12 in the MCE group and 12 in the placebo group). The characteristics of the participants are presented in Table 1. The study design, procedures and informed consent were approved by The University of Sydney Human Research Ethics Committee.

Eligibility criteria

Participants were included if they had non-specific low back pain of at least 3 months duration, were currently seeking care for low back pain, were aged greater than 18 and less than 80 years, comprehended English, and they expected to continue residing within the study region for the study duration. Exclusion criteria were: suspected or

Table 1 Characteristics of study participants at baseline (n = 35)

Variable	
Gender	
Female (%)	22 (62.9%)
Age (years)	53.3 (11.27)
Low back pain duration (years)	6.2 (7.70)
Weight (kg)	69.3 (11.49)
Height (m)	1.6 (0.08)
Primary outcomes	
Pain intensity numerical rating scale (0–10)	6.6 (2.16)
Global perceived effect (−5 to +5)	−2.2 (2.08)
Patient specific function scale (0–30)	10.1 (5.84)
Secondary outcome	
Roland Morris disability questionnaire (0–24)	12.1 (5.45)

Continuous data are mean (SD), categorical data are N (%)

confirmed serious spinal pathology, pregnancy, nerve root compromise, previous spinal surgery, scheduled for major surgery during treatment or follow-up period, and presence of any contraindication to exercise [1]. We also excluded participants who were able to activate their transversus abdominis muscle for longer than 10 s (as preliminary evidence from a previous trial suggested that these patients were less likely to benefit from a MCE program) [25].

Procedures

Ultrasound measurements

The procedures used in this study followed a previously published protocol [11]. Ultrasound images were made with a 10 cm, 5–10 MHz linear wideband array transducer (Terason ultrasound systems, Teratech). The transducer was placed transversely across the abdominal wall on a point between the inferior angle of the rib cage and the iliac crest and $\sim$10 cm from the umbilicus. This position was then adjusted by slightly moving the transducer head to ensure that the anteriomedial aspect of the transversus abdominis, including its medial edge, was visualised. Additionally, some gentle pressure was applied to the transducer head over the abdominal wall in order to ensure that the orientation of the muscle fibres was perpendicular to the transducer head avoiding possible errors due to artefact anisotropy. The images were then frozen, saved and stored for later data extraction.

Participants were asked to perform a simple task, which is expected to automatically activate the abdominal muscles, which has been described in detail elsewhere [11]. Participants were positioned in supine with the hips flexed to $\sim$50°, and knees flexed to $\sim$90° with the lower legs supported by slings around the knees and ankles. Participants were instructed to perform isometric

knee flexion followed by isometric knee extension. Two images were recorded from the left and then the right abdominal wall during each task; the first images taken with the muscles at rest and the second with the patient performing the isometric movement of the knee equivalent to 7.5% of body weight. Images were recorded at the end of the patient's expiration (patients were instructed to stop breathing without closing the glottis). Two load cells were attached around the ankles in order to provide feedback to the patient about the target force. At each testing occasion 16 images were collected [2 tasks (4 images; being 2 at rest and 2 during activation) × 2 trials × 2 sides]. The order of the tasks and sides was counterbalanced.

Data extraction

The data from the images were extracted using custom-designed imaging software. An electronic grid was placed over the image and the thickness measurements of three muscles (obliquus internus, OI; obliquus externus, OE; and transversus abdominis, TrA) were made 1.0, 1.5 and 2.0 cm from the medial edge of the TrA (Fig. 1). The average of the three measures for each muscle was used for analysis; the change in thickness was expressed as a proportion of thickness at rest.

We estimated the reproducibility of thickness of ultrasound measures by comparing the measures taken from the first and the second static image (i.e. we did not remove the patient from the plinth between the images). The reproducibility of thickness changes (reflecting the activation of the muscle) was calculated by comparing pairs of percentage changes in thickness during the activation tasks at baseline. Finally we calculated the reproducibility of differences in thickness changes over time (representing the improvement or deterioration in muscle activation) by comparing the differences in muscle activation of the first trial against the second trial (Fig. 2) (we also calculated the

Fig. 1 Ultrasound image of the abdominal muscles, the *dashed line* represents the edge of the transversus abdominis muscle, the thickness measures are indicated by the *vertical arrows*. *EO* obliquus externus, *IO* obliquus internus, *TrA* transversus abdominis
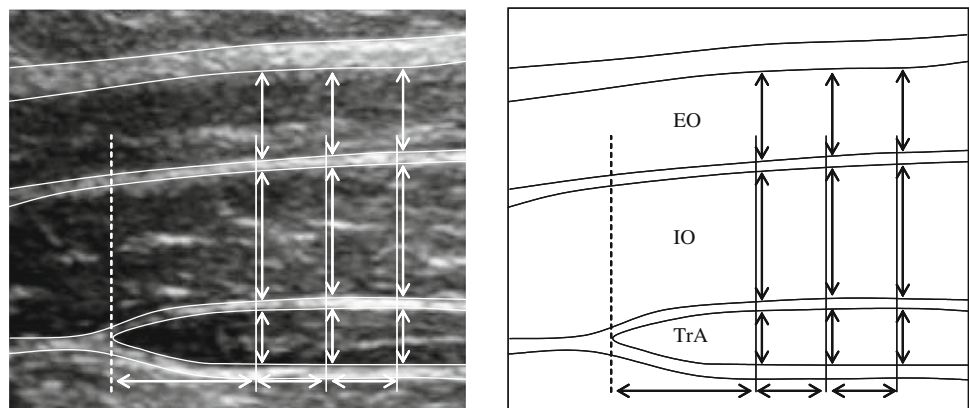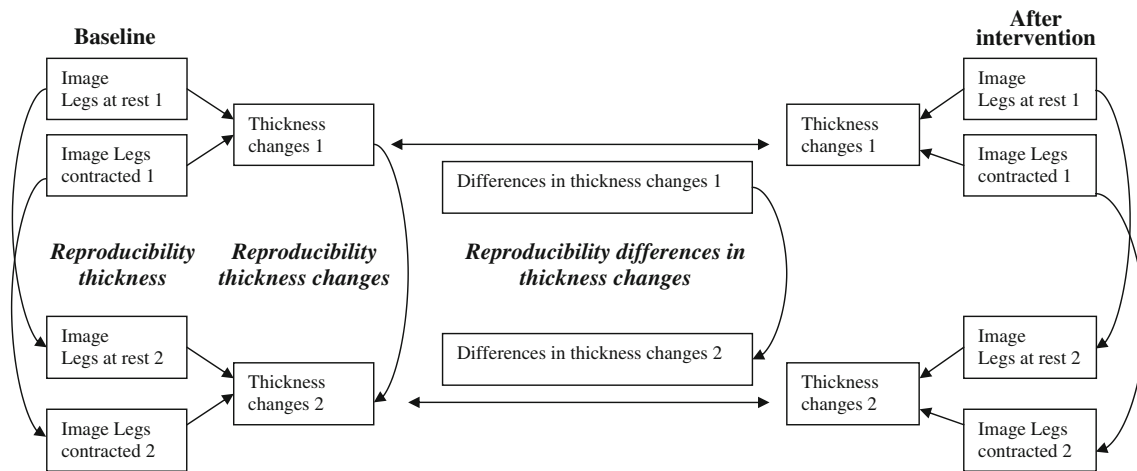
Fig. 2 Diagram representing how the reproducibility tests were performed

reproducibility of the differences in thickness changes across different combinations of baseline and post-intervention scores, finding similar results). It is important to note that this study evaluated the reproducibility of single measures of thickness change and differences in thickness change over time. Because some studies average duplicate measures (to take into account some of the trial-to-trial variability) we estimated the reliability of such measures using the Spearman–Brown formula [9].

Statistical analysis

To describe reliability we calculated the intraclass correlation coefficient (2,1) with 95% confidence intervals. We estimated the reliability of measures derived from the mean of two replicate measures using the Spearman–Brown formula. Our data fit the Spearman–Brown model assumptions [i.e. data were collected in a parallel (test–retest) design and the differences between the standard deviations of the first and the second set of measures were less than 15%] [9]. A guideline for the use of the Spearman–Brown formula advocates that this adjustment technique is accurate for adjustments up to two replicate measures.

We used two measures of agreement: the standard error of the measurement ($SEM_{consistency}$) and the smallest detectable change (SDC). The SEM was calculated by dividing the standard deviation of the mean differences between the two measurements by $\sqrt{2}$ (i.e. $SEM = SD_{differences}/\sqrt{2}$. The SEM reflects the error of the instrument itself. The SDC was calculated using the formula $SDC = 1.96 \times \sqrt{2} \times SEM$. The SDC reflects the smallest within person change in a score that, with $P < 0.05$, can be interpreted as a "real" change, above measurement error in one individual.

Results

We found excellent reliability and agreement values for thickness, moderate reliability and agreement values for thickness changes and finally poor reproducibility for differences in thickness changes over time. The reliability coefficients (i.e. ICCs and adjusted ICCs) and agreement values (SEMs and SDCs) are presented in Table 2.

Discussion

The use of ultrasound imaging by physiotherapists, as a feedback and measurement tool for patients with low back pain, has been increasing in the last decade. Our study aimed to estimate the reproducibility of ultrasound measurements of automatic activation of the lateral abdominal wall muscles during a leg force task in patients with chronic non-specific low back pain. The reproducibility scores were excellent for measuring thickness, moderate for thickness changes and poor for the differences in thickness changes over time.

In terms of reproducibility of measures of muscle thickness, the results of this study are consistent with the available literature on the topic, finding excellent ICC scores and also very small SEM and SDC scores [2–4, 7, 13, 14, 19, 20, 27–29, 32]. This demonstrates that the procedures used for obtaining US images from the abdominal wall and extracting data from the images are reproducible. Additionally it seems that the reliability and agreement of studies that recruited patients with low back pain are similar to the studies that recruited normal subjects. The reproducibility of thickness changes was lower compared to thickness, which is expected given the fact that two images (one with the muscle contracted and

Table 2 Reliability and agreement statistics for ultrasound measures of thickness, thickness changes and differences in thickness changes over time

	Thickness	Thickness changes	Differences in thickness changes over time
N (subjects)	35	35	24
N (images)	840	420	288
Single measure			
ICC$_{2,1}$ (95% CI)	Not applicable	0.56 (0.49–0.62)	0.31 (0.20–0.41)
Average of 2 measures			
ICC$_{2,1}$ (95% CI)	0.97 (0.96–0.97)	0.72 (0.65–0.77)[a]	0.44 (0.33–0.58)[a]
SEM	0.04 cm	15%	24%
SDC	0.11 cm	41.6%	66.5%

[a] Adjusted by using the Spearman–Brown formula

ICC Intraclass correlation coefficient, *SEM* standard error of the measurement, *SDC* smallest detectable change

another with the muscle at rest) are required to determine the level of activation (as reflected by thickness change). We also found a SEM of 15% and a SDC of 41.6% which means that the measurement error is around 15% and there would need to be 41.6% improvement in muscle activation to be sure that a true change had occurred. While measures of thickness changes taken from single images have only low reliability, we estimate that reliability would be moderate (ICC$_{2,1}$ = 0.72, 95% CI 0.65–0.77) for measures based upon the average of two measures. However, measuring differences in thickness changes over time seems to be more difficult and even with statistical adjustment the reproducibility was poor (ICC$_{2,1}$ = 0.44, 95% CI 0.33–0.58).

Although we considered the agreement values (both SEM and SDC) for thickness changes and differences in thickness changes over time high, it is important to consider that there is no normative data in the literature that could allow us to determine the minimum important change (MIC). The ideal scenario would be that the SDC is smaller than MIC [8], but unfortunately these values were not available in the literature.

To date, there is only one study that has estimated reliability of thickness changes for abdominal muscles in low back pain patients and asymptomatic controls [26]. The authors of this study measured the intrarater reliability of ultrasound measurements on 2 separate days in patients performing "abdominal hollowing" (voluntary abdominal task which contrasts to our study in which the recruitment of the abdominal muscles was automatic). Voluntary tasks are more likely to have lower reproducibility as these tasks depend on skill and motivation of the participant, whereas automatic tasks do not have these potential confounders. Their data were analysed in many different ways, but the conclusions were very similar to our study: excellent reliability for thickness (ICC ranging from 0.94 to 0.75); but moderate to poor reliability for thickness changes (0.72–0.26). A potential limitation of our study was that the participants were not repositioned during the process of obtaining the images (i.e. we did not take the participants

off the plinth and then reposition them) and therefore our reproducibility scores for thickness and thickness changes may be overestimated.

We attempted to optimise the precision of the ultrasound measures first by training the examiner intensively, second by using a unique software that uses an electronic grid which avoids visual distortions (most of the previous studies used a "grid over the screen" [11], which is prone to distortions due to the angle of the screen); and finally by using an accurate load cell system in a stable frame which standardises the forces generated by participants performing the leg force task. Additionally we are confident that a clinically relevant population (i.e. patients seeking care for LBP) was selected to participate in the current study.

One potential limitation of this study was that we estimated the reproducibility for duplicate measures based upon single measures. We chose to collect single measures so that the reproducibility study could be accommodated within an existing clinical trial and because many clinicians use single ultrasound measures in clinical practice. While other studies have taken measures based upon the mean of up to 20 replicate measures [4] these estimates of reliability may be artificially high and not representative of clinical practice or be feasible for implementation. We believe that the mean of duplicate measures is a feasible measurement protocol and used the Spearman–Brown formula to estimate the reliability for duplicate measures for operators who collect data in this way. Prior to using the adjustment we checked that the assumptions for the formula [9] held in our data set and so we are confident in the estimates of reliability we provide. Our data are the only data available for the reproducibility of the differences in thickness changes over time [6].

The whole process of performing US measures has multiple sources of error (e.g. accuracy of measurements of distance, identification of landmarks, ability to perform the tasks properly, and position of patient/transducer). Additionally it has to be acknowledged that trial-to-trial variation in performance of the activation tasks is expected. It

would be useful to consider whether modifications in the test protocol may enhance the reproducibility of US measures (especially for thickness changes and differences in thickness changes over time). One approach that has been shown to enhance reliability of other low back assessments is to further standardise the protocol [5, 23, 24]. We believe that the following aspects of the test may need further consideration in order to achieve better standardisation of our testing protocol.

- Although we performed the pre- and post-treatment measures at similar times of the day and asked the patients to empty their bladder before the US tests, there was no specific bowel or bladder preparation for the tests (i.e. diet restrictions, emptying the bowel etc.). This could impact on the abdominal volumes, distending the abdominal muscles and therefore changing the thickness of the muscles during the tasks.
- Although we used a load cell system that could identify whether the patient achieved the target force (i.e. 7.5%), it may have been difficult for patients to repeat the tasks in a similar fashion to that performed at baseline. It is expected that the performance in tasks could vary between trials.
- Although previous studies found similar reliability values for B and M (motion) modes of US [19, 27], we think that the M mode could provide better reproducibility for thickness changes. The main reason for this is that the operator would acquire one image only when using the M mode, compared to acquiring two images when using the B mode.
- Increase the number of repeated measures as this could decrease the variability, enhancing reproducibility. In this study, only single trials were recorded and then adjusted statistically to estimate the benefit of multiple measures. Further studies are needed in order to establish whether an increased number of repetitions for this specific test can improve the reproducibility of the measure to a more acceptable level.

References

1. ACSM (1995) ACSM's guidelines for exercise testing and prescription. Williams and Wilkins, Baltimore
2. Ainscough-Potts AM, Morrisey MC, Critchley DJ (2006) The response of the transverse abdominis and internal oblique to different postures. Man Ther 11:54–60. doi:10.1016/j.math.2005.03.007
3. Beazell JR, Grindstaff TL, Magrum EM, Cullaty M, Hart JM, Shen FH (2006) Comparison of clinical test and real time ultrasound evaluation of muscle contraction in normals and patients with low back pain. J Man Manip Ther 14:168–169
4. Bunce SM, Moore AP, Hough AD (2002) M-mode ultrasound: a reliable measure of transversus abdominis thickness? Clin Biomech (Bristol, Avon) 17:315–317. doi:10.1016/S0268-0033(02)00011-6
5. Chiradejnant A, Maher CG, Latimer J (2003) Objective manual assessment of lumbar PA stiffness is now possible. J Manip Physiol Ther 26:34–39. doi:10.1067/mmt.2003.3
6. Costa LOP, Maher CG, Latimer J, Smeets RJEM (2009) Reproducibility of rehabilitative ultrasound imaging for the measurement of abdominal muscle activity: a systematic review. Phys Ther (in press)
7. Critchley DJ, Coutts FJ (2002) Abdominal muscle function in chronic low back pain patients: measurement with real-time ultrasound scanning. Physiotherapy 88:322–332. doi:10.1016/S0031-9406(05)60745-6
8. de Vet HCW, Terwee CB, Knol DL, Bouter LM (2006) When to use agreement versus reliability measures. J Clin Epidemiol 59:1033–1039. doi:10.1016/j.jclinepi.2005.10.015
9. Feldt LS, Charter RA (2003) Estimating the reliability of a test split into two parts of equal or unequal length. Psychol Methods 8:102–109. doi:10.1037/1082-989X.8.1.102
10. Ferreira ML, Ferreira PH, Latimer J, Herbert RD, Hodges PW, Jennings MD, Maher CG, Refshauge KM (2007) Comparison of general exercise, motor control exercise and spinal manipulative therapy for chronic low back pain: a randomized trial. Pain 131:31–37. doi:10.1016/j.pain.2006.12.008
11. Ferreira PH, Ferreira ML, Hodges PW (2004) Changes in recruitment of the abdominal muscles in people with low back pain—ultrasound measurement of muscle activity. Spine 29:2560–2566. doi:10.1097/01.brs.0000144410.89182.f9
12. Ferreira PH, Ferreira ML, Maher CG, Herbert RD, Refshauge K (2006) Specific stabilisation exercise for spinal and pelvic pain: a systematic review. Aust J Physiother 52:79–88
13. Ferreira PH, Ferreira ML, Maher CG, Refshauge K, Latimer J (2003) Clinical ultrasound test for transversus abdominus thickness: investigation of reliability In: 13th biennial conference—musculoskeletal physiotherapy Australia. Sydney, Australia
14. Hides JA, Miokovic T, Belavy DL, Stanton WR, Richardson CA, Hides JA, Miokovic T, Belavy DL, Stanton WR, Richardson CA (2007) Ultrasound imaging assessment of abdominal muscle function during drawing-in of the abdominal wall: an intrarater reliability study. J Orthop Sports Phys Ther 37:480–486
15. Hodges PW, Pengel LHM, Herbert RD, Gandevia SC (2003) Measurement of muscle contraction with ultrasound imaging. Muscle Nerve 27:682–692. doi:10.1002/mus.10375
16. Hodges PW, Richardson CA (1999) Altered trunk muscle recruitment in people with low back pain with upper limb movement at different speeds. Arch Phys Med Rehabil 80:1005–1012. doi:10.1016/S0003-9993(99)90052-7
17. Hodges PW, Richardson CA (1998) Delayed postural contraction of transversus abdominis in low back pain associated with movement of the lower limb. J Spinal Disord 11:46–56. doi:10.1097/00002517-199802000-00008
18. Hodges PW, Richardson CA (1996) Inefficient muscular stabilization of the lumbar spine associated with low back pain—a motor control evaluation of transversus abdominis. Spine 21:2640–2650. doi:10.1097/00007632-199611150-00014

19. John EK, Beith ID (2007) Can activity within the external abdominal oblique be measured using real-time ultrasound imaging? Clin Biomech (Bristol, Avon) 22:972–979. doi:10.1016/j.clinbiomech.2007.07.005

20. Kidd AW, Magee S, Richardson CA (2002) Reliability of real-time ultrasound for the assessment of the transversus abdominis function. J Gravit Physiol 9:131–132

21. Macedo LG, Latimer J, Maher CG, Hodges PW, Nicholas M, Tonkin L, McAuley JH, Stafford R (2008) Motor control or graded activity exercises for chronic low back pain? A randomised controlled trial. BMC Musculoskelet Disord 9:65. doi:10.1186/1471-2474-9-65

22. Macedo LG, Maher CG, Latimer J, McAuley J (2009) Motor control exercises for persistent nonspecific low back pain: a systematic review. Phys Ther 89(1):9–25

23. Maher C, Adams R (1994) Reliability of pain and stiffness assessments in clinical manual lumbar spine examination. Phys Ther 74:801–809

24. Maher CG, Latimer J, Adams R (1998) An investigation of the reliability and validity of posteroanterior spinal stiffness judgments made using a reference-based protocol. Phys Ther 78:829–837

25. Maher CG, Latimer J, Hodges PW, Refshauge KM, Moseley GL, Herbert RD, Costa LOP, McAuley J (2005) The effect of motor control exercises versus placebo in patients with chronic low back pain. BMC Musculoskelet Disord 6:1–8. doi:10.1186/1471-2474-6-54

26. Mannion AF, Pulkovski N, Gubler D, Gorelick M, O'Riordan D, Loupas T, Schenk P, Gerber H, Sprott H (2008) Muscle thickness changes during abdominal hollowing: an assessment of between-day measurement error in controls and patients with chronic low back pain. Eur Spine J 17:494–501. doi:10.1007/s00586-008-0589-x

27. McMeeken JM, Beith ID, Newham DJ, Milligan P, Critchley DJ (2004) The relationship between EMG and change in thickness of transversus abdominis. Clin Biomech (Bristol, Avon) 19:337–342. doi:10.1016/j.clinbiomech.2004.01.007

28. Norasteh A, Ebrahimi E, Salavati M, Rafiei J, Abbasnejad E (2007) Reliability of B-mode ultrasonography for abdominal muscles in asymptomatic and patients with acute low back pain. J Body Mov Ther 11:17–20. doi:10.1016/j.jbmt.2005.11.002

29. Pietrek M, Sheikhzadeh A, Hagins M, Nordin M (2000) Evaluation of abdominal muscles by ultrasound imaging: reliability, and comparison to electromyography. Eur Spine J 9:309

30. Springer BA, Mielcarek BJ, Nesfield TK, Teyhen DS (2006) Relationships among lateral abdominal muscles, gender, body mass index, and hand dominance. J Orthop Sports Phys Ther 36:289–297

31. Terwee CB, Bot SDM, Boer MR, van der Windta DAWM, Knola DL, Dekkera J, Boutera LM, Veta HCW (2007) Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 60:34–42. doi:10.1016/j.jclinepi.2006.03.012

32. Teyhen DS, Miltenberger CE, Deiters HM, Del Toro YM, Pulliam JN, Childs JD, Boyles RE, Flynn TW (2005) The use of ultrasound imaging of the abdominal drawing-in maneuver in subjects with low back pain. J Orthop Sports Phys Ther 35:346–355

33. Teyhen DS, Rieger JL, Westrick RB, Miller AC, Molloy JM, Childs JD (2008) Changes in deep abdominal muscle thickness during common trunk-strengthening exercises using ultrasound imaging. J Orthop Sports Phys Ther 38:596–605