Reliability of real-time ultrasound measurement of transversus abdominis thickness in healthy trained subjects

Rafael Gnat · Edward Saulicz · Barbara Miądowicz

Received: 17 April 2011 / Revised: 28 December 2011 / Accepted: 28 January 2012 / Published online: 12 February 2012
© Springer-Verlag 2012

Abstract

Purpose To investigate intra- and inter-rater reliability of the ultrasound measurement of transversus abdominis (TrA) thickness and thickness change (difference between thickness at rest and during contraction) in asymptomatic, trained subjects. To define the number of repeated measurements that provide acceptable level of reliability. To investigate variability of the measurements over time of 5 days and the reliability of duplicate analysis of images. **Methods** A single-group repeated-measures design was used to assess reliability. Healthy volunteers ($n = 10$) were subjected to 1-week training in voluntary activation of TrA. Real-time ultrasound imaging and subsequent measurement of the TrA thickness at rest and during voluntary contraction were repeated on Monday, Wednesday and Friday of the next week.

Results Using a single repeated measurement, intraclass correlation coefficients (ICCs) for TrA thickness were: 0.86–0.95 (intra-rater), 0.86–0.92 (inter-rater); and for TrA thickness change: 0.34–0.56 (intra-rater), 0.47–0.61 (inter-rater). Using the mean of three repeated measurements respective values were: 0.97, 0.96–0.98; and 0.81–0.84, 0.80–0.90. No significant differences were found between mean values of TrA thickness as well as thickness change obtained on three consecutive measurement days. Duplicate analysis of the images was highly reliable with ICCs of 0.89–0.99.

Conclusions Two repeated measurements for TrA thickness and at least three measurements for TrA thickness

change are needed to achieve acceptable levels of intra- and inter-rater reliability. In healthy trained volunteers TrA thickness and thickness change are relatively stable parameters over a 5-day period. Duplicate analysis of the same images by two blinded observers is reliable.

Keywords Abdominal muscles · Ultrasonography · Reliability · Rehabilitation

Introduction

Ultrasound (US) tests of lateral abdominal muscles activation became popular together with the evolution of motor control exercises in the treatment of low back pain [1–3]. Most of the studies focus on measurements of transversus abdominis (TrA) thickness and thickness change between rest and contracted state. These parameters are regarded quick and easy to obtain, non-invasive indicators of the TrA performance which is crucial for stability and control of the lumbar spine [2, 3]. Authors conclude that US measurement of TrA thickness generally shows acceptable reliability in healthy persons [6–16] and patients with low back pain [12, 17–22]. Results concerning TrA thickness change are still contradictory or unclear.

A few studies have investigated percent TrA thickness change. For measurements taken on the same day Costa et al. [20] showed moderate intra-rater reliability, Kiesel et al. [12] reported good intra-rater reliability, Kopenhagen et al. [21] showed excellent intra-rater and good to excellent inter-rater reliability, and results from Mannion et al. [19] ranged from poor to excellent.¹

¹ Interpretation of the ICCs is as follows: 0.00–0.50 = poor, 0.50–0.75 = moderate, 0.75–0.90 = good and 0.90–1.00 = excellent reliability.

R. Gnat (✉) · E. Saulicz · B. Miądowicz
Department of Physiotherapy, Research Team Collegium
Magicum, University of Physical Education,
ul. Mikolowska 72, 40-065 Katowice, Poland
e-mail: rafal.gnat@interia.pl

Only three studies [16, 20, 21] have investigated the reliability of TrA thickness change over time. Costa et al. [20], using single measurements, reported poor between-day reliability, whereas Koppenhaver et al. [21] provide evidence for moderate to good reliability using averaged outcomes of two repeated measurements. In turn, Jhu et al. [16] demonstrated moderate to good reliability, which was further improved by means of compensation for motion of the array.

To our knowledge only one study [23] mentions the learning effect associated with motor control tests and exercises. Herbert et al. [23] demonstrate quick improvement in lumbar multifidus activation using constant US feedback and link it to learning effect. Indeed, it may happen that rapid improvement occurring on the beginning of the training is in fact associated with better control, not with physical change of the muscle structure. In this study, we introduced a training of voluntary activation of the TrA prior to measurements. In this way we tried to minimize possible impact of the learning effect on the results. We assumed that after 1 week of training the influence of learning effect on TrA performance would reach its plateau. After eliminating factors that interfere with motor control (pain and learning effect) we could uncover the normal, ‘clear’ variability of TrA thickness and thickness change in healthy subjects which would be beneficial for the purpose of clinical reasoning in patients.

The objectives of the present study, regarding measurements of TrA thickness and thickness change, are:

1. To define the number of repeated measurements that provide acceptable level of reliability and verify results of Koppenhaver et al. [21, 22].
2. To investigate intra- and inter-rater reliability of the measurements. To define cumulated influence of inter-rater source of error and of time-related variables, special attention is paid to between-day inter-rater reliability. No study has previously investigated this aspect, although in clinic or research US measurements are often repeated over a period of time by different raters.
3. To investigate whether measurements are stable over time of 5 days in asymptomatic, trained subjects. It is known that pain interferes with motor control of the abdominal muscles [4, 5, 24] and that there is strong learning effect on the beginning of motor control training [23], both able to influence conclusions drawn from the measurements.
4. To investigate the reliability of duplicate analysis of US images using ‘on screen’ measurement technique in Photoshop 8.0 software (Adobe, San Jose, USA) which minimised visual distortions typical for the ‘frame over the screen’ method used previously [26].

Subjects and methods

Participants

Of the 19 healthy volunteers who responded to our written call for participation, 12 (4 women) were eligible. During the procedure, two participants dropped out. Data are presented from 10 subjects (3 women): mean age 28.35 ± 2.58 years, weight 69.78 ± 5.34 kg, and height 175.23 ± 8.44 cm.

Exclusion criteria were: recent (1 month prior to measurement) history of any low back pain, history of serious low back pain (duration of ≥ 2 weeks and/or resulting in disability and absence from work, and/or obliging to seek medical advice), history of surgery in the area of the lumbar spine, pelvis and lower extremities. All participants gave written informed consent. The study was approved by the local bioethical committee.

Raters

Two raters with limited experience in acquiring and analysing US images participated in the study. Both were graduated specialist in physiotherapy, Ph.D., students at our Department. The raters underwent a 2-week training (3 h/day on weekdays) in US imaging of the muscles of the lateral abdominal wall. This training was guided by an experienced specialist who was not directly involved in the study. After the training, the specialist decided the raters were skilled enough to fulfil their role in the study.

Measurement

A US imaging device Mindray DP 6600 (Mindray, Shenzhen, China) equipped with the 75L38EA linear array was used for the measurements (B-mode, frequency 5 MHz).

With their bladders and bowels empty, participants wearing non-restrictive clothes were positioned supine on the couch with their abdomen exposed. Their legs were parallel to each other, a roll (diameter 15 cm) was placed under the knees, arms were straight at the sides of the trunk, the head, neutral in frontal and transverse planes, in slight flexion. A line perpendicular to the midline of the body was drawn on the right side of the subject’s abdomen at the height of the navel. Then, a crossing point with a vertical line running cranially from anterior superior iliac spine was marked on this line. The medial border of the array was placed at this point and the array was then moved laterally along the transverse line until three muscles of the lateral abdominal wall were properly visualized. The distance between the crossing point and final location of the

medial border of the array was recorded for further measurements. During the procedure small adjustments of the array pressure and angle were allowed without moving it further on the surface of the skin. To recollect all sensations associated with voluntary contraction of the TrA, subjects performed 5×5 -s contractions before measurements started. To minimise the effect of breathing, all images were recorded at the end of exhalation. The rater provided verbal commands: ‘inhale’, ‘exhale’ and, in case of TrA contraction, ‘contract’. While one of the raters was operating the array, the second rater was responsible for recording the image on the sign given by the first rater.

Structure

All volunteers attended the introductory meeting during which they were checked against the exclusion criteria. Those who qualified for the study received detailed information on the objectives and procedures. Since we decided to use voluntary contraction of the TrA, at this point a convenient date was established to start the 1-week training procedure.

Training was based on an abdominal hollowing exercise preferentially activating TrA in relation to more superficial muscles [3, 17, 25]. On Monday, a volunteer was instructed by one of the raters using US guidance about voluntary activation of the TrA. After stable contraction was achieved it was repeated in two series of 10×5 -s repetitions with a 2-min break between the series. Guided training sessions (with US feedback) were then performed

on Wednesday and Friday. On the remaining days the subjects exercised at home using the same number of series and repetitions. In addition, they were given a brochure explaining the main points of the exercise. They were also requested to maintain their usual level of physical activity and to immediately inform the researchers about any pain/ailments in the low back and/or legs.

Measurements, performed by the two raters (order randomized), started 1 week after commencing the training. Each rater acquired an image of the TrA at rest and during voluntary contraction. The cycle of rest–contraction was repeated three times giving a total of six images collected by each rater on 1 day. This same procedure was repeated on Wednesday and Friday. Figure 1 presents an outline of the study structure.

Data analysis

All images were transferred to the computer where they were subjected to analysis in Photoshop 8.0. After the brightness and contrast of the image were adjusted, the layers of fascia intersecting the abdominal muscles were selected using Photoshop tools and shaded in white. Then, three lines were drawn in $1/4$, $1/2$ and $3/4$ of the width of the image along which measurements were taken (Fig. 2). This procedure minimised visual distortions typical for the ‘frame over the screen’ method used previously [26]. The mean value of these three measurements was subjected to statistical analysis. All images were analysed separately by the two raters. To reduce bias, the names of image files were coded by a third person.

Fig. 1 Structure of the measurements and of calculation of the intraclass correlation coefficients. In the left-hand section showing blocks with raters on consecutive measurement days, to see all comparisons taken into account, all *linear arrows* (inter-rater reliability) connecting blocks of rater A and rater B should be mirrored to the opposite side. Similarly, all *solid arrows* (intra-rater reliability) interconnecting consecutive blocks of rater B should be

mirrored to the opposite side. In the upper right-hand section (*solid gray background*) structure of single set of measurements is illustrated. Intra-rater reliability 2/5 = within-rater reliability of measurements performed with 2- and 5-day break; inter-rater reliability 0/2/5 = between-rater reliability of measurements performed on 1 day and with a 2- and 5-day break, respectively

Fig. 2 *Left* raw real-time ultrasound image; *Right* an image after editing in Photoshop software showing three layers of lateral abdominal muscles. Brightness and contrast of the image were adjusted, fascial intersections were shaded in *white*. Transversus abdominis muscle thickness was measured along the three *black lines*

depicted on the edited image. The mean value of these three measurements was subjected to further analysis. *EO* external oblique muscle, *IO* internal oblique muscle, *TrA* transversus abdominis, *ST* subcutaneous tissue

Statistical analysis

Percent differences in TrA thickness between rest and contracted state were calculated for all subjects and are further referred to as ‘thickness change’.

To compare mean values of TrA thickness, we used a mixed model of ANOVA [repeated factor: condition (rest/contraction); independent factors: measurement day, rater]. To compare mean values of TrA thickness change factorial ANOVA was employed (independent factors: measurement day, rater). To compare mean values obtained in duplicate image analysis performed by the two raters, an independent Student’s *t* test was used.

To calculate intraclass correlation coefficients (ICCs) a mixed model of ANOVA was applied to the data with results of consecutive measurements as repeated factor and subjects as independent factor. The reliability of measurements taken on the same day by one rater is not presented since it is inherently included in other types of intra-rater reliabilities (see Fig. 1 for reference). Due to significant similarities of reliability estimates calculated during separate analyses, measurements of TrA muscle thickness at rest and during contraction were pooled and analysed together. This helped to avoid redundancy of results. Outcomes of this reliability analysis are referred to as ‘muscle thickness’.

In addition, the standard errors of measurement ($SEM = SD \times \sqrt{1 - ICC}$) and the smallest detectable differences ($SDD = 1.96 \times SEM \times \sqrt{2}$) were calculated.

Results

Mean TrA thickness (across all performed measurements) was 3.61 ± 0.75 mm at rest and 5.43 ± 1.03 mm during contraction; the difference between these two conditions was significant ($p < 0.001$). Over the course of 5 days TrA thickness showed no significant change. There was also no significant between-day difference in percentage thickness change. Mean thickness change (across all performed measurements) was $52.68 \pm 20.04\%$. No significant inter-rater differences were recorded. Table 1 shows mean values (based on the three repeated measurements) obtained by the two raters on three consecutive measurement days.

Measurements of muscle thickness showed good and excellent reliability irrespective of whether they were taken by one rater with a 2- and 5-day break (Table 2), or by different raters on the same day and with a 2- and 5-day break (Table 3). ICCs were 0.86 and higher. The highest SEM was 0.48 mm, and the highest SDD was 1.34 mm. Averaging the outcomes of repeated measurements had no

Table 1 Mean values ($\pm$ standard deviations; based on the three repeated measurements) obtained by the two raters on three consecutive measurement days

Day	Rater	Thickness rest (mm)	Thickness contracted (mm)	Change in thickness (%)
Monday	1	3.52 ($\pm$ 0.81)	5.34 ($\pm$ 1.09)	53.80 ($\pm$ 19.86)
	2	3.70 ($\pm$ 0.66)	5.43 ($\pm$ 1.03)	47.91 ($\pm$ 18.34)
Wednesday	1	3.53 ($\pm$ 0.74)	5.41 ($\pm$ 1.14)	55.05 ($\pm$ 21.88)
	2	3.58 ($\pm$ 0.86)	5.57 ($\pm$ 1.13)	59.23 ($\pm$ 25.34)
Friday	1	3.58 ($\pm$ 0.71)	5.34 ($\pm$ 1.06)	50.00 ($\pm$ 18.38)
	2	3.73 ($\pm$ 0.87)	5.49 ($\pm$ 0.96)	50.11 ($\pm$ 18.63)

Table 2 Estimates of intra-rater reliability

Reliability	N measurements	Thickness			Thickness change		
		ICC ($\pm$ 95%CI)	SEM (mm)	SDD (mm)	ICC ($\pm$ 95%CI)	SEM (%)	SDD (%)
Intra-rater 2	1	0.95 (0.93–0.97)	0.30	0.83	0.56 (0.30–0.74)	15.36	42.57
	2 (mean)	0.97 (0.95–0.98)	0.22	0.62	0.75 (0.52–0.87)	10.99	30.47
	3 (mean)	0.97 (0.95–0.98)	0.22	0.62	0.81 (0.65–0.90)	9.06	25.12
Intra-rater 5	1	0.86 (0.75–0.92)	0.48	1.34	0.34 (–0.96–0.67)	18.91	52.43
	2 (mean)	0.96 (0.93–0.98)	0.25	0.68	0.71 (0.30–0.89)	10.47	29.01
	3 (mean)	0.97 (0.95–0.99)	0.22	0.60	0.84 (0.60–0.94)	7.28	20.18

Presented are intraclass correlation coefficient (ICC: $\pm$ 95% confidence interval; models: 2,1 for single measurements 2,2 for two repeated measurements, and 2,3 for three repeated measurements), standard error of measurement (SEM) and smallest detectable difference (SDD). Intra-rater reliability 2 describes reliability of measurements performed with a 2-day break (i.e. Monday–Wednesday, Wednesday–Friday); intra-rater reliability 5 describes reliability of measurements performed with a 5-day break (i.e. Monday–Friday) (see Fig. 1)

Table 3 Estimates of inter-rater reliability

Reliability	N measurements	Thickness			Thickness change		
		ICC ($\pm$ 95% CI)	SEM (mm)	SDD (mm)	ICC ($\pm$ 95% CI)	SEM (%)	SDD (%)
Inter-rater 0	1	0.92 (0.87–0.95)	0.37	1.04	0.61 (0.32–0.79)	14.53	40.27
	2 (mean)	0.98 (0.96–0.99)	0.18	0.50	0.85 (0.69–0.93)	8.24	22.84
	3 (mean)	0.98 (0.97–0.99)	0.18	0.50	0.90 (0.78–0.95)	6.34	17.56
Inter-rater 2	1	0.90 (0.85–0.94)	0.42	1.17	0.58 (0.32–0.75)	15.01	41.59
	2 (mean)	0.97 (0.95–0.98)	0.22	0.60	0.82 (0.66–0.90)	8.87	24.59
	3 (mean)	0.96 (0.94–0.98)	0.26	0.72	0.80 (0.63–0.90)	9.30	25.78
Inter-rater 5	1	0.86 (0.76–0.93)	0.48	1.34	0.47 (0.06–0.75)	16.95	46.98
	2 (mean)	0.96 (0.93–0.98)	0.25	0.68	0.73 (0.33–0.89)	10.10	27.99
	3 (mean)	0.97 (0.92–0.98)	0.22	0.60	0.83 (0.59–0.93)	7.51	20.80

Presented are intraclass correlation coefficient (ICC: $\pm$ 95% confidence interval; models: 2,1 for single measurements 2,2 for two repeated measurements and 2,3 for three repeated measurements), standard error of measurement (SEM) and smallest detectable difference (SDD). Inter-rater reliability 0 describes reliability of measurements performed on the same day; inter-rater reliability 2 describes reliability of measurements performed with a 2-day break (i.e. Monday–Wednesday, Wednesday–Friday); intra-rater reliability 5 describes reliability of measurements performed with a 5-day break (i.e. Monday–Friday) (see Fig. 1)

marked influence on ICC values, but decreased SEMs and SDDs. Almost no difference was found between reliability estimates calculated from the means of two and three repeated measurements.

Measurements of percent TrA thickness change showed poor to good reliability depending on whether the calculations were performed using outcomes of single measurements, or the mean of two and three repeated

measurements. Comparable results were registered for measurements taken by one rater with a 2- and 5-day break (Table 2) and by different raters on the same day and with a 2- and 5-day break (Table 3). ICCs calculated from single measurements were the lowest and associated with the largest confidence intervals. In such cases the SEM was higher than 14.53%, and the SDD higher than 40.27%. Using the mean of three repeated measurements, only ICCs

higher than 0.80 were registered; with SEMs and SDDs being less than 9.30 and 25.78%, respectively.

Duplicate analysis of the same images by two raters showed excellent reliability for TrA thickness ($ICC_{(2,1)} = 0.99$; 95% CI 0.99–0.99) and good for thickness change ($ICC_{(2,1)} = 0.89$; 95% CI 0.87–0.90). No significant inter-rater differences between mean values were recorded.

Discussion

Our results indicate that the reliability of TrA US measures can be increased by averaging the outcomes of multiple repeated measurements. In case of muscle thickness, even single measurements proved to have good and excellent reliability; however, we recommend the use of two repeated measurements in clinical and research settings. No significant increase of ICC was registered after introducing a third repeated measurement. These findings are consistent with previous studies [6–14, 16–22].

Single measurements do not provide satisfactory outcomes for TrA thickness change, with ICCs of 0.34–0.61, and SEMs of 14.53–18.91%. However, reliability based on average of three repeated measurements is good, with ICCs ≥ 0.80 . This finding is in agreement with the results of Koppenhaver et al. [21, 22]. In summary, to obtain acceptable reliability we recommend two repeated measurements for muscle thickness and three repeated measurements for thickness change.

In our study, the reliability of the presented measurements remained relatively stable over time. Using a strictly standardized procedure and taking the average values of three measurements we encountered no significant decrease in the reliability estimates. To our knowledge no other study has investigated inter-rater reliability over time. It may be, however, particularly valuable in clinics where measurement of TrA thickness in a given patient is performed by one examiner and later repeated by another. After appropriate training and standardization of the procedure, reliable outcomes can be expected.

We have demonstrated that in healthy trained people TrA thickness at rest and during contraction, as well as TrA thickness change, are relatively stable parameters. Over the 5-day period we did not find significant differences in mean TrA thickness (both at rest and during contraction) and in thickness change. Mean thickness change for all measurements pooled was circa 53%. This is comparable to the results of Critchley and Coutts [27], who recorded thickness change of 49.71% in pain-free controls, however, disagrees with values reported by Koppenhaver et al. [21] who found thickness change of about 80% in symptomatic subjects performing voluntary contraction of the TrA. Since evidence suggest that it is unlikely that patients

would activate their muscles ‘better’ than asymptomatic subjects [27, 28], we can speculate that it was positioning of the subject what determined observed differences. In the study of Koppenhaver et al. [21], the largest TrA thickness change occurred in hook-lying while in supine position it ranged only from 13.1 to 17.9%. Since those authors do not provide a detailed description of subjects’ positioning, we can only guess that, in our study, arrangement of the participants’ body (supine with slight flexion of the knee joints) placed them somewhere between Koppenhaver’s et al. [21] supine and hook-lying. Moreover, basing on our clinical experience, which may be similar to experiences of other specialists in the area, we can also speculate that the participants of the latter study might have forcefully activated TrA together with more superficial abdominal muscles. Patients with low back pain or those unfamiliar with voluntary activation of the TrA (participants of Koppenhaver’s et al. [21] study performed only five repetitions of TrA contraction before measurement) sometimes try to ‘brace’ their abdomens using all muscles instead of activating TrA alone. In our trained asymptomatic subjects, more isolated contraction could have occurred, however, we are unable to provide appropriate references to support this hypothesis. In conclusion, we suggest that in the diagnostic process of the lateral abdominal muscles performance a pattern of activation of all muscles might be more useful than activation of TrA alone.

The proposed technique of image analysis by two different and blinded raters proved reliable. It seems to avoid distortions due to angle of observation typical for the ‘frame over the screen’ methods [26] and agrees with results of previous studies [20]. In case of muscle thickness, the reliability was excellent ($ICC = 0.99$). This finding confirms that most of the measurement error occurs during image acquisition and is not associated with further image analysis. Reliability of TrA thickness change (calculated separately by the two raters) was lower but still good ($ICC = 0.89$). For both TrA thickness and thickness change, we demonstrated that image analysis can be reliably performed by different but similarly trained raters.

Limitations

Results of our study are limited to asymptomatic, young people who underwent specific training of voluntary activation of the TrA. However, the findings seem clinically useful since they create a good reference (with no pain and learning effect interference) for symptomatic populations. The voluntary activation of the TrA may be considered a limitation as well. Evidence shows that a difference exists between TrA muscle response to voluntary or reflex activation [21]. Subjective and arbitrary adjustments of

brightness and contrast during image analysis might constitute a source of error as well because when used inappropriately such adjustments may cause blurring of the muscle border and alter its perceived boundaries.

Conclusions

1. To achieve acceptable levels of intra- and inter-rater reliability, two repeated measurements of TrA thickness and at least three repeated measurements of TrA thickness change are needed.
2. Based on the average of three repeated measurements, intra-rater and inter-rater reliability of TrA thickness and thickness change range from good to excellent. Time interval between measurements (2 or 5 days) does not influence the values of reliability estimates significantly.
3. In asymptomatic subjects trained to voluntarily activate TrA, mean values of muscle thickness and thickness change do not seem to change significantly, at least over the period of 5 days. Mean thickness change between rest and contracted state is about 53%.
4. Estimates of inter-rater reliability of duplicate image analysis performed separately by two raters (TrA thickness ICC = 0.99; thickness change ICC = 0.89) indicate that most of the measurement error occurs during image acquisition and is not associated with further image analysis.

Conflict of interest None.

References

1. Ferreira PH, Ferreira ML, Maher CG, Herbert RD, Refshauge KM (2006) Specific stabilization exercise for spinal and pelvic pain: a systematic review. *Aust J Physiother* 52:79–88
2. 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
3. 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
4. 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(22):2640–2650
5. Hodges PW, Richardson CA (1998) Delayed postural contraction of transversus abdominis associated with low back pain. *J Spinal Disord* 11:46–56
6. McMeeken JM, Beith ID, Newham DJ, Milligan P, Critchley DJ (2004) The relationship between EMG and change in thickness of transversus abdominis. *Clin Biomech* 19:337–342
7. Ainscough-Potts AM, Morrissey MC, Critchley D (2006) The response of the transversus abdominis and internal oblique muscles to different postures. *Man Ther* 11:54–60
8. Rankin G, Stokes M, Newham DJ (2006) Abdominal muscle size and symmetry in normal subjects. *Muscle Nerve* 34:320–326
9. 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
10. Van K, Hides JA, Richardson CA (2006) The use of real-time ultrasound imaging for biofeedback of lumbar multifidus muscle contraction in healthy subjects. *J Orthop Sports Phys Ther* 36:920–925
11. 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
12. Kiesel KB, Underwood FB, Maticola C, Nitz AJ, Malone TR (2007) A comparison of select trunk muscle thickness change between subjects with low back pain classified in the treatment-based classification system and asymptomatic controls. *J Orthop Sports Phys Ther* 37:596–607
13. Endleman I, Critchley DJ (2008) Transversus abdominis and obliquus internus activity during Pilates exercises: measurement with ultrasound scanning. *Arch Phys Med Rehabil* 89:2005–2012
14. Stetts DM, Freund JE, Allison SC, Carpenter G (2009) A rehabilitative ultrasound imaging investigation of lateral abdominal muscle thickness in healthy aging adults. *J Geriatr Phys Ther* 32(2):60–66
15. Whittaker JL, Warner MB, Stokes MJ (2009) Induced transducer orientation during ultrasound imaging: effects on abdominal muscle thickness and bladder position. *Ultrasound Med Biol* 35(11):1803–1811
16. Jhu JL, Chai HM, Jan MH, Wang CL, Shau YW, Wang SF (2010) Reliability and relationship between 2 measurements of transversus abdominis dimension taken during an abdominal drawing-in maneuver using a novel approach of ultrasound imaging. *J Orthop Sports Phys Ther* 40(12):826–832
17. 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
18. Norasfeh A, Ebrahimi E, Salavali M, Rafiei J, Abbasnejad E (2007) Reliability of B-mode ultrasonography for abdominal muscles in asymptomatic and patients with acute low back pain. *J Bodyw Mov Ther* 11:17–20
19. 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
20. Costa LOP, Maher CG, Latimer J, Hodges PW, Shirley D (2009) An investigation of the reproducibility of ultrasound measures of abdominal muscles activation in patients with chronic non-specific low back pain. *Eur Spine J* 18:1059–1065
21. Koppenhaver SL, Hebert JJ, Fritz JM, Parent EC, Teyhen DS, Magel JS (2009) Reliability of rehabilitative imaging of the transversus abdominis and lumbar multifidus muscles. *Arch Phys Med Rehabil* 90:87–94
22. Koppenhaver SL, Parent EC, Teyhen DS, Hebert JJ, Fritz JM (2009) The effect of averaging multiple trials on measurement error during ultrasound imaging of transversus abdominis and lumbar multifidus muscles in individuals with low back pain. *J Orthop Sports Phys Ther* 9(8):604–611

23. Herbert WJ, Heiss DG, Basso DM (2008) Influence of feedback schedule in motor performance and learning of a lumbar multifidus muscle task using rehabilitative ultrasound imaging: a randomized clinical trial. *Phys Ther* 88:261–269
24. Moseley GL, Hodges PW (2005) Are the changes in postural control associated with low back pain caused by pain interference? *Clin J Pain* 21:323–329
25. Richardson CA, Jull GA (1995) Muscle control-pain control. What exercises would you prescribe? *Man Ther* 1:2–10
26. 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
27. Critchley DJ, Coutts FJ (2002) Abdominal muscle function in chronic low back pain patients. Measurement with real-time ultrasound scanning. *Physiotherapy* 88:322–331
28. Teyhen DS, Williamson JN, Carlson NH, Suttles ST, O’Laughlin SJ, Whittaker JL, Goffar SL, Childs JD (2009) Ultrasound characteristics of the deep abdominal muscles during the active straight leg raise test. *Arch Phys Med Rehabil* 90:761–767