‘This is the peer reviewed version of the following article: Connie Jean Whittle, Carol Ann Flavell & Susan Jayne Gordon (2017): Methodological consistency and measurement reliability of transversus abdominis real time ultrasound imaging in chronic low back pain populations: a systematic review, Physical Therapy Reviews, DOI: 10.1080/10833196.2017.1287151

Methodological consistency and measurement reliability of transversus abdominis real time ultrasound imaging in chronic low back pain populations: a systematic review

Running heads:

A systematic review
Whittle et al.

Whittle Connie Jean [1, *]

Flavell Carol Ann [1]

Gordon Susan Jayne [1, 2]

[1] *College of Healthcare Sciences, James Cook University, Douglas Campus, Townsville, Australia*

[2] *School of Health Sciences, Flinders University, Bedford Park, Australia*

[*] Correspondence to: College of Healthcare Sciences, James Cook University, Douglas Campus, Townsville, Qld, 4811, Australia. Email: connie.whittle@my.jcu.edu.au

Background: Real time ultrasound imaging (RTUI) is used to measure transversus abdominis (TrA) thickness in low back pain (LBP) populations. However, individuals with chronic low back pain (CLBP) pose specific imaging challenges, such as older age and higher body mass index, compared to asymptomatic populations or acute and sub-acute LBP groups. These challenges potentially increase measurement error and may require different imaging methods.

Objectives: This review describes the methodologies and reported reliability for RTUI measurement of TrA specific to CLBP populations.

Methods: A systematic database search of Medline, CINAHL, PEDro, the full Cochrane library, Scopus, and Informit identified 20 studies that used RTUI to measure TrA of CLBP participants. Two independent raters appraised the quality of the studies using the QualSyst and the QAREL critical appraisal tools.

Results: Methodological quality varied from low to high. Methods for patient and transducer positioning and muscle measurement were inconsistent between studies. Eight articles cited reliability results from past studies of non-CLBP populations. Only two studies reported reliability in CLBP populations specifically and found higher Intraclass Correlation Coefficients for thickness measures at rest (0.63–0.97), compared to thickness change over time or contraction ratios (0.28–0.80).

Conclusions: Inconsistency of methodology, variable methodological quality, and limited and variable reliability reporting was highlighted in this review. This LBP subgroup poses challenges for RTUI, therefore future research should include standarized methods for image acquisition. This will improve the quality of study methods, reliability of TrA measurement, and improve the applicability and comparibility of research evidence available to clinicians.

Keywords

Transversus abdominis
Reliability
Real time ultrasound
Back pain

Introduction

Real time ultrasound imaging (RTUI) is used in research and clinical practice to measure various musculoskeletal structures.[1–3] In the last 20 years, a plethora of studies have measured transversus abdominis (TrA).[4–8] Through its attachment to the linear alba and thoracolumbar fascia, TrA provides spinal stability via multiple mechanisms, including feed-forward activation.[9] Hence it has been investigated in sports performance, and in relation to low back pain (LBP).[10]

Several studies agree that TrA morphology and function is altered in patients with low back pain.[11–14] When compared to healthy populations, both delayed anticipatory activation and reduced muscle thickness has been reported in individuals who had experienced at least one previous episode of LBP, but were asymptomatic at the time of RTUI.[12] In part, this provides insight into the progression from acute to recurrent low back pain and then chronic pain via central sensitization.[15]

Construct validity of RTUI has been demonstrated in pain-related conditions and specific anthropometric and demographic groups.[16,17] Furthermore, reviews have been conducted to evaluate the validity and reliability of RTUI specifically to measure TrA.[18,19] Validity has been established through comparison of RTUI to magnetic resonance imaging (MRI) and electromyography (EMG).[6,7,20] Validity can be affected by methods; for example, evidence shows that low-level contractions increase the accuracy of RTUI measures of TrA.[7,21,22]

Much current information regarding RTUI reliability is based on research with asymptomatic individuals or low back pain populations of varying durations.[19] In these populations, differences such as significantly larger TrA muscle thickness in males and participants with higher body mass index (BMI) has been reported previously.[23,24] Such demographic and anthropometric differences may affect RTUI measurement reliability, and this has particular relevance in CLBP populations. Individuals with CLBP have an increased average BMI, and RTUI poses extra challenges to quality and reliable TrA image production.[25]

Reliability of RTUI is affected by ultrasound mode, transducer type and placement, ultrasound frequency, patient positioning and task conducted, image acquisition, measurement method, and operator skill level. Each of these potential confounders should be controlled in clinical and research situations[24] and may need to be altered to address subgroup factors, such as higher BMI in the CLBP population. In recent years, RTUI methods used to investigate TrA have become more diverse[17,18] and advances in RTUI technology, if adopted, provide opportunities to address methodological inconsistencies in previous studies.

This systematic review identifies the current evidence for measurement methods and reliability reporting of TrA using RTUI in the CLBP population.

Method

This systematic review was registered with the international prospective register of systematic reviews (CRD42014013522). The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement was used to guide this review.[26]

Search strategy

Studies were identified through a systematic search of Medline via OvidSP (1946-September 2014), CINAHL, PEDro, the full Cochrane library, Scopus, and Informit with no date restrictions, finalized on the 10 May 2016 (CW & CF). Search terms included ‘ultrasound,’ ‘transversus,’ and ‘abdominal,’ grouped, using Boolean operators, and truncated as necessary.

Study selection

Suitable articles identified by title were exported from each database to Endnote (Version 17.3.1), and duplicates were removed. Two independent reviewers (CW & CAF) examined the abstracts and assessed for full-text suitability. Searching and cross-reference of the reference lists was conducted to identify further suitable articles. Consensus on full-text inclusion was reached and both reviewers proceeded to independently appraise full-text articles against eligibility criteria. Authors were contacted for clarification regarding eligibility if it was otherwise unclear. Reviewer disagreements on article selection were resolved via discussion and consensus.

Inclusion criteria

Participants: Over 18 years of age with CLBP, defined as LBP lasting more than three months.[15]

Outcome measures: Studies, which measured muscle size and/or morphology of TrA using RTUI.

Types of studies: Studies published in English and in peer-reviewed journals.

Date of publication: No restrictions were placed on the date of publication.

Exclusion criteria

Non-peer reviewed publications, trial protocols, unpublished work, opinion or discussion papers, single case studies, conference papers, and review articles were excluded. Studies were also excluded if the duration of low back pain symptoms for included participants was not explicitly stated.

Data extraction, synthesis, and methodological quality analysis

Both reviewers extracted relevant information and assessed the included articles for methodological rigor and bias using two critical appraisal tools; namely the QualSyst and the QAREL.[27,28] All included articles were assessed using the QualSyst, produced by the Alberta Heritage Foundation for Medical Research.[27] This appraisal form covers participant selection and demographics, randomization and blinding, outcome measure definition, analysis, and confounding. The tool produces a summary score for each article between zero (lowest) and one (highest). A score of 0.75 was interpreted as moderate quality.[29]

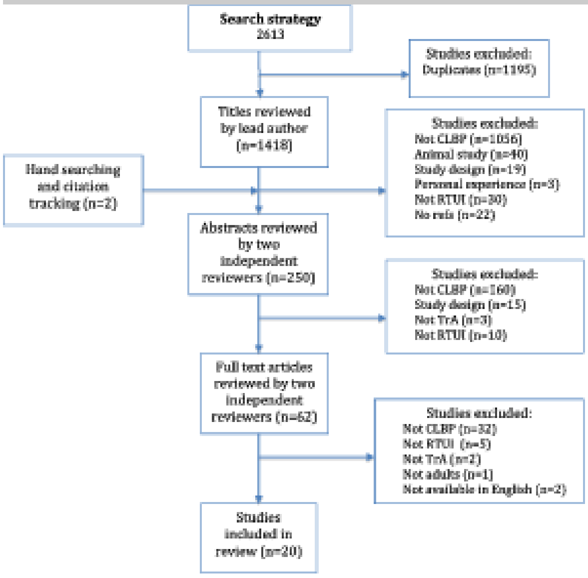
Reliability studies were also assessed using the QAREL.[28] Studies which reported reliability results without reporting the methods to achieve these results were not examined using the QAREL as there was not enough information to determine a fair score. This tool examined participants, examiners, blinding, timing of tests, analysis, and appropriateness of statistics. This form scores questions with yes, no, unclear, or not applicable (N/A) and produces a score between zero and 11. A score > 7 was classified as low risk of bias, <7 as high risk of bias, and 7 as moderate risk of bias in accordance with the classification system used in a previous systematic review.[30] All discrepancies were resolved via consensus. The scores from the two critical appraisal tools were used to assess and compare the overall quality of the included articles and to determine the credibility of the results.

Results

Study selection

The electronic database search yielded 2613 articles. Following exclusion of duplicates, 1418 remained and title screen yielded 248 eligible articles. A further two studies were identified through hand searching and citation tracking.[31,32] Both reviewers appraised the abstracts of 248 studies and jointly approved the eligibility of 62 for full-text screening. Following full screening, 20 studies were eligible for review (Figure 1).

Figure 1 Study selection process



Risk of bias within studies

The QualSyst assessment[27] was applied to all 20 articles and the QAREL tool[28] was applied to the four reliability studies.[11,32–34] QualSyst scores ranged from 0.19 to 1. Most studies were found to have moderate to high quality,[11,13,32–44] with only 5 studies scoring below 0.75[31,45–48] (Table 1). Results of the QAREL ranged from three to seven (Table 2). One study demonstrated a moderate risk of bias,[33] while the remaining three studies demonstrated a high risk of bias.[11,32,34]

Table 1 QualSyst critical appraisal results

Citation	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Score
Akbari et al[35]	2	2	1	2	2	2	0	1	1	2	0	1	1	1	0.75
Cho et al[31]	1	1	1	2	NA	1	0	1	1	1	0	2	1	0	0.46
Costa et al[32]	2	2	2	2	NA	0	NA	2	2	2	2	NA	2	2	0.91
Critchley and Coutts[11]	2	2	1	2	NA	2	NA	2	2	2	1	1	2	1	0.83
Ferreira et al[36]	2	1	2	2	NA	2	NA	1	2	2	2	2	2	2	0.92
Hides et al[13]	2	1	1	2	NA	1	NA	2	2	2	2	2	2	2	0.88

Citation	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Score
Huang et al[47]	1	1	0	1	0	0	NA	0	0	1	0	0	0	1	0.19
Kim et al[45]	2	1	1	0	NA	0	NA	1	2	2	2	0	2	1	0.58
Kim et al[37]	2	1	1	2	NA	0	NA	2	1	2	2	2	2	2	0.79
Lariviere et al[33]	2	2	2	2	NA	2	NA	2	2	2	2	2	2	2	1
Mannion et al[38]	2	1	2	2	NA	0	NA	2	2	2	2	1	2	2	0.83
Mannion et al[34]	2	1	2	2	NA	2	NA	2	2	2	1	0	2	2	0.83
Ohe et al[48]	1	2	1	2	2	0	0	2	2	2	0	1	1	1	0.61
Ota and Kaneoka[39]	2	1	1	2	NA	0	NA	1	2	2	2	2	2	2	0.79
Pinto et al[41]	2	1	1	2	NA	2	NA	2	2	2	2	1	2	2	0.88
Pinto et al[40]	2	1	1	2	NA	2	NA	1	2	2	2	2	2	2	0.88
Pulkovski et al[42]	2	1	2	2	NA	2	NA	2	1	2	2	0	2	2	0.83
Roddey et al[46]	2	2	1	1	NA	0	NA	1	0	2	2	0	2	2	0.63
Unsgaard-Tøndel et al[43]	2	2	2	2	NA	0	NA	1	2	2	2	2	2	2	0.88
Vasseljen and Fladmark[44]	2	2	2	2	2	2	NA	1	2	2	2	1	2	2	0.92

Table 2 QAREL critical appraisal results

	Citation			
Questions:	**Costa et al**[32]	**Critchley and Coutts**[11]	**Lariviere et al**[33]	**Mannion et al**[34]
1. Representative sample	Yes	No	Yes	Yes
2. Representative raters	Unclear	Unclear	Unclear	Unclear
3. Blinding (other raters)	Unclear	NA	Unclear	Unclear
4. Blinding (own findings)	Unclear	Unclear	Unclear	Unclear
5. Blinding (reference/disease)	NA	NA	NA	NA
6. Blinding (clinical information)	NA	NA	Yes	NA
7. Blinding (additional cues)	Unclear	Yes	Yes	Unclear
8. Examination order varied	Yes	NA	Yes	NA
9. Appropriate time interval	Unclear	Yes	Yes	Unclear
10. Test appropriate	Yes	Yes	Yes	Yes
11. Statistics appropriate	Yes	Yes	Yes	Yes
Internal validity (%) (Q:3–9)	1/5 = 20	2/3 = 66.67	4/7 = 57.14	0/5 = 0
External validity (%) (Q:1–2, 10)	2/3 = 66.67	1/3 = 33.33	2/3 = 66.67	2/3 = 66.67
Statistical methods (%) (Q:11)	1/1 = 100	1/1 = 100	1/1 = 100	1/1 =100
Overall total number of ‘Yes’	4	3	7	3

Risk of bias	High	High	Moderate	High
Questions:	Costa et	Critchley and	Lariviere et	Mannion et

Citation

Note: NA= not applicable.

Methodological appraisal of studies

Methodological aspects extracted, recorded and analyzed included ultrasound mode, transducer type and placement, ultrasound frequency, patient positioning and task conducted, operator skill level, image acquisition, and measurement (Table 3).

Table 3 Description of study methods

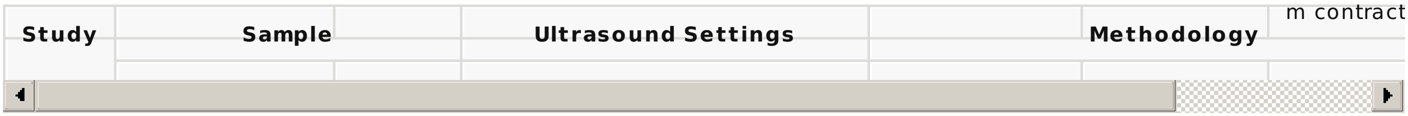
Study	Sample		Ultrasound Settings	Methodology		
	Demographics	BMI (kg/m²)	Transducer/frequency/mode	Position/Task	Transducer Placement	Measurer
Akbari et al[35]	**49 CLBP**; Gender not reported; Age (years) 40 (4) 39.8 ± 3.6	25.2 ± 1.36	7.5 MHz transducer B-mode	Nil activation Supine	Midway between costal margin and iliac crest along anterior axillary line	Muscle thic ss
Cho et al[31]	**17 CLBP**; 9 male, 8 female; Age (years) 34 (9.5)	CLBP 21.85 ± 2.72	8 MHz linear transducer	Automatic activation (a) Supine, knees 90° (b) Sitting, hips and knees 90° (c) Standing, feet shoulder width apart Right side only Images taken at rest, at the end of normal expiration and at the end of forced expiration	2.5 cm distal from a midpoint between the iliac crest and rib	Muscle thic ss Taken vert 1 cm from aponeuros Average of easures us
	17 healthy; 9 male, 8 female; Age (years) 29 (7)	Control 22.75 ± 2.86				
Costa et al[32]	**24 CLBP**; 13 male, 22 female; Age (years) 53 (11)	Not reported	5-10 MHz linear transducer	Automatic activation Supine, hips ~50°, knees ~90° and supported by slings Isometric knee flexion/extension Images taken at rest and contraction twice right and left	Between inferior angle of rib cage and iliac crest ~10 cm from the umbilicus Adjusted so that muscles were visualized Gentle pressure applied Participants were not repositioned between images	Muscle thic ss Taken at 1 and 2 cm fr TrA aponeu s using ele nic grid Average of easures us

Study	Sample		Ultrasound Settings	Two trials	Methodology	
	Demographics	**BMI (kg/m²)**	**Transducer/frequency/mode**	**Position/Task**	**Transducer Placement**	**Measurer**
Critchley and Coutts [11]	**20 CLBP**; 10 male, 10 female; Age (years) 40 (11)	CLBP 24.92 ± 4.32	7 MHz linear transducer	Voluntary activation	2.5 cm anterior to a midpoint between ribs and iliac crest on the mid-axillary line	Muscle thic ss
		Control 22.09 ± 4.97		AH in 4PTK		Taken usin tomatic ca s
	24 healthy; 9 male, 13 female; Age (years) 32 (11)			Standardized instructions		Images fro at the end halation at and at the of AH
				Standardized lumbar lordosis		
				No visual or verbal feedback		
				Images taken at rest and contraction twice		
Ferreira et al [36]	**34 CLBP**; 11 male, 23 female; Age (years) 49 (15.5)	Not reported	5 MHz linear transducer	Automatic activation	On a line midway between the inferior angle of the rib cage and the iliac crest	Muscle thic ss
				Supine, arms across chest, hips 50°, knees 90°		Taken in m e and 1 cm er side of r ne using g
				Isometric knee flexion/extension to 7.5% body weight measured by spring scale	Adjusted so TrA aponeurosis was in the right one-third of the image	Average of e 3 measu sed
				Images taken at rest and contraction twice for both movements	Location recorded for standardization across sessions	
				Order of testing randomized		
Hides et al [13]	**19 CLBP**; 12 male, 7 female; Age (years) 28 (10)	CLBP 24.3 ± 4.0	5 MHz convex transducer	Automatic activation	Perpendicular to muscle fibers on a line midway between the inferior angle of rib cage and iliac crest	Muscle thic ss and slid
	20 healthy; 6 male, 14 female; Age (years) 24.5 (6)					Videos stor or offline m ures
				Supine, knees 60°, ASIS and PSIS aligned vertically	Adjusted so TrA aponeurosis was 2 cm from medial edge of image	Thickness n between al lines at le of image
		Control 22.8 ± 2.7		Isometric leg press to 50% of body weight		Slide taker uperimpos est and co tion image d measurir stance bet n aponeuro
				Images taken at rest, 25% and 45% body weight force, 3 times each leg		
				Order of testing randomized		
Huang et al [47]	**12 CLBP**: 5 male, 7 female Age (years) 26 (5)	Not reported	5 MHz linear transducer	Nil activation	Adjacent and perpendicular to the abdominal wall, 25 mm anteromedial to the midpoint between the ribs and ilium on the mid-axillary line	Saved as s mages
			B-mode			Measured l een fascial ndaries
					Parallel to the muscle fibers of TrA	

Study	Sample		Ultrasound Settings	Methodology		
	Demographics	BMI (kg/m^2)	Transducer/frequency/mode	Position/Task	Transducer Placement	Measurer
Kim et al[45]	**20 CLBP**; Gender not reported; Age not reported **20 healthy**; Gender not reported; Age not reported	Not reported	7.5 MHz linear transducer B-mode	Voluntary activation Sitting, AH (10 sec rest, 10 sec AH, 3 repetitions)	Between iliac crest and the edge of the rib	Muscle thic ss Vertical me ement 15 r rom TrA ap urosis
Kim et al[37]	**15 CLBP**; 9 male, 6 female; Age (years) 35 (10.5) **15 healthy**; 9 male, 6 female; Age (years) 29 (6)	Not reported	8 MHz linear transducer	Automatic activation Crook lying, complete expiration with 5 second hold Images taken at rest and at the end of expiration	2.5 cm anterior to center line between iliac crest and lower rib	Muscle thic ss Taken perp cularly, 1 c om TrA apo rosis Average of easures us
Larivière et al[33]	**15 CLBP**; 5 male, 10 female; Age (years) 49.5 (6) **15 healthy**; 8 male, 7 female; Age (years) 39 (6.5)	CLBP 23 ± 2.5 Control 25.5 ± 3	3-12 MHz linear transducer	Automatic activation (a) Contralateral SLR (b) Bilateral hook lying leg raise (hips 135° and knees 90°) No feedback Images taken 3 times (twice on one day, and once 1-2 weeks later) 2 tasks, 2 cube conditions (with/without), 2 sides (right/left) = 8 experimental conditions and 2 videos of each	Between two lines at level of umbilicus (one where medial TrA is aligned with the lateral border of the image and the other 3 cm lateral) Foam cube used to control pressure and orientation Transducer repositioned between trials	Muscle thic ss Taken from ddle and 1 either side ween the in borders of muscle usi oftware Two, 7-seco videos fror t to contrac collected
Mannion et al[34]	**14 CLBP**; 7 male, 7 female; Age (years) 46 (9) **14 healthy**; 7 male, 7 female; Age (years) 31 (10)	CLBP 24.9 ± 2.4 Control 21.8 ± 2.6	5-12 MHz linear transducer M-mode	Voluntary activation Supine, AH, hips 30° 5 times right and left with 5 second hold, 1-2 minute rest between contractions Standardized instruction No feedback 1-2 weeks between measurement sessions	2.5 cm anteromedial to midpoint between iliac crest and costal margin on mid-axillary line Confirmed using B-mode ultrasound Foam cube used to control pressure	Muscle thic ss Leading ec oints mark regular inte s Images tak sec prior to during con on
Mannion et al[38]	**32 CLBP**; 11 male, 21 female; Age (years) 44 (12)	25.45 ± 4.1	5-12 MHz linear transducer M-mode	Voluntary and automatic activation (a) AH in hook lying. 10 x 5 second holds. 1-2 minute rest between (b) Rapid shoulder flexion (60°), abduction (60°) or extension (40°) in respons	2.5 cm anteromedial to midpoint between iliac crest and costal margin along mid axillary line Confirmed using B-mode ultrasound	Muscle thic ss Leading ec oints mark regular inte s. Software d to autom ally track b rs Taken at re ust prior to raction anc ng maxima

Study	Sample		Ultrasound Settings		Methodology	
	Demographics	BMI (kg/m2)	Transducer/frequency/mode	Position/Task	Transducer Placement	Measurer
				40°) in response to visual stimulus 10 movements each side Order of testing randomized No feedback		ng maxima ntraction
Ohe et al[48]	**30 CLBP**; 13 male, 17 female Age (years) 36 (9) **30 healthy**; 14 male, 16 female Age (years) 33 (9)	CLBP 21.3 ± 2.5 Control 21.2 ± 2.6	7 MHz linear transducer M-mode	Automatic activation Unilateral leg raising	2.5 cm anterior to midpoint between iliac crest and costal margin on mid axillary line Transducer placed in custom-made pad to minimize variability in angle and pressure Position and gain adjusted	Data collec was repeat times
Ota and Kaneoka[39]	**50 CLBP**; 36 male, 14 female; Age (years) 31.5 (9) **50 healthy**; 32 male, 18 female; Age (years) 30 (6)	Not reported	8 MHz transducer B-mode	Nil activation Supine, taken at end of expiration	2.5 cm anterior to axillary line at height of the umbilicus Minimum pressure	Muscle thic ss Calipers us o draw vert ine on scre
Pinto et al[41]	**30 CLBP**; 10 male, 20 female; Age (years) 43 (13.5) **30 healthy**; 17 male, 13 female; Age (years) 41 (11)	CLBP: 24.7 ± 3.3 Control: 24.6 ± 2.7	7.5 MHz linear transducer	Automatic activation Supine, arms crossed over chest, pre-selected wedges under pelvis or torso to simulate neutral sitting or flexed sitting, hips 50° and knees 90°, isometric knee flexion/extension up to 7.5% of body weight Right side only Order of testing randomized 3 flexion and extension efforts	On a line midway between inferior angle of the rib cage and iliac crest	Muscle thic ss Taken from ddle, and 1 either side ng grid ove age Average of easures us
Pinto et al[40]	**30 CLBP**; 10 male, 20 female; Age (years) 43 (13.5) **30 healthy**; 17 male, 13 female; Age (years) 41 (11)	Not reported	7.5 MHz linear transducer	Voluntary activation Supine, AH, hips and knees resting on pre-selected wedge. Wedge under pelvis or torso to simulate neutral or flexed sitting	On a line midway between rib cage and iliac crest approximately 10 cm from the midline	Muscle thic ss Taken from ddle and 1 either side ng grid ove age Average of easures us

Study	Sample		Ultrasound Settings	Right side only	Methodology	
	Demographics	BMI (kg/m2)	Transducer/frequency/mode	Position/Task Order of images randomized	Transducer Placement	Measure
				3 images taken at rest and maximum contraction		
Pulkovski et al[42]	**50 CLBP**; 18 male, 32 female; Age (years) 43 (12.5) **50 healthy**; 18 male, 32 female; Age (years) 43.5 (13)	CLBP: 26.0 ± 4.5 Control: 24.0 ± 4.3	5-12 MHz linear transducer M-mode	Voluntary activation Supine, AH No visual feedback 5 contraction images	Midway between costal margin and iliac crest along anterior axillary line Adjusted so TrA, EO and IO appeared parallel on screen Foam belt used to hold transducer in place	Muscle thic ss Leading ec oints mark regular inte s
Roddey et al[46]	**18 CLBP**; 5 male, 13 female; Age (years) 44 (7) **35 healthy**; 13 male, 22 female; Age (years) 39 (13)	Not reported	4-11 MHz linear transducer	Voluntary activation Supine, crook lying, AH Standardized instructions Taken twice at end of expiration and during contraction	Midway between iliac crest and ribs, approximately 2.5 cm from the side of the body Confirmed using B-mode ultrasound	Muscle thic ss
Unsgaard-Tøndel et al[43]	**87 CLBP**; 25 male, 62 female; Age (years) 41 (11)	24.7 ± 3.2	10 MHz linear transducer B-mode	Voluntary activation AH	Halfway between the 11th costal cartilage and iliac crest approximating TrA muscle fibers Adjusted so TrA and IO appeared on screen and the TrA aponeurosis appeared toward one side of the image Light pressure maintained Transducer was maintained in a constant position	Muscle thic ss and slid Thickness n within fas boundarie Slide take crolling vid o obtain di ce between ting and co cted positi
Vasseljen and Fladmark[44]	**109 CLBP**; 33 male, 76 female; Age (years) 40 (11)	24.5 ± 2.9	10 MHz linear transducer B-mode	Voluntary activation Supine, AH held for 5-10 sec Standardized instruction Taken twice	Halfway between 11th costal cartilage and the iliac crest approximating TrA muscle fibers Adjusted so EO, IO and TrA were visible and the TrA aponeurosis was seen on one side of the ultrasound image Examiner sat on right side of the participant	Muscle thic ss and slid Thickness n using on en calipers nd 2 cm lat to TrA apo osis betwe scial borde side hyper ic region Taken at re nd during c action Slide take m TrA apo osis at rest oint of ma



Study	Sample	Ultrasound Settings		Methodology	m contract

Notes: CLBP=chronic low back pain, M=male, F=female, BMI=body mass index, 4PTK=four point kneeling, AH=abdominal hollowing, ASIS=anterior superior iliac spine, PSIS=posterior superior iliac spine, TrA=transversus abdominis, EO=external oblique, IO=internal oblique, SLR=straight leg raise.

Ultrasound mode, transducer type and placement

All studies reported details of their RTUI methods. Six studies used brightness (B) mode,[35,39,43–45,47] four used motion (M) mode[34,38,42,48] and ten did not report the mode of ultrasound used.[11,13,31–33,36,37,40,41,46] Seventeen studies used a linear transducer,[11,31–34,36–38,40–48] making this the most common choice. One study used a curvilinear transducer[13] and two did not report this information.[35,39]

In 18 studies, the ribs and iliac crests were used as reference points between which the transducer was placed.[11,13,31,32,34–38,40–48] Three studies used the umbilicus as a measurement reference point for transducer positioning.[32,33,39] Other studies identified a palpable landmark while viewing the ultrasound image to confirm positioning of the transducer.[13,32–34,36,38,42–44,46,47]

Only seven articles reported on transducer pressure applied to the skin.[32–34,38,39,42,48] Three reported this subjectively as ‘gentle,’ ‘light,’ or ‘minimum.’[32,39,43] Others attempted to standardize pressure using a foam cube to house the transducer[33,34,38,48] and aligned the edge of the cube with a line on the transducer to assist control of position, tilt and pressure.[33] Only one study reported recording the transducer position to ensure accurate repositioning on a second trial.[36]

Frequency

Reported frequencies were inconsistent and ranged between 3 and 12 MHz across studies. Some reported single frequency values of 5 MHz[13,36,47] or 7 MHz.[11,35,40,41,45,48] Others reported ranges of 5–10 MHz[33] or 5–12 MHz.[34,38,42]

Participant positioning

Nine studies acquired images at rest in crook lying or in a non-standardized supine position.[31,32,35,37,39–42,46] Patient movements or tasks included abdominal hollowing in four point kneeling,[11] sitting,[45] crook lying,[34,38,44,46] or were not specified.[43,47] Automatic activation was assessed in some studies using isometric lower limb movements[13,32,36,40] or concentric upper and lower limb movements, including leg raises[33,48] and rapid shoulder movements.[38,43]

Operator skill level

The majority of included studies provided no information on the skill level or previous training of operators.[11,13,31,34–42,44,45,47,48] Of the four studies that included this information, two provided brief explanations, such as experienced[43] or intensively trained.[32] One study reported that reliability of their assessors had been previously demonstrated, however, the cited study was on participants without CLBP.[46] Only one study provided an in-depth explanation of a clinically applicable minimal training protocol for novice operators involving an initial training session, three months of hands-on training, and a final session to ensure a valid protocol.[33]

Measurement

The most common reference point for measuring muscle thickness on the ultrasound image was between fascial lines.[13,33,34,38,42–44,47] Three studies used on-screen calipers[11,39,44] for this purpose and four placed a grid over the image.[32,36,40,41] Four studies took thickness measures in the middle of the image and at set distances either side[33,36,40,41] and others reported taking measures at varying distances from the TrA aponeurosis.[31,32,37,44,45] In contrast, some studies calculated muscle slide during contraction. Measurement techniques for this measured the distance between TrA aponeuroses at rest and during contraction[13,44] or scrolled the M-mode video to obtain the change in thickness.[43]

Measurement reliability

Seven studies reported measurement reliability results in LBP populations[11,13,32–34,40,41] and only two were conducted with CLBP participants.[32,34] (Table 4). Intraclass correlation coefficients (ICC) for static thickness measures were 0.97[32] and 0.63–0.89.[34] Thickness changes during contraction ranged between 0.56–0.72[32] and 0.41–0.88.[34] Both studies found higher reliability for static measures of muscle thickness at rest compared to thickness change, changes over time or contraction ratios.[32,34]

Table 4 Measurement reliability results

Study	Population	Reliability	Measurement	Values	Risk of bias
Costa et al[32]	CLBP participants only	Intrarater reliability of single measure	ICC [2,1]; 95% CI; SEM [%]	TrA change: .56; .49 to .62; 15	High
		Intrarater reliability average of 2 measures	ICC [2,1]; 95% CI; SEM [mm]; SDC [mm]	TrA change over time: .31; .20 to .41; 24	
				TrA thickness: .97; .96 to .97; 0.04; 0.11	
		Intrarater reliability of adjusted average of 2 measures	ICC [2,1]; 95% CI; SDC [%]	TrA change: .72; .65 to .77; 41.6	
				TrA change over time: .44; .33 to .58; 66.5	
Critchley and Coutts[11]	Separate control subjects for pilot reliability study	Intrarater reliability	ICC; 95% CI [mm]	TrA thickness: .94; 4.3 to 4.6	High
	No demographic information reported				
Hides et al[13]	Mixed control and CLBP	Intrarater reliability – *10 random participants*	ICC [1,3]	TrA thickness and slide combined: .93 to .99	Results reported only therefore not assessed using the QAREL
			ICC [2,3]	TrA thickness: .50 to .81	
		Intersession reliability - *2-7 days between sessions. 20 random participants*		TrA slide: .87 to .91	
Lariviere et al[33]	Mixed control and CLBP	Intrarater reliability	Ø; SEM	TrA rest: .86 to .87; .3	Moderate
				TrA contraction: .83 to .84; .5 to .6	
		Interrater reliability	Ø; SEM	TrA % change: .70 to .72; 15.2 to 17.9	
				TrA rest: .78 to .82; .4	
				TrA contraction: .79 to .80; .6 to .7	
				TrA % change: .61 to .68; 16.6 to 18.9	
Mannion et al[34]	CLBP participants only	Intrarater reliability	ICC [3,1]; SEM [mm]; CV [%]	TrA rest: .63 to .89; .27 to .46; 7.2 to 11.5	High
				TrA contraction: .41 to .88; .41 to .78; 7.7 to 14.3	
				TrA contraction ratio: .28 to .80; .09 to .16; 6.0 to 11.6	
Pinto et al[41]	Mixed control and CLBP	Intrarater reliability	ICC [3,1]; 95% CI; SEM [%]	TrA (unclear): .92; .97 to .78; 5.7	Results reported only therefore not assessed using the QAREL
Pinto et al[40]	Mixed control and CLBP	Intrarater reliability	ICC [3,6]; 95% CI; SEM[%]	TrA change: .89; .76 to .96; 2.34	Results reported only therefore not assessed using the QAREL

Notes: ICC = intraclass correlation coefficient, 95% CI = 95% confidence interval, SEM = standard error of measurement, SDC = smallest detectable change, CV = coefficient of variation, Ø = dependability coefficient, TrA = transversus abdominis, CLBP = chronic low back pain.

Four studies reported combined results for CLBP and control participants [13,33,40,41] and one study reported reliability from image measurements of an asymptomatic pilot study population.[11] Statistics used to calculate and report reliability in the reviewed studies varied and included 95% confidence intervals,[11,32,40,41] standard error of measurement,[32–34,40,41] dependability coefficient,[33] and coefficient of variations.[34] Hence, a meta-analysis was not possible due to heterogeneity of statistical tests.

Discussion

The majority of studies using RTUI were of a moderate to high methodological quality, reported methodologies were inconsistent and very few studies reported the reliability of the method used. Methodological limitations included bias due to reproducibility of methods, appropriateness of statistics, and controlling of confounding factors, including blinding participants to the ultrasound monitor and controlling for BMI and other patient demographics.

Methodological consistency is important in clinical settings and research to ensure reliability and confident comparison of results within and between examiners, and between studies, especially considering studies include participants from varying subgroups and demographic profiles. Methodological inconsistencies throughout the reviewed literature limit the capacity for clinical interpretation and excluded meta-analysis. Although the need for TrA RTUI methodological consensus has previously been suggested,[30] further research is needed to determine if one method would be appropriate for all subgroups of the LBP population.

In particular in the CLBP population, ultrasound frequency may need to vary to accommodate for higher BMI and body fat percentage.[42,49] It is possible that the wide frequency ranges found in the included literature [32–34,38,42,46] were used to compensate for individual disparities. Taking the included literature and past recommendations into consideration, it is possible that a linear or curvilinear transducer sets between 5–10 MHz would achieve the best quality image in the CLBP population.[7,50] However, no research is currently available on the effects of frequency choice in individuals with excess subcutaneous fat when imaging TrA.

The descriptions of technique for transducer positioning were mostly non-specific or ambiguous. Landmarks used for transducer placement included bony or palpable landmarks,[11,13,31,32,34–38,40–48] visual landmarks such as the umbilicus [33,38,39] and theoretical landmarks, such as axillary lines.[11,34,35,38,39,42,47,48] Increased difficulties in palpation of landmarks in people with higher BMI [51] is likely to influence reliability of TrA RTUI in the CLBP population. It may explain the varied reliability of TrA reported in the review articles. For example, a pilot study using healthy participants found an ICC of 0.94 for TrA thickness measures [11] whereas a study using CLBP patients found a variation of 0.63–0.89 for the same measure.[34] B-mode ultrasound was also used to assist transducer placement.

Use of B-mode ultrasound for transducer positioning is likely to be most reproducible in a CLBP population, when used in combination with other techniques such as palpation or visualization of anatomical landmarks as it can allow for individual variations and alleviates the reliance solely on palpable landmarks. For example, placing the transducer at a standardized distance laterally from the umbilicus (e.g. 10 cm) and using B-mode ultrasound to confirm the position. Despite these inferences, no reliability data currently exist for CLBP populations to support this theory. Additionally, increased pressure controlled objectively should be placed through the transducer to compensate for increased subcutaneous tissue.

The results of this review indicate that TrA activation in CLBP is most reliably imaged using automatic activation tasks while providing visual or verbal feedback on the force of the contraction in order to produce smaller isometric contractions. Costa et al.[32] Hides et al.[13] and Ferreira et al.[36] all provide examples of such tools. It is possible that large uncontrolled automatic activation movements, such as a straight leg raise may be less reliable, as RTUI has been shown to be most effective at identifying low-level contractions up to 20% of maximal.[42] Voluntary activation tasks such as abdominal hollowing have shown to have high between trial variations due to patients' fluctuating ability to perform the task, which in turn affects reliability.[19] Compared to research settings, clinically automatic activation tasks cannot be easily applied, and therefore voluntary contraction in an appropriate, standardized position has been advocated as the most appropriate clinical method.[34]

This review identified limited evidence to determine the best method for patient positioning when imaging TrA in a CLBP population. It is anticipated that imaging in four point kneeling may be less reliable due to the difficulty in controlling lumbar lordosis, which is known to have an effect on TrA thickness.[41] More research is required to establish the most appropriate positioning for CLBP patients and provide conclusive recommendations for optimal and consistent scanning methods appropriate for both clinical and research settings.

It is clinically relevant to investigate muscle contraction ratios or the difference between resting and contracted thickness, compared to singular muscle thickness measures at rest.[41] These measures provide information on percentage of change which relates to muscle activation. This is especially relevant in CLBP populations where factors, such as age and BMI cause thicker TrA muscles at rest[24,35,41] meaning measurements are less comparable than a relative measure when looking at different patient subgroups. In addition, the extended duration of symptoms in CLBP can cause increased fatigability and persistent delayed muscle activation[35] making between trial variations more pronounced.

Only a few studies reported information on operator skill level[32,33,43,46] which is a common theme reported in past systematic reviews of RUSI methods and can be linked to lower methodological quality.[19,52] Although it has been shown that good intra and inter-rater reliability is possible between novice and experience raters,[53] evidence to support this assumption is currently insufficient for measurement in CLBP. Compared to other imaging gold standards, RUSI is highly operator-dependent and can affect reliability and validity of measurement, therefore, studies should include a description of operators experience and skill level to assist in comparisons across the literature.[19,54]

This review identified great variation in the types of TrA measures taken and where the measures were taken from on the image. Measures included thickness measures at rest and during contraction, contraction ratios, and muscle slide. Several studies vaguely reported taking measurements from between fascial lines,[13,33,34,38,42–44,47] or, using a grid over the ultrasound image[32,36,40,41] which leaves much room for interpretation. Measurements differ if they are taken from the upper or lower fascial borders introducing measurement error and false positives or negatives when assessing the effect of intervention or exercise. Future studies should use and report consistent measuring points to standardize their methods. Grids also risk measurement error depending on how and where it is placed. Studies using a manual or electronic grid reported a higher variation in ICC results (0.56–0.92),[32,40,41] while those using other measurement techniques did not (ICC = 0.83–0.94)[11,13,33] and is therefore not recommended.

It has been shown that averaging a single measure across three images improves reliability, whereas, averaging three parallel measures from one image has little effect.[33] Therefore the findings of this review suggest that in CLBP patients, TrA thickness and activation should be measured using consistent reference points, using the on-screen image to confirm transducer position and measure an average across single measures from three images.

Overall reliability was poorly and inconsistently reported. The two studies reporting CLBP population reliability[19,38] used different methods and statistics making it impossible to directly compare their results. The four studies that reported reliability for mixed healthy and CLBP populations[13,33,40,41] also used a variety of statistics to report their results, including different ICC models. The majority of ICC scores reported for both intra and inter-rater reliability of TrA measurement suggest excellent reliability (ICC > 0.75).[55] Despite this, only two studies in the literature report results specifically for the CLBP population, both used different methods for image acquisition and statistical analysis and both demonstrated a high risk of bias through poor reporting of appropriateness of raters and blinding.[32,33] Hence, the results must be interpreted with caution.

Conclusion

This review demonstrated that there has been a lack of methodological consistency between studies using RTUI to measure TrA in CLBP populations. This is despite the fact that individually much of the evidence is of moderate to high quality. In addition, reliability was poorly and inconsistently reported and appraisal demonstrated a moderate to high risk of bias. This highlights the insufficiency of current research to accurately establish reliable methodologies for RTUI measurement of TrA in the chosen population. RTUI is routinely used in clinical practice by physiotherapists to measure TrA, therefore, it is important for clinicians to acknowledge the limitations of current evidence when applied to CLBP and for researchers to understand the need to establish methodological standardization in future studies.

Funding

No external funding was received for this project

Disclosure statement

The authors know of no existing conflicts of interest.

References

1 Arab AM , Rasouli O , Amiri M , Tahan N . Reliability of ultrasound measurement of automatic activity of the abdominal muscle in participants with and without chronic low back pain. Chiropr Man Therap. 2013;21(37):1–7.

2 Hides JA , Stokes MJ , Saide M , Jull GA , Cooper DH . Evidence of lumbar multifidus muscle wasting ipsilateral to symptoms in patients with acute/subacute low back pain. Spine. 1994;19(2):165–72.

AQ1 patients with acute/subacute low back pain. Spine. 1994;19(2):165–72.

3 Lariviere C , Gagnon D , De Oliveira E Jr, Henry SM , Mecheri H , Dumas JP . Ultrasound measures of the lumbar multifidus: effect of task and transducer position on reliability. PM&R. 2013;5(8):678–87.

4 Gnat R , Saulicz E , Miądowicz B . Reliability of real-time ultrasound measurement of transversus abdominis thickness in healthy trained subjects. Eur Spine J. 2012;21:1508–15.

5 Hides J , Wilson S , Stanton W , McMahon S , Keto H , McMahon K , et al. An MRI investigation into the function of the transversus abdominis muscle during ‘drawing-in’ of the abdominal wall. Spine. 2006;15(31):175–8.

6 Koppenhaver SL , Hebert JJ , Fritz JM , Parent EC , Teyhen DS , Magel JS . Reliability of rehabilitative ultrasound imaging of the transversus abdominis and lumbar multifidus muscles. Arch Phys Med Rehabil. 2009;90(1):87–94.

7 McMeeken JM , Beith ID , Newham DJ , Milligan P , Critchley DJ . The relationship between EMG and change in thickness or transversus abdominis. Clin Biomechan. 2004;19(4):337–42.

8 Saliba S , Croy T , Guthrie R , Grooms D , Weltman A , Grindstaff T . Differences in transverse abdominis activation with AQ2 stable and unstable bridging exercises in individuals with low back pain. N Am J Sports Phys Ther. 2010;5(2):63-73.

9 Rasouli O , Arab AM , Amiri M , Jaberzadeh S . Ultrasound measurement of deep abdominal muscle activity in sitting positions with different stability levels in subjects with and without chronic low back pain. Man Ther. 2011;16(4):388–93.

10 Hyde J , Stanton WR , Hides JA . Abdominal muscle response to a simulated weight-bearing task by elite Australian Rules football players. Hum Mov Sci. 2012;31:129–38.

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

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

13 Hides JA , Belavý DL , Cassar L , Williams M , Wilson SJ , Richardson CA . Altered response of the anterolateral abdominal muscles to simulated weight-bearing in subjects with low back pain. Eur Spine J. 2009;18(3):410–8.

14 Kiesel KB , Underwood FB , Mattacola CG , Nitz AJ , Malone TR . 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. 2007;37(10):596–607.

15 Geertzen JHB , Wilgen CPV , Schrier E , Dijkstra PU . Chronic pain in rehabilitation medicine. Disabil Rehabil. 2006;28(6):363–7.

16 Teyhen DS , Miltenberger CE , Deiters HM , Del Toro YM , Pulliam JN , Childs JD , et al. The use of ultrasound imaging of the abdominal drawing-in maneuver in subjects with low back pain. J Orthop Sports Phys Ther. 2005;35(6):346–55.

17 Koppenhaver SL , Hebert JJ , Parent EC , Fritz JM . Rehabilitative ultrasound imaging is a valid measure of trunk muscle size and activation during most isometric sub-maximal contractions: a systematic review. Aust J Physiother. 2009;55(3):153–69.

18 Hebert JJ , Koppenhaver SL , Parent EC , Fritz JM . A systematic review of the reliability of rehabilitation ultrasound imaging for the quantitative assessment of the abdominal and lumbar trunk muscles. Spine. 2009;34(23):848–56.

19 Costa LOP , Maher CG , Latimer J , Smeets RJEM . Reproducibility of rehabilitative ultrasound imaging for the measurement of abdominal muscle activity: a systematic review. Phys Ther. 2009;89(8):756–69.

20 Hides JA , Richardson CA , Jull GA . Magnetic resonance imaging and ultrasonography of the lumbar multifidus muscle: comparison of two different modalities. Spine. 1995;20(1):54–8.

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

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

23 Rankin G , Stokes M , Newham DJ . Abdominal muscle size and symmetry in normal subjects. Muscle Nerve. 2006;34(3):320–6.

24 Springer BA , Mielcarek BJ , Nesfield TK , Teyhen DS . Relationships among lateral abdominal muscles, gender, body mass index, and hand dominance. J Orthop Sports Phys Ther. 2006;36(5):289–97.

25 Miller JC , Lee SI [Internet]. Radiology rounds: imaging and obese patients. 2005 Jul. [cited 2014 Jun 25]. Available from: http://www.mghradrounds.org/index.php?src=gendocs&link=2005_july

26 Moher D , Liberati A , Tetzlaff J , Altman DG . Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol. 2009;62:1006–12.

27 Kmet LM , Lee RC , Cook LS [Internet]. Standard quality assessment criteria evaluating primary research papers from a variety of fields. 2004 Feb 1. [cited 2014 Jun 25]. Available from: http://www.biomedcentral.com/content/supplementary/1471-2393-14-52-S2.pdf

28 Lucas NP , Macaskill P , Irwig L , Bogduk N . The development of a quality appraisal tool for studies of diagnostic reliability (QAREL). J Clin Epidemiol. 2010;63(8):854–61.

29 Walburn J , Vedhara K , Hankins M , Rixon L , Weinman J . Psychological stress and wound healing in humans: A systematic review and meta-analysis. J Psychosom Res. 2009;67:253–71.

30 Flavell CA , Gordon S , Marshman L , Watt K . Inter-rater reliability of classification systems in chronic low back pain
AQ3 populations: a systematic review. Phys Ther Rev. 2014;19(3):204-12.

31 Cho SH , Kim KH , Baek IH , Goo BO . Comparison of contraction rates of abdominal muscles of chronic low back pain patients in different postures. J Phys Ther Sci. 2013;25(8):907–9.

32 Costa LOP , Maher CG , Latimer J , Hodges PW , Shirley D . An investigation of the reproducibility of ultrasound measures of abdominal muscle activation in patients with chronic non-specific low back pain. Eur Spine J. 2009;18(7):1059–65.

33 Larivière C , Gagnon D , De Oliveira E Jr, Henry SM , Mecheri H , Dumas JP . Reliability of ultrasound measures of the transversus abdominis: effect of task and transducer position. PM & R. 2013;5(2):104–13.

34 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. Eur Spine J. 2008;17(4):494–501.

35 Akbari A , Khorashadizadeh S , Abdi G . The effect of motor control exercise versus general exercise on lumbar local stabilizing muscles thickness: randomized controlled trial of patients with chronic low back pain. J Back Musculoskelet Rehabil. 2008;21(2):105–12.

36 Ferreira PH , Ferreira ML , Maher CG , Refshauge K , Herbert RD , Hodges PW . Changes in recruitment of transversus abdominis correlate with disability in people with chronic low back pain. Br J Sports Med. 2010;44(16):1166–72.

37 Kim KH , Cho SH , Goo BO , Baek IH . Differences in transversus abdominis muscle function between chronic low back pain patients and healthy subjects at maximum expiration: measurement with real-time ultrasonography. J Phys Ther Sci. 2013;25(7):861–3.

38 Mannion AF , Caporaso F , Pulkovski N , Sprott H . Spine stabilisation exercises in the treatment of chronic low back pain: a good clinical outcome is not associated with improved abdominal muscle function. Eur Spine J. 2012;21(7):1301–10.

39 Ota M , Kaneoka K . Differences in abdominal muscle thicknesses between chronic low back pain patients and healthy subjects. J Phys Ther Sci. 2011;23(6):855–8.

40 Pinto RZ , Ferreira PH , Franco MR , Ferreira MC , Ferreira ML , Teixeira-Salmela LF , et al. The effect of lumbar posture on abdominal muscle thickness during an isometric leg task in people with and without non-specific low back pain. Man Ther. 2011;16(6):578–84.

41 Pinto RZ , Ferreira PH , Franco MR , Ferreira ML , Ferreira MC , Teixeira-Salmela LF , et al. Effect of 2 lumbar spine postures on transversus abdominis muscle thickness during a voluntary contraction in people with and without low back pain. J Manipulative Physiol Ther. 2011;34(3):164–72.

42 Pulkovski N , Mannion AF , Caporaso F , Toma V , Gubler D , Helbling D , et al. Ultrasound assessment of transversus abdominis muscle contraction ratio during abdominal hollowing: a useful tool to distinguish between patients with chronic low back pain and healthy controls? Eur Spine J. 2012;21(SUPPL. 6):S750–S9.

43 Unsgaard-Tøndel M , Lund Nilsen TI , Magnussen J , Vasseljen O . Is activation of transversus abdominis and obliquus internus abdominis associated with long-term changes in chronic low back pain? A prospective study with 1-year follow-up. Br J Sports Med. 2012;46(10):729–34.

44 Vasseljen O , Fladmark AM . Abdominal muscle contraction thickness and function after specific and general exercises: A randomized controlled trial in chronic low back pain patients. Man Ther. 2010;15(5):482–9.

45 Kim HI , Kim SY , Kim TY . Comparison of changes in abdominal muscle thickness using ultrasound imaging during the abdominal drawing-in maneuver performed by patients with low back pain and healthy subjects. J Phys Ther Sci. 2012;24(5):383–5.

46 Roddey TS , Brizzolara kJ , Cook KF . Side-to-side differences in the transverse abdominus muscle measured by real time ultrasound in persons with and without chronic low back pain. OPTP. 2008;20(2):56–9.

47 Huang Q , Li D , Zhang J , Yang D , Huo M , Maruyama H . Comparison of the efficacy of different long-term interventions on chronic low back pain using the cross-sectional area of the multifidus muscle and the thickness of the transversus abdominis muscle as evaluation indicators. J Phys Ther Sci. 2014;26(12):1851–4.

48 Ohe A , Kimura T , Goh A-C , Oba A , Takahashi J , Mogami Y . Characteristics of trunk control during crook-lying unilateral leg raising in different types of chronic low back pain patients. Spine. 2015;40(8):550–9.

49 Mengiardi B , Schmid MR , Boos N , Pfirrmann CWA , Brunner F , Elfering A , et al. Fat content of lumbar paraspinal muscles in patients with chronic low back pain and in asymptomatic volunteers: quantification with MR spectroscopy. Radiology. 2006;240(3):786–92.

50 Teyhen DS , Gill NW , Whittaker JL , Henry SM , Hides JA , Hodges P . Rehabilitative ultrasound imaging of the abdominal muscles. J Ortho Sports Phys Ther. 2007;37(8):450–66.

51 Snider KT , Snider EJ , Degenhardt BF , Johnson JC , Kribs JW . Palpatory accuracy of lumbar spinous processes using multiple bony landmarks. J Manipulative Physiol Ther. 2011;34(5):306–13.

52 Kwah LK , Pinto RZ , Diong J , Herbert RD . Reliability and validity of ultrasound measurements of muscle fascicle length and pennation in humans: a systematic review. J Appl Physiol. 2013;114(6):761–9.

53 Wallwork TL , Hides JA , Stanton WR . Intrarater and interrater reliability of assessment of lumbar multifidus muscle thickness using rehabilitative ultrasound imaging. J Orthop Sports Phys Ther. 2007;37(10):608–12.

54 Whittaker JL , Teyhen DS , Elliott JM , Cook K , Langevin HM , Dahl HH , et al. Rehabilitative ultrasound imaging: understanding the technology and its applications. J Orthop Sports Phys Ther. 2007;37(8):434–49.

55 Hallgren KA . Computing inter-rater reliability for observational data: an overview and tutorial. Tutor Quant Methods Psychol. 2012;8(1):23–4.