

Assessment of pelvic floor and abdominal muscles three months postpartum: A reliability study

Sabine Vesting (✉ Sabine.vesting@gu.se)

Goteborgs universitet Institutionen for neurovetenskap och fysiologi <https://orcid.org/0000-0002-6523-1833>

Monika Fagevik Olsén

Goteborgs universitet Institutionen for neurovetenskap och fysiologi

Annelie Gutke

Goteborgs universitet Institutionen for neurovetenskap och fysiologi

Gun Rembeck

Goteborgs universitet Institutionen for medicin

Maria EH Larsson

Goteborgs universitet Institutionen for neurovetenskap och fysiologi

Research article

Keywords: Postpartum, Assessment, Reliability, Inter-rater reliability, Pelvic floor muscle, Diastasis recti abdominis

Posted Date: June 12th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-34672/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background Pregnancy and childbirth often result in alterations of core muscles, and women may require postpartum assessment of pelvic floor muscle function and abdominal wall changes, e.g. diastasis recti abdominis (DRA). However, there is currently no gold standard for postpartum assessment of these muscles' function. Here we aimed to evaluate the reliability of clinically applicable methods for assessing pelvic floor muscles and DRA after pregnancy. Methods We recruited 222 postpartum women from Swedish antenatal and childbirth centers, and via social media. Pelvic floor and DRA assessment were performed via observation and palpation at three rehabilitation centers in Sweden. At each center, two independent physiotherapists performed the assessments in random order. Results Assessment of the maximal voluntary contraction and pelvic floor muscle endurance revealed kappa values ranging from 0.49–0.69. Assessments of voluntary contraction by observation, involuntary contraction, and voluntary relaxation yielded inconsistent results, with slight-to-moderate weighted kappa values ranging from 0.10–0.51. DRA assessment by caliper yielded ICC values of 0.73–0.83 after physiotherapists underwent 2 months of training in applying this assessment method. The standard error of measurements for this method was about 4 mm, and the minimal detectable change was 12 mm. Assessments of DRA depth and bulging showed moderate kappa values of 0.43–0.51, with reservation for some inconsistency between the centers. Conclusions Palpation of maximal voluntary contraction and pelvic floor muscle endurance are reliable postpartum assessment methods. With some experience and training, a caliper is a reliable instrument for assessing the postpartum DRA width. Additional research is needed to identify a reliable assessment method for pelvic floor muscle functions other than voluntary contraction, and for DRA depth and bulging.

Background

The pelvic floor and abdominal muscles are parts of the human core, i.e., the anatomic and functional center of the human body [1]. Their function and activation play various roles in securing spine stability [2–4], posture [3], and continence [5, 6]. In healthy women, these muscles work together in voluntary or reflexive co-contractions [7]. During pregnancy and childbirth, these muscles are greatly stretched, resulting in postpartum muscular alterations [8, 9], which can lead to feelings of insecurity.

Women are increasingly seeking help and advice regarding postpartum core muscle changes from physiotherapists at primary health care centers. One recent publication reported that 12% of women seek emergency care during the early weeks after giving birth, primarily due to pelvic floor problems [10]. Childbirth is associated with pelvic floor traumas, such as perineal tears and levator ani injuries, which can lead to incontinence, pelvic organ prolapse, and decreased quality of life [11, 12]. Approximately one-third of women experience persistent stress urinary incontinence after their first delivery. Another postpartum concern is a persistent separation of the two parts of the rectus abdominis, termed a diastasis recti abdominis (DRA). At 12 months postpartum, 33% of women exhibit a space between these muscles greater than the width of two fingers [13]. A DRA is reportedly correlated with impaired quality of life, negative body image, and abdominal pain [14, 15].

Physiotherapists who manage women's health use various methods to assess the pelvic floor and DRA after pregnancy [16, 17]; however, there is currently no gold standard. Pelvic floor muscles (PFM) can be assessed by observation and by digital palpation, defined as "the process of using fingers/hands as part of assessment, to gather information about the tissues" [18]. The PFM can be involuntarily and voluntarily contracted, and can also be voluntarily relaxed [19]. These functions are defined by the international society of incontinence [19]; however, there are not yet any standardized rating scales or other clear outcome measures. One study reported the use of a Delphi scheme to identify the optimal protocol for assessing these functions and tested their inter-rater reliability [20], but these assessments were not performed in postpartum women.

Ultrasound assessment is the most reliable and valid method for DRA measurement [21]. However, most women who are concerned about their DRA seek help at primary healthcare centers, where ultrasound is seldom available. About 96% of American physiotherapists specialized in women's health assess the DRA using the finger-width method [13, 22], which is

Loading [MathJax]/jax/output/CommonHTML/jax.js

and has weaker inter-rater reliability than instrumental assessment methods

[21]. Less than 2% of American physiotherapists use calipers for the assessment of postpartum women [22]. DRA assessment using a caliper is reported to be nearly as accurate as ultrasound assessment [21, 24], although the inter-rater reliability of this method has not yet been tested.

One experimental study shows that the tendon between the two parts of the rectus abdominis—the linea alba—exhibits a deviant behavior in a curl-up movement in women with DRA [25], and these findings are strengthened by another similar study [26]. Both research groups argue that the DRA depth and bulging are more relevant than the width. These studies were performed using assessment by ultrasound [25] and shear-wave elastography [26]. However, there are currently no validated and clinically applicable assessment methods or rating scales for the parameters of depth and bulging of the linea alba.

In the present study, we aimed to evaluate the reliability of different aspects of the clinical assessment of PFM and DRA using observation, calipers, and digital palpation at 3 months postpartum.

Methods

This study included 222 women from the Region Västra Götaland, Sweden. Assessments were conducted at three rehabilitation centers. Based on the guidelines of Ko & Li, we aimed to assess at least 30 participants at each center [27]. The women were invited to participate at antenatal and childcare centers, and via social media. Inclusion criteria were age of ≥ 18 years, vaginal delivery or caesarean section within the past 3 months, and ability to understand and respond in Swedish. Exclusion criteria were chronic pelvic girdle pain and/or low back pain (defined as pelvic or low back pain for over 3 months, not related to pregnancy) and/or pelvic floor tear grade III/IV.

The participants were contacted and booked for assessment at one of the three rehabilitation centers in the Västra Götaland region within 3 months after giving birth. Prior to the assessments, the participants completed a questionnaire about their age, BMI, mode of delivery, number of delivered children, self-reported pelvic floor tears, most recent baby's birth weight, and the birth weights of previous children (if applicable).

Assessments were performed by six physiotherapists—two at each rehabilitation center. These physiotherapists had each completed a four-day (or longer) course in PFM assessment and treatment methods, and all had between one and nine years of experience in assessing PFM. During the design phase of this study, four hours of training in DRA measurement was planned. All included physiotherapists were novices at measuring the DRA by caliper, and on using the rating scales for depth and bulging. Two months after the start of the study, we conducted a preliminary data analysis because the physiotherapists expressed strong uncertainty regarding the right technique for using the caliper. This preliminary analysis showed low-to-negative ICC values and large differences between the measurements. Thus, the 61 measurements acquired between September and November of 2018 were excluded from the final analysis. The physiotherapists at all centers underwent additional training. At this time, rehabilitation center 3 had not yet started their assessments.

Clinical assessment of PFM

The PFM was assessed with the patient in the supine position, with the legs flexed and slightly abducted on a plinth, and a pillow under the head. Participants were assessed by observation and digital palpation.

During observation, the physiotherapist stood beside the plinth, holding the participant's legs and observing the movement of the perineum. To observe involuntary contraction, the participant was asked to cough forcefully, and the physiotherapist rated the movement as moving downwards, perineal in-drawing, or no movement. To observe voluntary contraction, the participant was given the verbal cue "contract your pelvic floor muscles like you want to prevent the escape of gas/urine". The physiotherapist then observed the movement of the perineum, and rated it as moving downwards, perineal in-drawing, or no movement.

Digital palpation of the PFM was performed by physiotherapists using their index and middle finger, with examinations
Loaders were inserted 2–3 cm into the vagina, with the palmar side directed to the

caudal part of the vagina. To assess involuntary contraction, the participant was asked to forcefully cough three times. The physiotherapist noted the absence or presence of a correct contraction, defined as a squeeze around the pelvic openings and an inward lift [17].

To assess maximal voluntary contraction (MVC), the participant was asked to contract the PFM. In the event of a downward movement, the participant was again given the verbal cue “contract your pelvic floor muscles like you want to prevent the escape of gas/urine”. If the physiotherapist felt a correct contraction, the participant was encouraged to activate their PFM “as strong and as long you can”. Of three MVCs, the strongest was rated on a 6-point modified oxford scale (Appendix 1). The participants rested 15 seconds between the contractions. If a participant, despite several attempts and verbal cues, failed to squeeze and lift and was instead straining, their PFM function was rated as “-1” and the participant was excluded from the statistical analysis of MVC and PFM endurance.

To assess PFM endurance by digital palpation, after 15 seconds of rest, the participant was asked to contract the PFM for as long as possible at approximately 50% of the previous contraction strength. The physiotherapist rated PFM endurance as positive if the participant was able to hold this contraction for longer than 30 seconds. Finally, to assess voluntary relaxation, the participant was given the verbal cue “try to relax your pelvic floor, let the vagina get larger and go downwards”. This function was rated as absent, partial, or complete.

Clinical assessment of DRA

DRA assessment was conducted in the same position as described above for PFM. The physiotherapists assessed DRA width using an electronic digital caliper (150 mm, carbon fiber, accuracy ± 0.2 mm, 24 se Sverige AB, Kalmar, Sweden). Caliper application is explained in Appendix 2. At the start, the physiotherapist used a water-soluble marker to mark the three measurement points: at the umbilicus, and at 4.5 cm above and 4.5 cm below the umbilicus [28, 29]. For accurate assessment, the participant had to lift her head 2–3 cm from the plinth, with no pillow. Before the assessment began, the physiotherapist assured that the participant correctly lifted her head 2–3 cm, which was trained by several repetitions.

To assess DRA width, the participant was asked to lift her head and then lower it slowly. During this movement, the physiotherapist palpated the outer edges of the linea alba with their index and middle finger, without examination gloves. Next, the participant was instructed to relax her muscles and perform the trained head lift of 2–3 cm. During this movement, the physiotherapist identified the distance between the two parts of the rectus abdominis with her fingers, and measured this felt distance using the caliper (Fig. 2a). The same procedure was conducted at all three measurement points.

To measure the linea alba depth, the participant was asked to repeat the exact same head lift of 2–3 cm. At all three measurement points, the physiotherapist palpated the resistance, and rated it as “good resistance at all points”, “resistance in the depth at measurement point x”, or “bottomless resistance at measurement point x”. To assess linea alba bulging under load, the participant performed a 3-step sit-up test [30]. During this test, the physiotherapist observed whether the linea alba bulged during the movement (Fig. 2b).

Upon completion of the assessment, the participants rested for 30 minutes in a sitting or lying position. After the 30-minute rest, the second physiotherapist conducted the same assessment as described above. The two investigating physiotherapists were blinded to each other’s findings, and were not allowed to talk about their assessments.

Statistical analysis

Statistical analyses were performed using IBM SPSS statistical package version 25 (SPSS Inc., Chicago, IL) and the Svensson Excel template from <http://avdic.se/svenssonsmetod.html>. Descriptive statistics are presented as mean, standard deviation (SD), and range for ratio data, and as number and percentage for nominal and ordinal data. To calculate statistically significant differences between the three rehabilitation centers, we used the one-way ANOVA test for interval and ratio data, and the Kruskal-Wallis test for ordinal data. A p value of ≤ 0.05 was regarded as statistically significant.

All PFM functional measures were rated on ordinal scales, except for PFM endurance and involuntary contraction by palpation. DRA depth was also rated on an ordinal scale. Ratings on ordinal scales were evaluated by Cohen's weighted kappa values. PFM endurance, involuntary contraction by palpation, and linea alba bulging were rated on nominal scales, and these ratings were evaluated by Cohen's kappa values. For interpretation of kappa values, we used the categories of Landis and Koch: <0.2, slight; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, substantial; and 0.81–1.0, almost perfect agreement [31]. Percentage agreement was calculated and presented for all nominal and ordinal data, and < 60% agreement was defined as faulty agreement [32].

We used the Svensson method, developed by Elisabeth Svensson [33], to distinguish the position and concentration variance between physiotherapists. Position variance was defined by the fact that one physiotherapist is systematically using a higher or lower value on a rating scale than the other physiotherapist. Concentration variance was defined by the fact that one physiotherapist was systematically using a smaller part of the scale. Relative rank variance explains an individual variation that cannot be explained by a systematic bias.

For assessment of DRA width, a continuous scale (in mm) was used. To evaluate the inter-rater reliability of the assessments on a continuous scale, we calculated the intraclass correlation coefficient (ICC) and 95% confidence interval (CI). ICC values were calculated in SPSS based on absolute agreement and a 2-way mixed effects model. ICC values of < 0.50 indicate poor reliability, 0.50–0.75 indicate moderate reliability, 0.75–0.90 good reliability, and values of > 0.90 indicate excellent reliability [27].

To further evaluate reliability, we calculated the standard error of measurements ($SEM = SD \times \sqrt{1 - ICC}$), which represent the typical error in a single measurement and the minimal detectable change ($MDC = SEM \times 1.96 \times \sqrt{2}$). For calculation of the standard error of measurements (SEM), we used the standard deviation (SD) from the scores of all subjects. SEM and minimal detectable change values are presented in mm.

Results

A total of 222 women were assessed, with measurements conducted from September 2018 through February 2020. Table 1 presents the participants' characteristics. The participating women at rehabilitation center 3 were significantly younger than the women at rehabilitation center 2, and had significantly more children and more vaginal deliveries than the participants at rehabilitation centers 1 and 2.

Table 1
 Characteristics of the 222 participating women at three months postpartum

	Total (n = 222)				Rehabilitation center 1 (n = 90)				Rehabilitation center 2 (n = 103)				Rehabilitation center 3 (n = 29)			
Age in years	33.1 (± 3.3)				32.6 (± 3.5)				33.8 (± 2.9)				32.0 (± 3.8)			
m (± SD)	24–42				24–42				27–40				25–39			
range																
BMI	24.5 (± 3.0)				24.4 (± 3.1)				24.3 (± 2.9)				25.1 (± 3.2)			
mean (± SD)	17–34				19–34				17–31				21–32			
range																
Delivery mode	28 (13)				9 (10)				16 (16)				3 (10)			
n (%)	194 (87)				81 (90)				87(85)				26 (90)			
C-section																
Vaginal delivery																
Number of children	1	2	3	>3	1	2	3	>3	1	2	3	>3	1	2	3	>3
n (%)	137	74	9	3	51	33	7	0	75	26	1	1	11	15	1	2
	(61)	(33)	(4)	(1)	(57)	(37)	(6)		(73)	(25)	(1)	(1)	(37)	(50)	(3)	(7)
Self-reported pelvic floor tear																
n (%)																
No tear	54 (24)				24 (27)				23 (22)				7 (24)			
First-degree perineal tear	52 (23)				16 (18)				24 (23)				12 (41)			
Second-degree perineal tear /episiotomy	74 (33)				33 (38)				35 (34)				6 (20.0)			

	Total (n = 222)	Rehabilitation center 1 (n = 90)	Rehabilitation center 2 (n = 103)	Rehabilitation center 3 (n = 29)
Neonatal birth weight	3574.5 (± 507.0)	3517.7 (± 533.7)	3604.7 (± 489.2)	3641.1 (± 484.0)
mean (± SD)	730–4725	730–4500	2235–4725	2592–4682
range				
<i>n</i> , number; BMI, body mass index; m, mean; SD, standard deviation.				

Assessment of PFM

Table 2 presents the results of PFM assessment. The evaluation of MVC showed substantial agreement (weighted kappa value, 0.69), and assessment of PFM endurance showed moderate agreement (kappa value, 0.49). Seven participants (3.3%) were excluded from the analyses of MVC and PFM endurance due to incorrect PFM contraction (straining). We found that the three rehabilitation centers significantly differed in their application of the modified Oxford scale. Rehabilitation center 1 did not use the full scale for MVC assessment, and only one participant was rated as higher than 3 on the modified Oxford scale. Assessment of voluntary contraction by observation showed moderate agreement, with a weighted kappa of 0.45. Among all assessments, about 89% were rated as “perineal inward movement”

The assessment of involuntary contraction by observation exhibited slight-to-fair weighted kappa values. About 70% of participants were rated as “downward movement”, and 9–11% as upward movement. Fair-to-moderate kappa values were found for evaluation of involuntary contraction by palpation. Over 80% of participants were rated as “absence of correct contraction”.

Ratings of voluntary relaxation showed large variations between different rehabilitation centers, with weighted kappa values ranging from – 0.08 to 0.56. The application of the scale significantly differed between rehabilitation center 3 and rehabilitation centers 1 + 2. At rehabilitation center 3, the physiotherapists rated 25 of 29 assessments as showing complete voluntary relaxation. In contrast, at rehabilitation centers 1 + 2, 10–12% of participants were rated as absent voluntary relaxation, 66% as partly relaxed, and 20–24% as complete voluntary relaxation. Rehabilitation center 2 showed a negative kappa value of – 0.08, indicating an agreement worse than expected or no agreement [32].

Comparing all centers with the specific centers revealed some deviant results. The Svensson method showed the following findings. For assessment of PFM endurance, rehabilitation center 2 showed a lower agreement (kappa value, 0.31), and a position variance was found (0.23 [95% CI: 0.12; 0.33]). For voluntary contraction by observation, we found fair agreement, with weighted kappa values of 0.29–0.40 at rehabilitation center 1 + 3, and we detected no position or concentration variance. With regards to involuntary contraction by observation, we found the lowest weighted kappa value (0.11) at rehabilitation center 2. Moreover, we identified a position variance (– 0.36 [95% CI: –0.46; –0.24]) and a concentration variance (0.15 [95% CI: 0.03; 0.29]). Assessment of involuntary contraction by palpation showed deviant results at rehabilitation center 1 (kappa value, 0.26), and a position variance was found (– 0.09 [95% CI: –0.16; –0.02]). For voluntary relaxation by palpation, rehabilitation center 2 exhibited a negative weighted kappa value. We found a position variance (– 0.31, [95% CI: –0.42; –0.19]) and a concentration variance (0.10 [95% CI: 0.01; 0.34]). The Svensson method was also used to analyze whether there was a possible learning or fatigue effect between the first and second assessments. We found a

position variance for assessment of MVC and PFM endurance, with the second assessment showing higher values than the first: MVC, 0.04 [95% CI: <0.01; 0.10]; and PFM endurance, 0.08 [95% CI: 0.01; 0.15].

Table 2
Results of PFM assessment

Parameters	Total group (<i>n</i> = 222)			Rehabilitation center 1 (<i>n</i> = 90)			Rehabilitation center 2 (<i>n</i> = 103)			Rehabilitation center 3 (<i>n</i> = 29)		
	Kappa	95% CI	PA %	Kappa	95% CI	PA %	Kappa	95% CI	PA %	Kappa	95% CI	PA %
Voluntary contraction												
Observation	0.45	0.28; 0.62	90	0.40	0.11; 0.70	91	0.55	0.33; 0.77	90	0.29	-0.13; 0.70	86
Palpation	0.69	0.62; 0.76	71	0.70	0.58; 0.82	77	0.59	0.46; 0.71	67	0.67	0.50; 0.84	62
MVC*												
PFM endurance*	0.49		74	0.69		84	0.31		65	0.55		83
Involuntary contraction												
Observation	0.10	-0.02; 0.22	57	0.20	-0.01; 0.42	77	0.11	-0.05; 0.27	48	0.38	0.10; 0.67	68
Palpation	0.51		85	0.26		87	0.47		85	0.47		75
Voluntary relaxation												
Palpation	0.26	0.15; 0.37	57	0.30	0.16; 0.50	63	-0.08	-0.23; 0.07	45	0.56	0.07; 1.02	89
<i>n</i> , number; CI, confidence interval; PA, percentage agreement; MVC, maximal voluntary contraction; PFM, pelvic floor muscle.												
* reduced number of participants due to incorrect PFM contraction (straining): total group (<i>n</i> = 215), rehabilitation center 1 (<i>n</i> = 88) rehabilitation center 2 (<i>n</i> = 98), rehabilitation center 3 (<i>n</i> = 29).												

Assessment of DRA

DRA width, depth, and bulging were assessed in 159 women. Table 3 presents the measured DRA widths, which were significantly wider at rehabilitation center 3 compared to at rehabilitation centers 1 and 2.

Table 3
Width of diastasis recti abdominis (DRA) at 3 months postpartum (in mm) measured with a caliper

	Total group (<i>n</i> = 159)		Rehabilitation center 1 (<i>n</i> = 61)		Rehabilitation center 2 (<i>n</i> = 69)		Rehabilitation center 3 (<i>n</i> = 29)		<i>p</i> value
	m	SD	m	SD	m	SD	m	SD	
Width at 4.5 cm above the umbilicus	22.0 [20.9; 23.2]	7.4	19.8 [18.5; 21.2]	5.3	20.6 [19.5; 21.7]	4.6	29.9 [25.8; 34.1]	10.9	< 0.01
Width at the umbilicus	25.9 [24.8; 27.1]	7.2	24.1 [22.5; 25.8]	6.5	24.7 [23.5; 25.9]	5.2	32.8 [29.3; 36.2]	9.0	< 0.01
Width at 4.5 cm below the umbilicus	19.6 [18.4; 20.7]	7.3	14.8 [13.3; 16.2]	5.8	21.1 [20.0; 22.2]	4.6	26.1 [22.7; 29.6]	9.0	< 0.01

n, number; m, mean; SD, standard deviation.

Table 4 presents the results of width assessment, which showed good reliability when measured at the umbilicus and 4.5 cm below the umbilicus, and moderate reliability at 4.5 cm above the umbilicus. For the total group, the SEM was between 4.05–4.75 mm, and the minimal detectable change was 11.23–13.17 mm. Sub-analysis of the different rehabilitation centers revealed two negative outliers. At rehabilitation center 2, assessment at 4.5 cm below the umbilicus showed an ICC value of 0.51 [95% CI: 0.20; 0.70], which is at the lower boundary of the definition for moderate reliability. Assessment at 4.5 cm above the umbilicus at rehabilitation center 3 showed much lower ICC values compared to the other values. An ICC value of 0.40 indicates low reliability. At this measurement point, the SEM was 8.3 mm, and the minimal detectable change 23.01 mm.

Table 4
Results of the assessment of diastasis recti abdominis (DRA) width in mm

Parameters	Total group (n = 159)		Rehabilitation center 1 (n = 61)		Rehabilitation center 2 (n = 69)		Rehabilitation center 3 (n = 29)	
	ICC	SEM MDC	ICC	SEM MDC	ICC	SEM MDC	ICC	SEM MDC
	95% CI		95% CI		95% CI		95% CI	
Width at 45 mm above the umbilicus	0.73 [0.63; 0.80]	4.75 13.17	0.78 [0.63; 0.87]	5.32 14.75	0.60 [0.36; 0.75]	3.46 9.59	0.40 [- 0.32; 0.72]	8.30 23.01
Width at the umbilicus	0.83 [0.76; 0.87]	4.05 11.23	0.85 [0.75; 0.91]	3.29 9.12	0.62 [0.39; 0.77]	4.34 12.03	0.82 [0.61; 0.91]	4.93 13.66
Width at 45 mm below the umbilicus	0.80 [0.72; 0.85]	4.40 12.20	0.75 [0.58; 0.85]	3.64 10.09	0.51 [0.20; 0.70]	4.03 11.17	0.74 [0.43; 0.88]	6.16 17.07

n, number; ICC, intraclass correlation coefficient; CI, confidence interval; SEM, standard error of measurements; MDC, minimal detectable change.

Table 5 presents the results of the assessment of DRA bulging and depth. The assessment of DRA depth showed fair weighted kappa values ranging from 0.34–0.43. In the assessment of the depth, 21% were assessed as “good resistance at all points”, 67% as “resistance in the depth”, and 12% as “bottomless”. Assessment of linea alba bulging in the 3-step sit-up test showed kappa values ranging from 0.35–0.77. Among the participants, about 81% were rated as “no bulging”, 12% as “bulging of the linea alba”, and ~ 7% as “cannot assess”.

These results were further analyzed using the Svensson method. For the assessment of depth, we found a small concentration variance at rehabilitation center 1 (0.10 [95% CI: 0.00; 0.21]). For DRA bulging, we found a relative position variance at rehabilitation center 2 (0.17 [95% CI: 0.07; 0.26]). Most of this variance has to be explained by individual variation and not by a systematic bias. We identified no learning or fatigue effect between the first and the second assessments of bulging or depth.

Table 5
Results of the assessment of diastasis recti abdominis (DRA) depth and bulging

	Total Group (n = 159)			Rehabilitation center 1 (n = 61)			Rehabilitation center 2 (n = 69)			Rehabilitation center 3 (n = 29)		
	Kappa	95% CI	PA %	Kappa	95% CI	PA %	Kappa	CI	PA %	Kappa	95% CI	PA %
Depth	0.43	0.29; 0.56	69	0.37	0.15; 0.59	70	0.34	0.11; 0.58	71	0.36	0.04; 0.69	62
Bulging*	0.51		88	0.77		94	0.35		83	0.36		88

n, number; CI, confidence interval; PA, percentage agreement.

*The physiotherapists had the rating option of “cannot assess”, and these assessments were excluded. Assessments for all rehabilitation center = 137; rehabilitation center 1 = 52; rehabilitation center 2 = 59; and rehabilitation center 3 = 26.

Discussion

Main findings

The main findings of this study are that physiotherapists managing women's health in primary care have reliable methods available to assess voluntary PFM contraction and DRA width during the postpartum period. On the other hand, the assessment of involuntary contraction by observation and voluntary relaxation had kappa values with slight-to-fair agreement. Data with such low agreement are not useful for clinical practice or research [32]. Further investigations are needed to improve the clinical applicability and reliable assessment of these factors.

PFM

Our present results showed weighted kappa values of 0.59–0.70 for MVC assessment, which are higher values compared to in previous studies [34–37]. One explanation could be differences between the rating scales used for this muscle function. For example, in two prior studies, they rated only the squeezing and not the lifting component of the contraction [34, 36], which may not be specific enough for reliable assessment [38]. Furthermore, the previous studies had smaller sample sizes than our present study. The different results could also be related to differences in the study population [34–36] and study design [37].

The assessment of PFM endurance showed moderate reliability. However, there were several inconsistencies in the PFM endurance data. Assessment of PFM endurance at rehabilitation center 2 showed only fair reliability, with a kappa value of 0.31. At this center gave physiotherapist 2 systematically higher ratings compared to physiotherapist 1. Our present results are lower compared with the findings of Devreese *et al.* [38]. Notably, in the study of Devreese, a contraction longer than 10 seconds was rated as positive rather than a contraction longer than 30 seconds, as in the present study. Additionally, their study population was older and was not 3 months postpartum. In a time period of 30 seconds, it may more difficult to assess the exact point of time when the contraction is subsiding. It is also possible that postpartum women have, on average, a weaker contraction, making it more difficult to assess PFM endurance. Further research is needed to decide whether a PFM endurance of 10 seconds or 30 seconds is more clinically relevant, for example, to hold in urine while exercising.

Voluntary contraction via observation showed fair-to-moderate weighted kappa values. A prior MRI study reported that the average inward movement of the perineum is about 1 cm while sitting, and it is more than 2 cm in the supine position, according to Kegel in 1952 [39]. It could be assumed that this large movement would be easy to observe, and a higher kappa value was expected. A previous study reported high inter-rater reliability in the observation of inward perineum movement [20]. Correspondingly, another study showed that inward perineal movement could be observed with a kappa value of 0.91 among continent women, and 0.93 among incontinent women [38].

As factors other than PFM strength may contribute to urinary leakage postpartum [40], it is important to assess other aspects of PFM function. Unfortunately, in our present study, we did not find a reliable method for assessing involuntary contraction by observation, and we demonstrated inconsistent findings for the assessment of involuntary contraction by palpation. Accordingly, another study reported only fair inter-rater reliability for the assessment of involuntary contraction by observation and palpation [20]. Further studies are needed to develop improved methods for the assessment and rating of these PFM functions in clinical practice.

When prescribing postpartum PFM training, it must be considered that some women have hypertonic, overactive, and possibly painful PFM. The estimated prevalence of women with vulvodynia is 8% [41]. Women with hypertonic muscles exhibit poorer PFM strength and control [42]. It remains unclear whether women with hypertonic pelvic floor muscles should be advised to do PFM training. In clinical practice, physiotherapists recommend an individualized approach; however, we have found no research about this topic. Our present results showed that the rating of voluntary relaxation had slight-to-fair

Loading [MathJax]/jax/output/CommonHTML/jax.js

inter-rater reliability. Slieker-ten *et al.* reported similar findings [20]. Another study used a five-step rating scale for relaxation after contraction, and reported a correlation of 0.34 between two raters [43]. Even, Slieker-ten *et al.* recommend the addition of more rating steps to the scale, e.g., incomplete relaxation in their discussion. It is important to continue this research. Regardless of whether women with hypertonic PFM need more support in PFM training or the recommendation of no PFM training, there remains a need for better methods of assessing this condition.

DRA

Our present results showed moderate-to-good reliability in measuring DRA width using a caliper, after 60 assessments and additional training. The highest ICC value of 0.85 at the umbilicus was calculated from the assessments at rehabilitation center 1. The data from rehabilitation centers 2 + 3 required more careful interpretation due to two outliers and the larger 95% confidence intervals of the ICC values. The characteristics of the DRA at 3 months postpartum measured with the caliper (Table 3) were comparable with the DRA characteristics measured by ultrasound [44], indicating the measured values in our present study are true values for this population.

The pre-analysis and the additional training were important components of this study, strengthening the assumption that this technique required some experience and a strictly standardized protocol for measuring the DRA with a caliper. Our analysis of the first 61 assessments from rehabilitation centers 1 + 2 revealed a significant discrepancy between the results of these two centers. One bias identified in our additional training was that the accurate head lift of just 2–3 cm was an important factor. This observation is supported by the study of Mota *et al.* which showed that the distance between the two parts of the rectus abdominis decreased during a sit-up movement [45].

The SEM and minimal detectable change were higher at rehabilitation center 3 compared to rehabilitation centers 1 + 2 (Table 4). The minimal detectable change measured at 4.5 cm above the umbilicus was over 2 cm, raising doubt about whether these results have any clinical relevance. At rehabilitation center 3, fewer than 30 participants were recruited during the study period of over one year. These results indicate that in addition to training and experience, some continuity in measuring the DRA with a caliper is necessary for reliable assessment. It is also possible that it is more difficult to measure a wider DRA, considering that rehabilitation center 3 had significantly larger DRA widths.

The SEM in our present study was about 4–5 mm, which is over twice the SEM of 1.54 mm reported in an intra-rater reliability study [46]. We found no other studies reporting the minimal detectable change. However, a study comparing ultrasound and caliper assessment reported that the limits of agreement were between 1–2 cm [24], which is a comparable value to define the boundaries of error of a measurement method [47]. It is uncertain whether the assessment of DRA width is clinical relevant for normal values, considering that the 20th to 80th percentiles of a normal DRA at 6 months postpartum are 17–28 mm [44]. Both the limits of agreement of the study comparing ultrasound and caliper assessment, and the minimal detectable change in our present study, are nearly as large as or even larger than the expected variation of normal values. However, researchers have recently shown greater interest in the screening, assessment, and follow-up of moderate or large DRA (≥ 3.5 cm [48]), considering that no correlations have been found between mild and moderate DRA and low back pain or pelvic floor disorders [49]. Caliper measurement may be a useful method for the assessment and screening of women with moderate or large DRA after pregnancy.

We found fair-to-moderate weighted kappa values for the assessment of DRA depth. To our knowledge, there is no other study with which to compare these results. About 63–66% of the assessments were rated as “resistance in the depth”. The physiotherapists hypothesized that “resistance in the depth” was felt and assessed as soon as the participants did not activate their deeper abdominal muscles during the head lift. Accordingly, in a conference paper, Lee and Hodges described the increased tension in the linea alba caused by activation of the deep abdominal muscles [50]. Other studies have also shown that activation of the deep system changes the behavior of the linea alba [25, 45]. Future studies must more precisely define the pre-activation of the deep abdominal muscles in the assessment of DRA depth.

Similar considerations are raised about abdominal muscle pre-activation and insufficient experience with this kind of rating when interpreting the data regarding “bulging of the linea alba” during a sit-up curl. Three weeks after the start of the study, the physiotherapists requested an additional option to rate as “unsure” since the assessment could be complicated by overhanging skin, abdominal fat, or the inability to perform a sit-up curl. Analysis using the Svensson method could not confirm the hypothesis that there was less bulging or depth during the second of the two assessments due to the learning effect of doing this kind of exercise.

Strengths and limitations

One strength of this study was the large sample size and the quantity of different aspects of muscle function. A comparable study assessing different aspects of PFM function in women with and without pelvic floor disorders included only 41 participants [20]. A review about DRA assessment methods included studies with 20–106 participants, and these studies only examined DRA width [21]. Another strength of this study was that we were able to perform the same tests at three different centers in different parts of west Sweden. This makes our results transferable for different physiotherapists using the same methods, and for a large group of postpartum women.

In reliability studies, it is important to discuss whether a deviation in the agreement between two measurements is caused by the assessment method (instrumental reliability) or by the investigators (rater reliability) [51]. Here we utilized the Svensson method for all assessments performed on nominal and ordinal scales, to distinguish between weaknesses in the scales or assessment methods and systematic bias between the physiotherapists. Systematic bias between the physiotherapists was detected in all assessments with kappa or weighted kappa values of < 0.41 , except for the assessment of voluntary contraction via observation at rehabilitation centers 1 and 3.

The standardization of our assessments using a strict protocol can be discussed as both a strength and a limitation of this study. On one hand, a clinical test or assessment should be described in as much detail as possible to achieve high inter-rater agreement [52]. On the other hand, the individual approach of assessing and palpating a muscle may be an important aspect in the assessment of muscle function [53].

The present study also had several clear limitations. One was that we lacked access to the participants’ delivery records. Thus, our analysis of statistically significant differences associated with age, BMI, mode of delivery, pelvic floor tearing, and highest birth weight were based on self-reported data from the participants. Another issue was the uneven cell distribution seen in over 50% of the rating scales tested in this study. In the literature, it is controversial whether a low kappa value can be explained by uneven cell distribution or low prevalence of a condition [54, 55]. We also faced the problem of rating a negative status based on the difficulties in determining a true status [55]. Notably, mild cases in a healthy population of postpartum women could potentially lead to an overestimation of conditions [56].

Another limitation was that the assessment methods used in this study did not distinguish between superficial and deep PFM. We focused on assessing the levator ani muscle, which may be affected by micro and macro damages during delivery [9, 57, 58]]. However, other muscles and structures can be affected by pregnancy and vaginal birth, leading to pelvic floor disorders, such as urinary incontinence [59]. For example, Devreese *et al.* reported that women with mild-to-moderate urinary stress incontinence exhibited weak superficial PFM [38].

Future research

For a next step, we must investigate the extent to which these values are clinically relevant for postpartum women—for example, if the assessment outcomes are associated with pain and dysfunction. Furthermore, we must determine a cut-off point for DRA severity relative to pain and dysfunction. It will also be important to determine what training advice should be given to postpartum women based on the results of these examinations, and to define how much DRA must be changed to substantially improve pain and quality of life. Until these questions are answered, it is difficult to decide how clinically relevant these assessment methods are for postpartum women.

There also remains a need for further research on how to assess and rate involuntary contraction by observation and palpation, and voluntary relaxation, in the clinical assessment of women after pregnancy. Furthermore, additional research is needed regarding the assessment of the DRA, in terms of depth and bulging of the linea alba in movement.

Conclusion

Women are increasingly demanding assessment of their pelvic floor muscles and DRA after pregnancy. Our present results revealed moderate-to-substantial reliability for the assessment of MVC and PFM endurance after pregnancy. Furthermore, DRA width can be measured by caliper, with an SEM of 4–5 mm and a minimal detectable change of about 1.2 cm. However, assessment using this instrument requires some experience and training.

Abbreviations

CI

Confidence interval; BMI:Body mass index; DRA:Diastasis recti abdominis; ICC:Intraclass correlation coefficient; MVC:Minimal detectable change; PA:Percentage agreement; PFM:Pelvic floor muscle*; SD:Standard deviation; SE:Standard error; SEM:Standard error of measurements

Declarations

Acknowledgements

We thank Lisa Altvall, Ulrika Hansson, Anna-Lena Magnberg, Vanesa Rufete Bernal, Ute Jesberg, and Johanna Ekberg, who assessed the 222 women in the current study.

Funding

This study was funded by the Research and Development Centre, Region Västra Götaland, Gothenburg, Sweden.

Ethics declarations

Ethics approval and consent to participate

All included women gave their written informed consent to participate. The study protocol was approved by the Swedish ethical review authority in Gothenburg in April 2018 (Dnr 088-18) and the local data protection service in January 2019. The study protocol was registered at clinicaltrials.gov in October 2018 (registration number: NCT03703804) and at the Research and Development Centre in Sweden (registration number: 243071).

Consent for publication

The individuals shown in Figure 2 a+b gave their written informed consent to have these images published in a scientific publication.

Contributions

Each author of this paper meets the criteria for authorship. SV, MFO, AG, GR, and MEHL designed the study. SV collected the data, analyzed and performed a first interpretation of the results, and drafted the article. MFO, AG, GR, and MEHL participated in the final interpretation of the results, and critically revised the article for important intellectual content. The final manuscript was seen and approved by all authors.

Competing interests

Loading [MathJax]/jax/output/CommonHTML/jax.js

The authors declare that they have no competing interests.

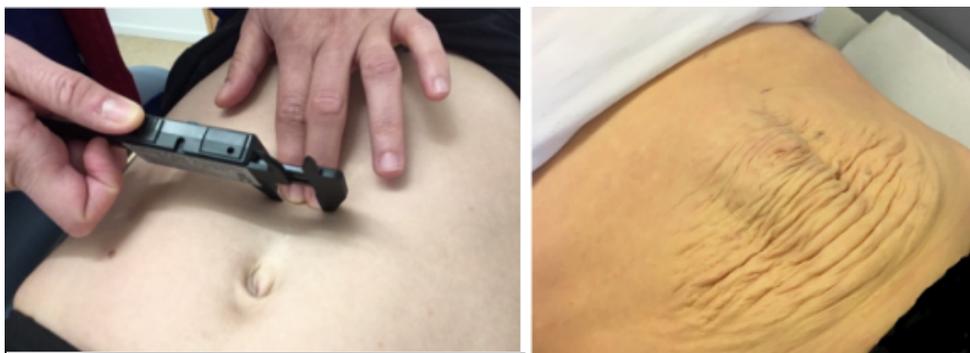
References

1. Bliss LS, Teeple P. Core Stability: The Centerpiece of Any Training Program. *Curr Sports Med Rep*. 2005;4(3):179–83.
2. Sahrmann S. *Diagnosis and treatment of movement impairment syndromes*. St. Louis: Mosby; 2002.
3. Hodges PW, Sapsford R, Pengel LHM. Postural and respiratory functions of the pelvic floor muscles. *Neurourol Urodyn*. 2007;26(3):362–71.
4. Hodges PW, Eriksson AE, Shirley D, Gandevia SC. Intra-abdominal pressure increases stiffness of the lumbar spine. *J Biomech*. 2005;38(9):1873–80.
5. DeLancey JO. Structural support of the urethra as it relates to stress urinary incontinence: the hammock hypothesis. *Am J Obstet Gynecol*. 1994;170(6):1713–20. discussion 20 – 3.
6. Moser H, Leitner M, Baeyens JP, Radlinger L. Pelvic floor muscle activity during impact activities in continent and incontinent women: a systematic review. *Int Urogynecol J*. 2018;29(2):179–96.
7. Sapsford RR, Hodges PW, Richardson CA, Cooper DH, Markwell SJ, Jull GA. Co-activation of the abdominal and pelvic floor muscles during voluntary exercises. *Neurourol Urodyn*. 2001;20(1):31.
8. Fernandes da Mota PG, Pascoal AG, Carita AI, Bo K. Prevalence and risk factors of diastasis recti abdominis from late pregnancy to 6 months postpartum, and relationship with lumbo-pelvic pain. *Man Ther*. 2015;20(1):200–5.
9. Ashton-Miller JA, DeLancey JOL. On the Biomechanics of Vaginal Birth and Common Sequelae. *Annu Rev Biomed Eng*. 2009;11(1):163–76.
10. Vikström A, Johansson S-E, Barimani M. Postnatal ER visits within 30 days—Pattern, risk factors and implications for care. *J Clin Nurs*. 2018;27(3–4):769–76.
11. Cyr M-P, Kruger J, Wong V, Dumoulin C, Girard I, Morin M. Pelvic floor morphometry and function in women with and without puborectalis avulsion in the early postpartum period. *Am J Obstet Gynecol*. 2017;216(3):274. .e1-.e8.
12. Memon H, Handa VL. Pelvic floor disorders following vaginal or cesarean delivery. *Curr Opin Obstet Gynecol*. 2012;24(5):349–54.
13. Sperstad JB, Tennfjord MK, Hilde G, Ellström-Engb M, Bø K. Diastasis recti abdominis during pregnancy and 12 months after childbirth: prevalence, risk factors and report of lumbopelvic pain. *British Journal of Sports Medicine*. 2016.
14. Keshwani N, Mathur S, McLean L. Relationship Between Interrectus Distance and Symptom Severity in Women With Diastasis Recti Abdominis in the Early Postpartum Period. *Phys Ther*. 2018;98(3):182–90.
15. Vieillefosse S, Conté C, Deffieux X. Symptom associated with diastasis recti in the early postpartum period. *European Journal of Obstetrics Gynecology Reproductive Biology*. 2019;234:e133-e.
16. Keeler J, Albrecht M, Eberhardt L, Horn L, Donnelly C, Lowe D. Diastasis Recti Abdominis: A Survey of Women's Health Specialists for Current Physical Therapy Clinical Practice for Postpartum Women. *Journal of Women's Health Physical Therapy*. 2012;36(3):131–42.
17. Bo K, Sherburn M. Evaluation of female pelvic-floor muscle function and strength. *Phys Ther*. 2005;85(3):269–82.
18. Bo K, Frawley HC, Haylen BT, Abramov Y, Almeida FG, Berghmans B, et al. An International Urogynecological Association (IUGA)/International Continence Society (ICS) joint report on the terminology for the conservative and nonpharmacological management of female pelvic floor dysfunction. *Neurourol Urodyn*. 2017;36(2):221–44.
19. Messelink B, Benson T, Berghmans B, Bo K, Corcos J, Fowler C, et al. Standardization of terminology of pelvic floor muscle function and dysfunction: report from the pelvic floor clinical assessment group of the International Continence Society. *Neurourol Urodyn*. 2005;24(4):374–80.

20. Slieker-Ten Hove MCP, Pool-Goudzwaard AL, Eijkemans MJC, Steegers-Theunissen RPM, Burger CW, Vierhout ME. Face validity and reliability of the first digital assessment scheme of pelvic floor muscle function conform the new standardized terminology of the International Continence Society. *Neurourol Urodyn*. 2009;28(4):295–300.
21. van de Water ATM, Benjamin DR. Measurement methods to assess diastasis of the rectus abdominis muscle (DRAM): A systematic review of their measurement properties and meta-analytic reliability generalisation. *Man Ther*. 2016;21:41–53.
22. Keeler J, Albrecht M, Eberhardt L, Horn L, Donnelly C, Lowe D. Diastasis Recti Abdominis: A Survey of Women's Health Specialists for Current Physical Therapy Clinical Practice for Postpartum Women. *Journal of Women's Health Physical Therapy*. 2012;36(3):131–42.
23. Bursch SG. Interrater reliability of diastasis recti abdominis measurement. *Phys Ther*. 1987;67(7):1077–9.
24. Barbosa S, de Sa RA, Coca Velarde LG. Diastasis of rectus abdominis in the immediate puerperium: correlation between imaging diagnosis and clinical examination. *Archives of gynecology obstetrics*. 2013;288(2):299–303.
25. Lee D, Hodges PW. Behavior of the Linea Alba During a Curl-up Task in Diastasis Rectus Abdominis: An Observational Study. *J Orthop Sports Phys Ther*. 2016;46(7):580.
26. Beamish N, Green N, Nieuwold E, McLean L. Differences in Linea Alba Stiffness and Linea Alba Distortion Between Women With and Without Diastasis Recti Abdominis: The Impact of Measurement Site and Task. *J Orthop Sports Phys Ther*. 2019;49(9):656–65.
27. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*. 2016;15(2):155–63.
28. Chiarello CM, Falzone LA, McCaslin KE, Patel MN, Ulery KR. The effects of an exercise program on diastasis recti abdominis in pregnant women. *Journal of Women's Health Physical Therapy*. 2005;29(1):11–6.
29. Chiarello CM, McAuley JA. Concurrent validity of calipers and ultrasound imaging to measure interrecti distance. *J Orthop Sports Phys Ther*. 2013;43(7):495–503.
30. Hills NF, Graham RB, McLean L. Comparison of Trunk Muscle Function Between Women With and Without Diastasis Recti Abdominis at 1 Year Postpartum. *Phys Ther*. 2018;98(10):891.
31. Landis JR, Koch GG. A One-Way Components of Variance Model for Categorical Data. *Biometrics*. 1977;33(4):671–9.
32. McHugh ML. Interrater reliability: the kappa statistic. *Biochemia medica*. 2012;22(3):276–82.
33. Svensson E. Different ranking approaches defining association and agreement measures of paired ordinal data. *Stat Med*. 2012;31(26):3104–17.
34. Ferreira CH, Barbosa PB, de Oliveira Souza F, Antonio FI, Franco MM, Bo K. Inter-rater reliability study of the modified Oxford Grading Scale and the Peritron manometer. *Physiotherapy*. 2011;97(2):132–8.
35. Navarro Brazalez B, Torres Lacomba M, de la Villa P, Sanchez Sanchez B, Prieto Gomez V, Asunsolo Del Barco A, et al. The evaluation of pelvic floor muscle strength in women with pelvic floor dysfunction: A reliability and correlation study. *Neurourol Urodyn*. 2018;37(1):269–77.
36. Bø K, Finckenhagen HB. Vaginal palpation of pelvic floor muscle strength: inter-test reproducibility and comparison between palpation and vaginal squeeze pressure. *Acta Obstet Gynecol Scand*. 2001;80(10):883–7.
37. Neumann PB, Grimmer-Somers KA, Gill VA, Grant RE. Rater reliability of pelvic floor muscle strength. *Australian New Zealand Continence Journal*. 2007;13(1):8–14.
38. Devreese A, Staes F, De Weerd W, Feys H, Van Assche A, Penninckx F, et al. Clinical evaluation of pelvic floor muscle function in continent and incontinent women. *Neurourol Urodyn*. 2004;23(3):190–7.
39. Bo K, Lilleas F, Talseth T, Hedland H. Dynamic MRI of the pelvic floor muscles in an upright sitting position. *Neurourol Urodyn*. 2001;20(2):167–74.
40. Fritel X, Ringa V, Quiboeuf E, Fauconnier A. Female urinary incontinence, from pregnancy to menopause: a review of findings. *Acta Obstet Gynecol Scand*. 2012;91(8):901–10.

41. Rosen NO, Dawson SJ, Brooks M, Kellogg-Spadt S. Treatment of Vulvodynia: Pharmacological and Non-Pharmacological Approaches. *Drugs*. 2019;79(5):483–93.
42. Goldstein AT, Pukall CF, Brown C, Bergeron S, Stein A, Kellogg-Spadt S. Vulvodynia: Assessment and Treatment. *J Sex Med*. 2016;13(4):572–90.
43. Reissing ED, Brown C, Lord MJ, Binik YM, Khalifé S. Pelvic floor muscle functioning in women with vulvar vestibulitis syndrome. *Journal of Psychosomatic Obstetrics Gynecology*. 2005;26(2):107–13.
44. Mota P, Pascoal AG, Carita AI, Bo K. Normal width of the inter-recti distance in pregnant and postpartum primiparous women. *Musculoskelet Sci Pract*. 2018;35:34–7.
45. Mota P, Pascoal AG, Carita AI, Bo K. The Immediate Effects on Inter-rectus Distance of Abdominal Crunch and Drawing-in Exercises During Pregnancy and the Postpartum Period. *J Orthop Sports Phys Ther*. 2015;45(10):781–8.
46. Boxer S, Jones S. Intra-rater reliability of rectus abdominis diastasis measurement using dial calipers. *Aust J Physiother*. 1997;43(2):109–14.
47. Turner D, Schünemann HJ, Griffith LE, Beaton DE, Griffiths AM, Critch JN, et al. The minimal detectable change cannot reliably replace the minimal important difference. *J Clin Epidemiol*. 2010;63(1):28–36.
48. Lo T, Candido G. P J. Diastasis of the Recti abdominis in pregnancy: risk factors and treatment. *Physiother Can*. 1999;51(1):32–44.
49. Sperstad JB, Tennfjord MK, Hilde G, Ellström-Engb M, Bø K. Diastasis recti abdominis during pregnancy and 12 months after childbirth: prevalence, risk factors and report of lumbopelvic pain. *Br J Sports Med*. 2016;50(17):1092.
50. Lee D, Hodges PW. Diastasis rectus abdominis – Should we open or close the gap? *Musculoskeletal Science Practice*. 2017;28:e16.
51. Bruton A, Conway JH, Holgate ST. Reliability. What is it, and how is it measured? *Physiotherapy*. 2000;86(2):94–9.
52. Fritz JM, Wainner RS. Examining diagnostic tests: an evidence-based perspective. *Phys Ther*. 2001;81(9):1546–64.
53. Holmgren U, Waling K. Inter-examiner reliability of four static palpation tests used for assessing pelvic dysfunction. *Man Ther*. 2008;13(1):50–6.
54. Chmura Kraemer H, Periyakoil VS, Noda A. Kappa coefficients in medical research. *Stat Med*. 2002;21(14):2109–29.
55. Vach W. The dependence of Cohen's kappa on the prevalence does not matter. *J Clin Epidemiol*. 2005;58(7):655–61.
56. Lijmer JG, Mol BW, Heisterkamp S, Bossel GJ, Prins MH, van der Meulen JHP, et al. Empirical Evidence of Design-Related Bias in Studies of Diagnostic Tests. *JAMA*. 1999;282(11):1061–6.
57. Shek K, Dietz H. Vaginal Birth and Pelvic Floor Trauma. *Current Obstetrics Gynecology Reports*. 2019;8(2):15–25.
58. Dietz PH, Lanzarone PV. Levator Trauma After Vaginal Delivery. *Obstet Gynecol*. 2005;106(4):707–12.
59. Rikard-Bell J, Iyer J, Rane A. Perineal outcome and the risk of pelvic floor dysfunction: a cohort study of primiparous women. *Aust N Z J Obstet Gynaecol*. 2014;54(4):371–6.

Figures



Loading [MathJax]/jax/output/CommonHTML/jax.js

Figure 1

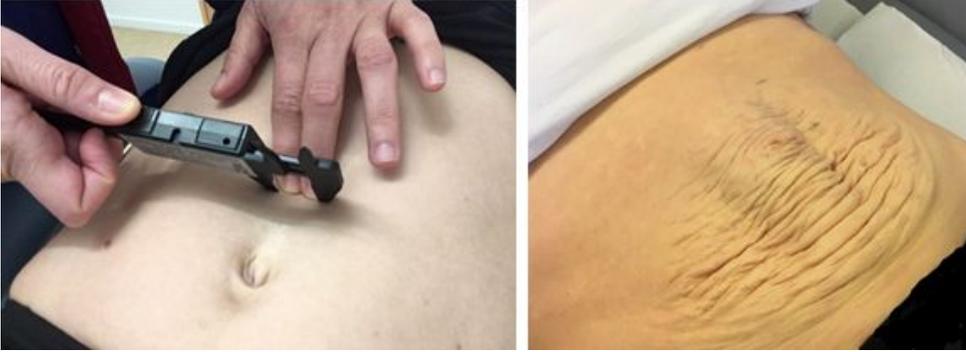


Figure 2

Figure 1

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplement4.docx](#)
- [supplement5.docx](#)