The Digital Caliper's Interrater Reliability in Measuring the Interrecti Distance and Its Accuracy in Diagnosing the Diastasis of Rectus Abdominis Muscle in the Third Trimester of Pregnancy

Maíra Belo, PT, MSc, [a] Adriana Melo, MD, PhD, [b] Alexandre Delgado, PT, MSc, [a, b] Adriana Costa, PT, [c] Valéria Anísio, PT, [c] and Andrea Lemos, PhD, PT [a]

ABSTRACT

Objective: The purpose of this study was to determine the digital caliper's interrater reliability in measuring the interrecti distance and its accuracy in diagnosing the diastasis of the recti abdomini muscles (DRAM) in the third trimester of pregnancy compared with ultrasound.

Methods: Fifty-four pregnant women were included. Measurements were taken in supine, during trunk flexion and expiration, in the supraumbilical region, in the umbilical scar (US), and in the infraumbilical (IU) region, at first by ultrasonography (USG) and then by the 2 digital caliper's examiners. The 3 evaluators were independent and blind.

Results: In the interrecti distance measurement, comparing the 2 calipers' evaluators, the instrument showed an intraclass correlation coefficient (ICC) of 0.86 (95% CI: 0.72-0.92) in the supra-umbilical and ICC of 0.96 (95% CI: 0.92-0.98) in the US. Compared with the USG, the instrument showed the worst interrater reliability in the US (ICC −0.14; 95% CI: −0.39 to 0.13). In diagnosing DRAM, comparing calipers' evaluators, kappa was 0.56 ($P < .01$) for the IU region and 0.12 ($P = .19$) for the US region. When compared with the USG, kappa was 0.02 ($P = .84$) for the IU region and 0.05 ($P = .59$) for the US. In the US, the caliper presented a positive likelihood ratio of 1.05 (95% CI: 0.86-1.30) and a negative likelihood ratio of 0.63 (95% CI: 0.12-3.43).

Conclusion: The digital caliper did not present good interrater reliability in measuring or in diagnosing the DRAM during the third trimester of pregnancy compared with ultrasound. (J Chiropr Med 2020;19;136-144)

Key Indexing Terms: *Abdominal Muscles; Pregnancy; Ultrasonography; Reproducibility of Results; Sensitivity and Specificity*

INTRODUCTION

The diastasis of the rectus abdominis muscle (DRAM) is defined as a separation of the muscle beams of the rectus abdominis from the linea alba, which increases the interrecti distance (IRD).[1-3] It is associated with genetic factors, and with conditions such as obesity, excessive weight loss, excess exercise, and pregnancy.[4] The prevalence in the pregnant population varies among studies from 30% to 70% in the third gestational trimester.[1,5]

The DRAM may lead to severe implications for the abdomen, resulting in sagging and bulging of the abdominal wall.[3,6] Although this condition does not directly cause discomfort or pain, in cases of excessive strain, it may interfere with abdominal muscle functionality and generate greater predisposition to various disorders such as back pain and aesthetic and urogynecological problems.[7-9]

Several methods to diagnose this condition have been reported in literature.[3] The variety of measurement methods used among the studies leads to a lack of standardization, making it difficult to compare their findings. At first, palpation was the method used to measure the IRD,[5,9,10] and in

[a] Department of Physical Therapy, Universidade Federal de Pernambuco, Recife-PE, Brazil.
[b] Instituto de Medicina Integral Professor Fernando Figueira, Recife-PE, Brazil.
[c] Departament of Physical Therapy, Faculdade Maurício de Nassau, Campina Grande-PB, Brazil.
Corresponding author: Andrea Lemos, PhD, PT, Physiotherapy Department, Federal University of Pernambuco, Av. Professor Morais Rego, 1235, Cidade Universitária, Recife 50740-600, Brazil.
(e-mail: *andrealemos4@gmail.com*).
Paper submitted May 9, 2019; in revised form December 28, 2019; accepted February 14, 2020.
1556-3707

https://doi.org/10.1016/j.jcm.2020.02.002

these cases it was considered DRAM if the IRD was greater than 2 fingers. However, the current literature points out that subcutaneous fat and excessive abdominal flabbiness can interfere with measurements and that the finger's width may vary among people, which makes palpation an unreliable method to diagnose DRAM.[11-13]

Currently, the method of choice considered the gold standard for evaluation of the abdominal wall and hence measurement of DRAM is the computed tomography scan. However, besides being an expensive method, it exposes the patient to radiation, and thus it cannot be used on pregnant women.[11,13] Therefore, ultrasonography (USG) has been proposed as an alternative to evaluate the abdominal wall, especially during pregnancy. One study demonstrated it as an accuracy method of measuring DRAM, especially in the supraumbilical (SU) level and at the umbilicus.[11]

Digital calipers have also been reported as a method to measure the IRD and thus to diagnose DRAM, and it has been used in several studies.[2,8,13-16] Recently, a systematic review[17] aimed to investigate the best method to screen for DRAM presence and monitor DRAM width. Only 3 studies[13,15,18] that evaluated calipers as a measurement tool to screen for DRAM were identified. All of them happened during puerperium, and only 1 study[15] evaluated the instrument's accuracy. According to the raw data presented by the study, a systematic review calculated the following: sensitivity of 89.7%, specificity of 75%, and positive predictive value of 82.5%, although there was no confidence interval (CI) and the study was biased.

Therefore, the purpose of the present study was to determine the interrater reliability of digital calipers in measuring the IRD and its accuracy in diagnosing the DRAM during the third trimester of pregnancy compared with the ultrasound.

Methods

This was an accuracy and reliability study, under Standards for Reporting of Diagnostic Accuracy[19] and Guidelines for Reporting Reliability and Agreement Studies[20] orientation. The Standards for Reporting of Diagnostic Accuracy was used to assess the diagnostic study and Guidelines for Reporting Reliability and Agreement Studies to evaluate the reliability. It took place in a maternity ward in the city of Campina Grande, Brazil.

The sample consisted of 54 pregnant women in a nonprobabilistic sequential manner, between the 28th and 41st weeks of gestation, aged from 18 to 35 years. The exclusion criteria were as follows: high-risk gestation with rest recommendation, incapacity to understand or carry out the procedure, and neuromuscular disorders. After being invited to join the study, the participants freely consented to participation by signing the informed consent term. The study had the approval of the Ethical and Research on Human Beings Committee of Hospital Universitário Alcides Carneiro under protocol number CAAE 30780314.2.0000.5182.

All women were submitted to an evaluation consisting of the collection of socioeconomic (age, education), clinical (gestational age, gestation, and parity number), and anthropometric (weight, pregestational weight, height, body mass index, nutritional status—according the criteria of Atalah et al[21]—weight gain, and waist circumference measure) data. One examiner who was previously trained assessed the DRAM by USG (Philips HD3xe, linear transducer of 5-9 MHz) (Fig 1).

Ultrasonography was used as the gold standard for DRAM evaluation, as magnetic resonance imaging is contraindicated during pregnancy. Ultrasonography has the advantage of not being invasive and ionizing, which is why it has been referred to as the gold standard for the diagnosis of DRAM in pregnant women.[11,14]

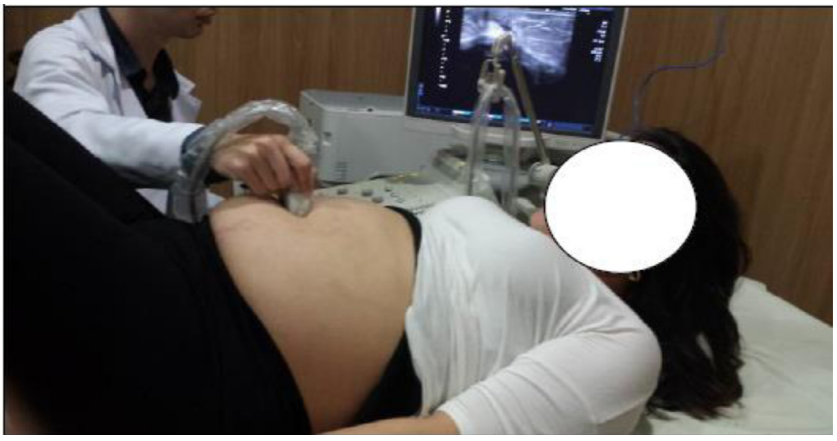

Fig 1. *DMRA measurement position. Patient in dorsal decubitus, hips and knees flexed, and feet supported on the supporting surface. Lateral flexion of the trunk during expiration.*

The participants were placed in a dorsal decubitus position, with hips and knees flexed, feet placed on bed, and arms resting on each side of the body.[3,10,13,14,22] In this position, the abdomen was marked at the points located at the SU region (3 cm above the umbilicus), in the umbilical scar (US), and in the infraumbilical (IU) place (2 cm bellow the umbilicus). The first examiner placed the USG transversally in each of the points previously marked and instructed the patient to flex the trunk during exhalation until the lower borders of the scapula no longer touched the bed, and then the IRD was measured (Fig 2).

After this first evaluation, 2 different examiners who were previously trained measured the IRD using a digital caliper (Starrett 799) to measure the IRD to identify the DRAM and compare it with the use of USG. The measurements were taken in the same points and position described before.

Each examiner performed measurements 3 times at each point, and there was blindness for all the measurements. The mean of the 3 measurements were used in the data interpretation.

To classify the participants with DRAM positive or negative, we took into consideration the study of Beer et al[23] of nulliparous women, to the supraumbilical and infraumbilical points. According to this study, an IRD greater than 2.2 cm SU and over 1.6 cm IU were considered DRAM positive. For the US it was considered DRAM positive at measurements greater than 2.0 cm, according to Boissonnault and Blaschak.[5]

The IRD measurements were presented as means, standard deviations (SDs), minimum, maximum, and CIs at 95%. The means of the 3 IRD measurements of the first examiner were calculated, as were the means of both calipers' evaluators. Then, the mean of the measurements obtained by the first examiner were compared with the mean of the 2 calipers' evaluators. The prevalence of DRAM was presented through absolute number and frequency.

The results of the benchmark test (digital caliper) for each local measurement (SU, US, and IU) were compared with the gold-standard test results (ultrasound) through a 2×2 table, and from there, the kappa was calculated, with CI at 95% to obtain the caliper's interrater reliability in diagnosing the DRAM condition (DRAM positive or negative). Also, the intraclass correlation coefficient (ICC), with CI at 95%, was calculated to present the instrument's interrater reliability in measuring the IRD. The accuracy variables, such as sensibility, specificity, negative and positive predictive values, and positive and negative likelihood ratios with CI at 95%, were also calculated.

For the statistical analysis, the SPSS software version 20.0 (IBM SPSS Statistics for Windows, Version 20.0, IBM Corp., Armonk, NY) and the Stata software (Statacorp LLC, College Station, TX) were used.

Results

The sample consisted of 54 pregnant women with a mean age of 24.6 (SD = 5.58) years, a mean gestational age of 34.3 (SD = 3.44) weeks, an average of 2.2 (SD = 1.41) gestations, and an average body mass index of 29.3 (SD = 5.14) kg/m^2.

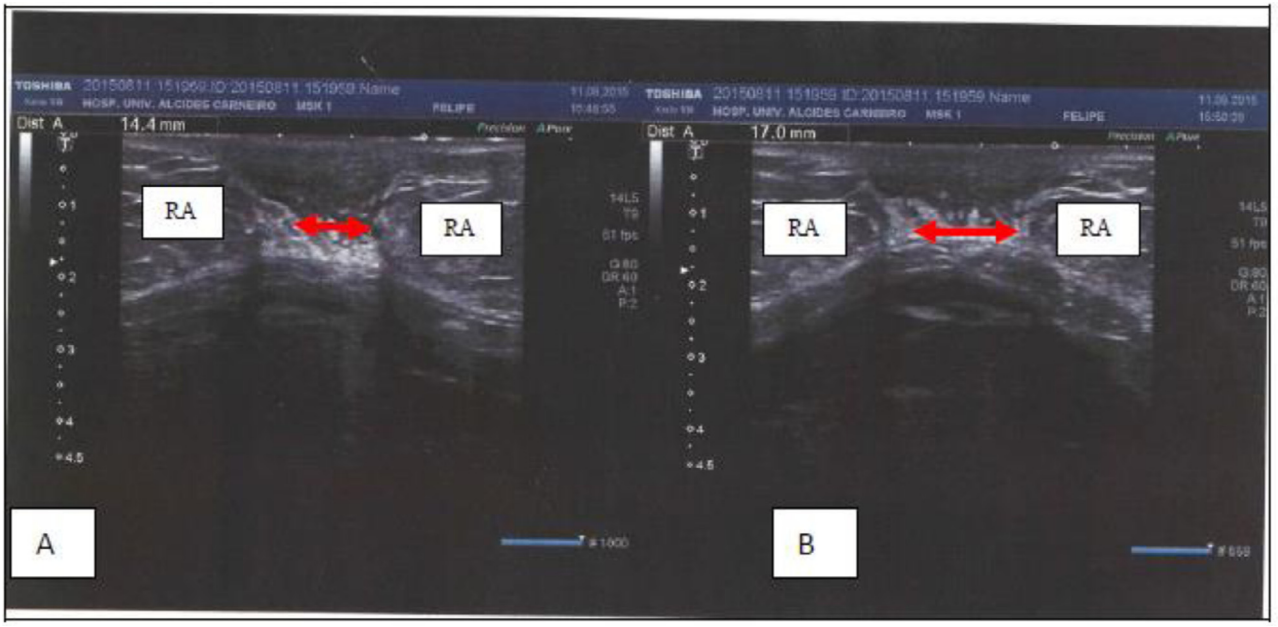

Fig 2. *Abdominal USG image of supraumbilical region of study participant, in dorsal decubitus position, with flexed hips and knees and feet resting on the bed. (A) The DIR of 1.44 cm at rest. (B) The DIR of 1.7 cm during expiration and contraction. Measurements of DIR obtained in abdominal contraction were taken.* IRD, *interrecti distance;* RA, *rectus abdominis;* USG, *ultrasonography.*

Table 1. *Description of Socioeconomic, Clinical, and Anthropometric Variables in Third Trimester Pregnant Women*

	Mean	SD	Minimum	Maximum	95% CI
Age (y)	24.6	5.58	18	37	23.22-26.06
Years of study	8.7	3.56	0	20	7.80-9.59
Gestational age (wk)	34.3	3.44	28	41	33.35-35.21
No. of gestation	2.2	1.41	1	7	1.87-2.61
No. of births	1.0	1.20	0	4	0.70-1.35
BMI (kg/m^2)	29.3	5.14	19.9	43.7	27.94-30.6.1
WC (cm)	108.0	19.04	88.1	190.0	103.46-114.03
Weight gain (kg)	11.2	6.48	2.5	37.0	9.62-12.94

BMI, body mass index; *CI,* confidence interval; *SD*, standard deviation; *WC*, waist circumference.

A total of 35.2% of the sample was classified as eutrophics, 27.8% were considered as presenting overweight, 24.1% were obese, and 13% were considered low weight by the criteria of Atalah et al[21] (Table 1).

Both USG and digital caliper showed that the greater IRD width was at the US spot (USG: 2.4 cm, SD = 0.74; digital caliper: 2.66 cm, SD = 0.5) (Table 2). The frequency of DRAM varied according to the measurement point, and the highest values were found at the US (70.3%) (Table 2).

Ultrasound vs Digital Calipers Interrater Reliability

The interrater reliability between digital calipers and USG in measuring the IRD showed an ICC of −0.11 (95% CI: −0.36 to 0.16) in the SU, −0.07 (95% CI: −0.33 to 0.20) in the IU, and −0.14 (95% CI: −0.39 to 0.13) in the US.

When considering the interrater reliability of the instrument for diagnosing DRAM, the instrument showed a kappa value of 0.02 ($P = .84$) for the IU and 0.05 ($P = .59$) for the US points (Table 3). It was not possible to calculate this variable for the SU point owing to the lack of positive cases diagnosed in that region by the digital caliper.

Table 2. *Description of the Interrecti Distance Obtained Through the Ultrasonography (USG) and Digital Caliper, During Exhalation Followed by Trunk Flexion, in the Anatomic Points Established in Literature*

	Mean	SD	Minimum	Maximum	95% CI
USG					
SU (3 cm)	1.60	0.72	0.49	4.00	141-1.78
CU	2.40	0.74	0.88	4.00	2.20-2.60
IU (2 cm)	1.80	0.68	0.34	3.50	1.63-1.98
Digital caliper					
SU (3 cm)	1.35	0.28	0,.96	2.01	128-1.42
US	2.66	0.50	1.84	3.78	2.54-2.80
IU (2 cm)	1.64	0.27	1.17	2.34	1.57-1.71

Digital Calipers Interrater Reliability

For the interrater reliability of the IRD measurements, the instrument presented a substantial reliability (>0.80), in the SU point (ICC 0.86; 95% CI: 0.72-0.92) and at the US (ICC 0.96; 95% CI: 0.92-0.98) (Table 4).

The analysis of Bland-Altman plots (Fig 3) indicates a substantial variability of the measures for the 3 points measured, especially in the US.

When comparing the 2 digital caliper evaluators in diagnosing DRAM, the instrument demonstrated a kappa value of 0.56 ($P < .01$) in the IU spot, and a kappa result of 0.12 ($P = .19$) in the US level. It was not possible to calculate the kappa in the SU region owing to the lack of DRAM cases diagnosed by the digital caliper (Table 3).

Table 3. *Digital Caliper's Reliability Regarding the Presence or Absence of RAD in Pregnant Women in the Third Trimester of Gestation*

Local	Kappa Paq/Paq[a]	*P* Value	Kappa USG/ Paq[b]	*P* Value
Supraumbilical (3 cm)	[c]	[c]	[c]	[c]
Umbilical scar	0.12	0.19	0.05	.59
Infraumbilical (2 cm)	0.56	0.00	0.02	.84

Pac, Pachymeter*; RAD*, recti abdomini diastasis; *USG*, ultrasonography.
[a] Between the 2 caliper's evaluators.
[b] Between USG and caliper.
[c] It was not possible to be calculated because of the lack of positive cases identified by calipers in this spot.

Table 4. *Digital Caliper's Interrater Reliability Regarding the Quantitative Measurements of the IRD, Between Caliper and USG and Between the 2 Caliper's Evaluators, in Pregnant Women in the Third Trimester of Gestation*

Measurement Spot	ICC[a] (95% CI) Cal/USG	ICC[b] (95% CI) Cal/Cal
Supraumbilical (3 cm)	−0.11 (−0.36 to 0.16)	0.86 (0.72-0.92)
Umbilical scar	−0.14 (−0.39 to 0.13)	0.96 (0.92-0.98)
Infraumbilical (2 cm)	−0.07 (−0.33 to 0.20)	0.79 (0.59-0.89)

The values in parentheses refer to the distance (cm) from the point where the measurements were taken to the umbilical scar.
Cal, caliper; *CI,* confidence interval; *ICC*, intraclass correlation coefficient; *IRD*, interrecti distance; *USG*, ultrasonography.
[a] ICC between caliper and USG.
[b] ICC between the 2 caliper's evaluators.

Accuracy Results

The digital caliper's accuracy varied according to the location of measurement. Considering the US, the instrument showed the following results: sensibility (92.1%; 95% CI: 78.6-98.2), positive predictive value (71.4%; 95% CI: 56.7-83.4), specificity (12.5%; 95% CI: 1.9-38.3), and negative predictive value (NPV) (40%; 95% CI: 6.4- 84.6) (Table 5).

The results seen in the SU point were specificity of 100% (95% CI: 91.9-100), negative likelihood ratio of 1.00, and NPV of 81.4% (95% CI: 68.5-90.7). In the SU analysis of sensibility, positive likelihood ratio and positive

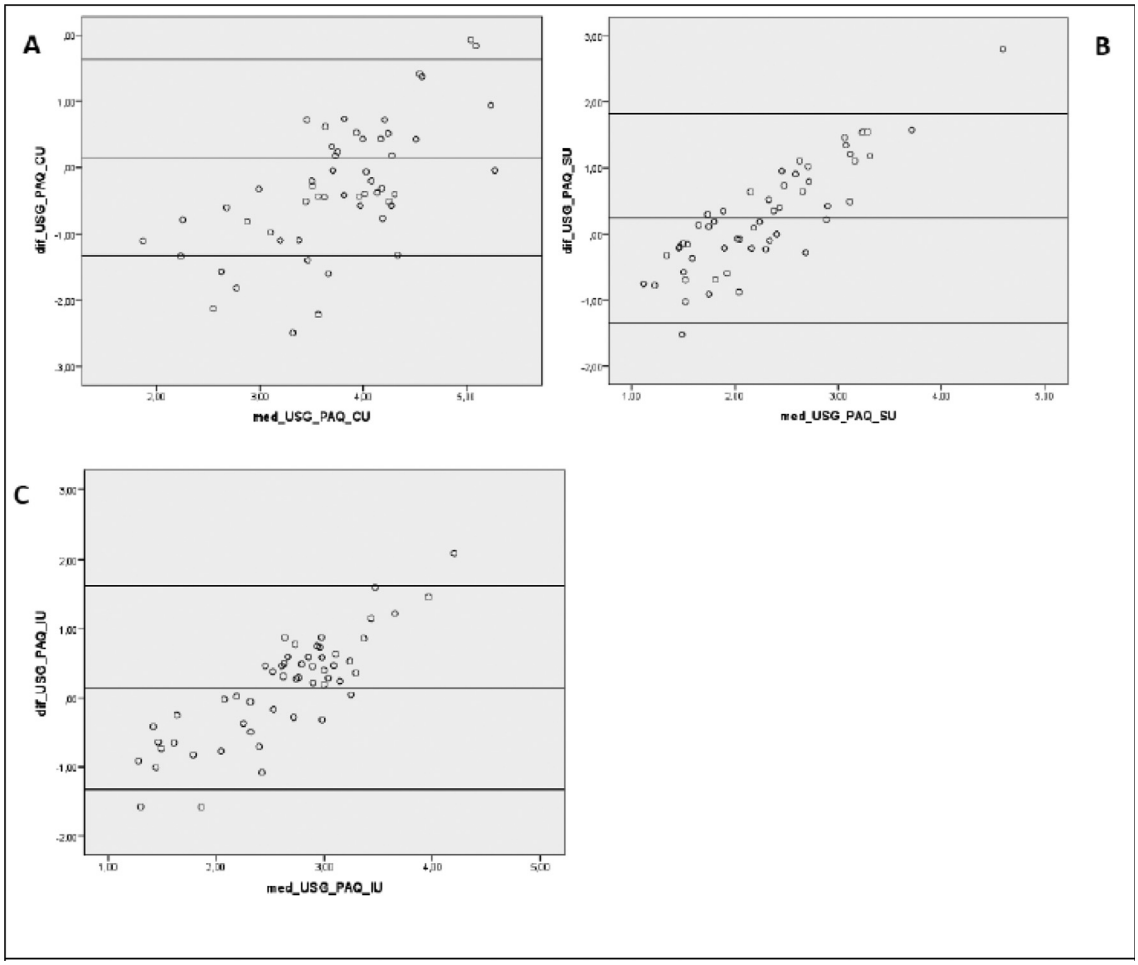

Fig 3. *Bland-Altman plots showing the reliability of digital calipers compared with USG: (A) umbilical scar, (B) 3 cm supraumbilical, and (C) 2 cm infraumbilical point.* USG*, ultrasonography.*

Table 5. *Description of Accuracy Variables Resulting From the Comparison Between Digital Caliper and Ultrasonography*

	Supraumbilical	Umbilical Scar	Infraumbilical
Sensibility % (95% CI)	0.0% (0.0-30.8)	92.1% (78.6-98.2)	56% (34.9-75.6)
Specificity % (95% CI)	100% (91.9-100)	12.5% (1.9-38.3)	51.7% (32.5-70.5)
Positive likelihood ratio	[a]	1.05 (0.86-1.30)	1.16 (0.6-1.9)
Negative likelihood ratio	1.00 (1.00-1.00)	0.63 (0.12-3.43)	0.85 (0.48-1.41)
Positive predictive value	[a]	71.4% (56.7-83.4)	16.6% (2.7-63.9)
Negative predictive value % (95% CI)	81.4% (68.5-90.7)	40% (6.4-84.6)	50% (30.6-69.3)
Prevalence % (95% CI)	18.5% (9.2-31.4)	70.3% (56.3-82)	57.6% (36.9-76.6)

CI, confidence interval.
[a] It was not possible to calculate because of the lack of cases identified by digital calipers.

predictive value were impossible to be done because of the lack of positive cases diagnosed by the digital caliper, using the cutoff point of 2.2 cm, suggested by Beer et al [23] (Table 5).

Discussion

The present study found the US as the point with the greater IRD width and thus the higher prevalence of DRAM. The instrument did not present good results concerning its accuracy, especially in the SU point. On the other hand, regarding the instrument's interrater reliability, digital calipers presented poor agreement in diagnosing DRAM and in measuring the IRD compared with USG.

Regarding the IRD width, these findings corroborate 3 studies[11,14,24] in which the authors found that the wider IRD is seen in the umbilical level; however, other authors[1,23-26] have chosen not to measure this owing to technical difficulties with ultrasound, and therefore some of them [23,25,26] reported the SU region as the point with the greater IRD width.

The fact that the second largest IRD was found in the IU spot also differs from what has been found in literature. Whereas the present study has found the smallest IRD above the US, several studies[11,14,23,24,27] have reported the IU region as the narrowest IRD. One of these studies[27] has suggested that the IU region of the linea alba has a major ability to resist stresses imposed over a greater period, as happens during pregnancy. Nevertheless, in the present study the population investigated consisted of pregnant women, while the ones mentioned above were conducted during puerperium.

Regarding DRAM frequency, the present research found higher frequencies in the US (70.3%). Analyzing the numbers reported in literature, the prevalence varied among studies from 30% to 70% in the last gestational trimester.[1,5] The wide range observed in DRAM prevalence is mainly due to the variability in the cutoff points used in different studies. One of them[14] showed different results when trying to describe a prevalence of DRAM depending on the reference used for its classification.

Regarding the accuracy and reliability data, the present study has dealt with the difficulty of the lack of standardization concerning the criteria for the diagnosis of DRAM. Among studies, it is suggested that there is DRAM when the IRD exceeds 1.5 cm,[28] 2 cm,[3,29] 2.5 cm,[30] 3 cm,[2] or greater than 2 fingers' width when measured in a crook lying position.[10]

After analyzing the literature, the present study took into consideration the study of Beer et al[23] because this was the only one that presented cutoff points that were not arbitrary, as they were derived from an analysis of 150 nulliparous women. After stratifying the sample into percentiles, they considered DRAM in those women who were beyond those cutoff points. The US was not evaluated by their research. Then, the only study that measured the US region and presented a cutoff point was Boissonnault and Blaschak.[5]

When analyzing the interrater reliability of the instrument in diagnosing the presence or absence of DRAM, the instrument varied between poor (compared with the USG) to moderate (comparing both digital caliper examiners) agreement. From these data, it is possible to infer that digital calipers may not be the best method to diagnose DRAM in the third trimester of pregnancy compared with USG.

Regarding the ability of digital calipers to measure the IRD, ICC values showed excellent interrater reliability, especially in the US point, when comparing both digital caliper examiners. This corroborates other studies[13,22] that presented an ICC greater than 90 for the instrument. Meanwhile, when comparing digital calipers with USG, the ICC

demonstrated that the instrument has poor agreement. Thus, the contrast between both ICC results shows that ultrasound may be more suitable for measuring the IRD.

When comparing the results of kappa and ICC between digital calipers, the digital caliper seems to be reliable in measuring the IRD, but not reliable in diagnosing DRAM. Compared with USG, however, the instrument does not seem to be reliable either in measuring the IRD or in diagnosing DRAM.

Finally, the Bland-Altman plot corroborates the reliability variables comparing the digital caliper and the USG presented and discussed in the present research. In all 3 points, there was great variability between the 2 instruments, showing no agreement in the measurement of the IRD in the third trimester of pregnancy, compared to the USG. It is possible to see in the plot that in the US the variability was more evident.

The digital caliper's reliability has been already investigated in the literature,[13,15,18] during puerperium, which makes it difficult to compare with the present findings. The studies available[13,18] showed ICCs varying from 0.40 to 0.95, depending on the measurement spot and on the situation in which the measurements were taken (at rest or during contraction). Though below the umbilicus, the instrument presented lower agreement. The authors added that the results indicated that the digital calipers consistently overestimated IRD when compared to USG.

The only systematic[17] review available in the literature searched (PubMed, LILACS, CINAHL, Embase, and Medline) that analyzed digital calipers as instruments to assess DRAM identified the studies discussed above as the only available ones that assessed the digital caliper's reliability, and just 2 of them[13,15] compared digital calipers and USG. The authors identified among studies an ICC of 0.78 to 0.97 for test−retest, interrater, and intrarater reliability for digital calipers. It was also reported that the available information supports ultrasound and digital calipers as adequate methods to assess DRAM, though all 3 studies had an increased risk of bias and it only investigated women in puerperium.

Regarding the digital caliper's accuracy, it was not possible to calculate these data for the SU point because the measurements taken by digital calipers did not reach the cutoff point of 2.2 cm used. Because of the lack of cutoff points in the pregnant women, this value was obtained from a study with women in puerperium, and it may not be suitable for pregnant women.

On the other hand, in the US spot, in spite of presenting a good sensibility and positive predictive value, there was a great variability in specificity and negative predictive value. The IU point, though, did not show good sensibility and specificity. Besides this, when analyzing the positive likelihood ratio and the negative likelihood ratio of these 2 points, the CI reached the null effect, demonstrating that the instrument did not show a good posttest probability in diagnosing DRAM condition in the 2 points measured.

Only 1 study[15] has been found that aimed to determine the accuracy of digital calipers compared to USG to assess DRAM. In this study, the authors evaluated points above the umbilicus (3 cm, 6 cm, 9 cm, and 12 cm), and the cutoff points used to diagnose DRAM were not specified. The variables concerning accuracy were not presented by the authors; they were calculated and presented by other authors in a letter to the editor[31] and in a systematic review[17]: sensibility of 90%, specificity of 75%, positive predictive value of 82.5%, and NPV of 84.6%. The CI was not shown, which made it impossible to interpret these data. In addition, the study had bias concerning the lack of blinding and that the same examiner used the digital caliper and USG.

The authors[17,30] emphasize that the choice of the instrument to assess DRAM should depend on the purpose of the evaluation. When evaluating DRAM width as a treatment outcome, a small measurement error is desired, for which the systematic review recommended ultrasound or digital calipers.

Thus, in cases in which more accurate measurements are necessary, the ultrasound must be preferred to evaluate DRAM. Thus, we suggest the necessity of establishing a reliable reference to define normal values of the IRD during pregnancy, as well as the parameter to consider presence of DRAM, so that the scientific community can better understand it.

Limitations

One limitation of the present research was that there was only 1 examiner handling the USG, and it is known as an operator-dependent instrument. Thus, we strongly suggest that future researchers check the interrater reliability of USG by using 2 examiners with this instrument. Another limitation was the lack of a calculation to estimate the sample size of the study.

Conclusion

From the data presented here, digital calipers were neither accurate to measure the IRD nor to diagnose the DRAM in the third trimester of pregnancy using the cutoff point of 2.2 cm for the SU region, 2.0 cm for the US, and 1.6 cm for the IU spot (used by most of the studies in this field research).

Funding Sources and Conflicts of Interest

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors declare no conflicts of interest.

Contributorship Information

Concept development (provided idea for the research): M. B., A.L., A.M.D., A.M., A.C., V.A.
Design (planned the methods to generate the results): M.B., A.L., A.M.D.
Supervision (provided oversight, responsible for organization and implementation, writing of the manuscript): M.B., A.L., A.M., A.C., V.A
.Data collection/processing (responsible for experiments, patient management, organization, or reporting data): M.B., A.M.D., A.M.
Analysis/interpretation (responsible for statistical analysis, evaluation, and presentation of the results): M.B., A.L,. A.M., A.C., V.A.
Literature search (performed the literature search): M.B., A. L., A.M.D.
Writing (responsible for writing a substantive part of the manuscript): M.B., A.M.D, A.M., A.C., V.A.
Critical review (revised manuscript for intellectual content, this does not relate to spelling and grammar checking): M. B., A.L., A.M.D., A.M.

Practical Applications

- This study found that calipers were not accurate to diagnose DRAM during pregnancy.
- Calipers did not show good interrater reliability in measuring the IRD.
- Calipers did not show good interrater reliability in diagnosing DRAM.
- We suggest that if more accurate measurements are necessary, USG should be considered.

References

1. Mota P, Pascoal AG, Sancho F, Bo K. Test-retest and intrarater reliability of 2D ultrasound measurements of distance between rectus abdominis in women. *J Orthop Sports Phys Ther*. 2012;42(11):940-946.
2. Leite ACNMT, Araújo KKBC. Diástase dos retos abdominais em puérperas e sua relação com variáveis obstétricas. *Fisioter Mov*. 2012;25(2):389-397.
3. Rett MT, Braga MD, Bernardes NO, Andrade SC. Prevalence of diastasis of the rectus abdominis muscles immediately postpartum: comparison between primiparae and multiparae. *Rev Bras Fisioter*. 2009;13(4):275-280.
4. Nahas FX, Ferreira LM, Ely PB, Ghelfond C. Rectus diastasis corrected with absorbable suture: a long-term evaluation. *Aesthetic Plast Surg*. 2011;35(1):43-48.
5. Boissonnault JS, Blaschak MJ. Incidence of diastasis recti abdominis during the childbearing year. *Phys Ther*. 1988;68 (7):1082-1086.
6. Mestak O, Kullac R, Mestak J, Nosek A, Krajcova A, Sukop A. Evaluation of the long-term stability of sheath plication using absorbable sutures in 51 patients with diastasis of the recti muscles: an ultrasonographic study. *Plast Reconstr Surg*. 2012;130(5):714-719.
7. Hichey F, Finch JG, Kihanna A. A systematic review on the outcomes of correction of diastasis of the recti. *Hernia*. 2011;15(6):607-614.
8. Parker MA, Millar AL, Dugan SA. Diastasis rectus abdominis and lumbo-pelvic pain and dysfunction—are they related. *J Wom Health Phys Ther*. 2008;33(2):15-22.
9. Spitznagle TM, Leong FC, Dillen LRV. Prevalence of diastasis recti abdominis in a urogynecological patient population. *Int Urogynecol J Pelvic Floor Dysfunct*. 2007;18(3):321-328.
10. Bursch SG. Interrater reliability of diastasis recti abdominis measurement. *Phys Ther*. 1987;67(7):1077-1079.
11. Mendes DA, Nahas FX, Veiga DF, et al. Ultrasonography for measuring rectus abdominis muscles diastasis. *Acta Cirúrgica Brasileira*. 2007;22(3):182-186.
12. Mota P, Pascoal AG, Sancho F, Carita AI, Bo K. Reliability of the inter-rectus distance measured by palpation. Comparison of palpation and ultrasound measurements. *Man Ther*. 2013;18(4):294-298.
13. Boxer S, Jones S. Intra-rater reliability of rectus abdominis diastasis measurement using dial calipers. *Aust J Physiother*. 1997;43(2):109-114.
14. Lemos A, Souza AI, Andrade AD, Figueiroa JN, Cabral-Filho JE. Pregnancy inter-recti abdominis distance has no impact on respiratory strength. *J Phys Ther Sci*. 2011;23 (5):757-760.
15. Barbosa S, de Sá RA, Coca Velarde LG. Diastasis of rectus abdominis in the immediate puerperium: correlation between imaging diagnosis and clinical examination. *Arch Gynecol Obstet*. 2013;288(2):299-303.
16. Nahas FX. Should diastasis recti be corrected? *Aesthetic Plast Surg*. 1997;21(4):285-289.
17. 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 generalization. *Man Ther*. 2016;21:41-53.
18. 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.
19. Bossuyt PM, Reitsma JB, Bruns DE, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Croat Med J*. 2003;44(5):639-650.
20. Kottner J, Audige´L, Brorson S, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol*. 2011;64(1):96-106. Republished in *Int J Nurs Stud*. 48(6):661-671.
21. Atalah S, Castillo C, Castro R, Aldea A. Proposal of a new standard of nutritional evaluation in pregnant women. *Rev Méd Chile*. 1997;125(12):1429-1436.
22. Liaw L, Hsu M, Liao C, et al. The relationships between inter-recti distance measured by ultrasound imaging and abdominal muscle function in postpartum women: a 6-month follow-up study. *J Orthop Sports Phys Ther*. 2011;41(6):435-443.
23. Beer GM, Schuster A, Seifert B, Manestar M, Mihic-Probst D, Weber SA. The normal width of the linea alba in nulliparous women. *Clin Anat*. 2009;22(6):706-711.

24. Coldron Y, Stokes MJ, Newham DJ, Cook K. Postpartum characteristics of rectus abdominis on ultrasound imaging. *Man Ther*. 2008;13(2):112-121.
25. Pascoal AG, Dionisio S, Cordeiro F, Mota P. Inter-rectus distance in postpartum women can be reduced by isometric contraction of the abdominal muscles: a preliminary case-control study. *Physiotherapy*. 2014;100(4):344-348.
26. Sancho MF, Pascoal AG, Mota P, Bø K. Abdominal exercises affect inter-rectus distance in postpartum women: a two-dimensional ultrasound study. *Physiotherapy*. 2015;101(3):286-291.
27. Mesquita LA, Machado AV, Andrade AV. Fisioterapia para redução da diástase dos músculos retos abdominais no pós-parto. *Rev Bras Ginecol Obstet*. 1999;21(5):267-272.
28. Gilleard WJ, Brown JM. Structure and function of the abdominal muscles in primigravid subjects during pregnancy and the immediate postbirth period. *Phys Ther*. 1996;76(7):750-762.
29. Lo T, Candido G, Janssen P. Diastasis of the recti abdominis in pregnancy: risk factors and treatments. *Physiother Can*. 1999;51:32-37.
30. Candido G, Lo T, Janssen PA. Risk factors for diastasis of the recti abdominis. *J Assoc Charter Physiother Women's Health*. 2005;97:49-54.
31. van de Water ATM, Benjamin DR. Measure DRAM with a purpose: diagnose or evaluate. *Arch Gynecol Obstet*. 2014;289(1):3-4.