Intra- and inter-tester reliability of diaphragm thickness and excursion measurements of healthy women using rehabilitative ultrasound imaging across novice examiners

Anastasia Skoura,[1] Evdokia Billis,[1] Maria Andriopoulou,[2] Elena Drakonaki,[3] Dimitra Tania Papanikolaou,[1] Maria Tsekoura,[1] Eleni Kortianou,[4] Ioannis Maroulis[5]

[1]Laboratory of Clinical Physiotherapy and Research, Department of Physiotherapy, School of Health Rehabilitation Sciences, University of Patras, Greece; [2]General Hospital Asklipieio Voulas, Greece; [3]Anatomy Department, Medical School, University of Crete, Heraklion, Greece; [4]Department of Physiotherapy, University of Thessaly, Greece; [5]Department of Medicine, University of Patras, Greece.

Abstract

Beyond respiration, the diaphragm contributes to trunk stability and posture. Rehabilitative Ultrasound Imaging (RUSI) assesses Diaphragm Thickness (DT) and Excursion (DE), but evidence on reliability across novice examiners is limited. This study investigated intra- and inter-tester reliability of diaphragmatic RUSI across novice examiners in healthy women and compared DT and DE between parous and nulliparous participants. Healthy parous and nulliparous women participated. Exclusion criteria included pregnancy, chronic conditions, and BMI > 30 kg/m2. Two physiotherapists, novices in RUSI, measured DT during full inspiration and expiration and DE during quiet breathing. An independent examiner calculated the Thickening Fraction (TF). Three within-session measurements were taken for intra-tester reliability, whereas, for inter-tester reliability, agreement between mean measurements was analyzed via Intraclass Correlation Coefficients (ICC) and 95% Confidence Intervals (CIs). Independent samples t-tests assessed DT and DE differences between groups. Twenty women (33±14.3 years, BMI: 23.4±3.3) participated. Intra-tester reliability was moderate to good for DT (ICC:0.511-0.691) and very good for DE (ICC:0.700-0.756). Inter-tester reliability was very good for DE (ICC=0.748), and moderate for DT (ICC: 0.573-0.706). TF reliability was poor. Only one tester found higher inspiratory DT in parous women ($p<0.01$). Novice examiners can reliably measure DT and DE using RUSI, supporting its use in rehabilitation. Further standardization is required to enhance RUSI practices and additional research to clarify parity-related diaphragm changes.

Key Words: diaphragm, ultrasound imaging, ultrasonography, RUSI, reliability study.

Eur J Transl Myol 35 (4) 13907, 2025 doi: 10.4081/ejtm.2025.13907

The diaphragm is the main respiratory muscle, working in coordination with other accessory muscles (such as intercostals), to support the work of breathing. Beyond its respiratory function, research highlights the diaphragm's crucial role in trunk stability and posture.[1] Located at the superior border of the "abdominal canister" the diaphragm works in synergy with the pelvic floor and abdominal muscles to modify or maintain intra-abdominal pressure during static and dynamic tasks, thus increasing spinal stability.[1-5] Its unique anatomy with its crurae attached to the lumbar vertebrae is also considered to contribute to spinal stability.[6,7]

Although Rehabilitative Ultrasound Imaging (RUSI) was initially proposed for the assessment of the abdominals, pelvic floor and multifidus muscles,[8,9] over the past decade, RUSI has gained significant popularity among respiratory therapists, Physical and Rehabilitation Medicine (PRM) professionals, physiotherapists, and researchers, offering a safe and cost-effective method to assess diaphragm function as opposed to traditional, high-cost, invasive or other non-invasive assessment methods (*i.e.* dynamic magnetic resonance imaging, fluoroscopy, electromyography, trans-diaphragmatic pressure assessment etc.).[10,11] Utilizing brightness- or motion-modulation, diaphragmatic RUSI is applied to assess diaphragm thickness and excursion parameters in critical care,[12] cardiac,[13,14] and pulmonary rehabilitation,[15,16] as well as within the musculoskeletal field.[17,18] Diaphragm thickness refers to the shortening and thickening of the diaphragm as the muscle fibers contract and relax during breathing, while excursion describes its craniocaudal displacement along its axis during the respiratory cycle. Indeed, reduced diaphragm thickness and lower diaphragm mobility have been observed amongst individuals with non-specific low back,[19-21] and lumbopelvic pain,[22] compared to healthy controls. In postpartum populations, Kharaji *et al.*[23] also found reduced diaphragm excursion in women with lumbopelvic pain compared to asymptomatics, during a low postural demanding task. It is, therefore, not unreasonable to assume that diaphragm thickness and excursion, measured via Ultrasound (US) may have a central role within respiratory and musculoskeletal rehabilitation.

Previous studies confirm strong intra-tester and inter-tester reliability of diaphragm thickness and excursion measurements among clinicians with varying backgrounds and experience levels with ultrasonography.[24-33] However, a crucial concern involves the ability to consistently measure diaphragm thickness and excursion, especially among novice (in RUSI) physiotherapists. Given the rapidly increasing interest of RUSI across physiotherapy, this study aims to fill this knowledge gap, by assessing the intra-tester and inter-tester reliability of physiotherapists' US measurements of diaphragm thickness and excursion across a healthy women sample. A secondary objective is to explore any differences in diaphragm measurements among healthy nulliparous and parous women, as this concept remains under-researched in relation to postpartum recovery and rehabilitation.

Materials and Methods

Study design and registration

This reliability study was carried out from May to October 2023 following the Guidelines for Reporting Reliability and Agreement Studies (GRRAS).[34] The Helsinki Declaration and human experimentation ethical requirements were respected.[35] The study was approved by the Ethics Committee of the University of Patras (number 15543/12-04-2023). All participants signed an informed consent form before participating in the study.

Sample size calculation

An *a priori* sample size calculation was performed using GPowerSoftware 3.1.9.7. A correlation coefficient of 0.6 was chosen, representing a strong association between measurements based on previously reported values.[27] With a significance level of 0.05, a power of 0.80, and a one-tailed hypothesis, the sample size calculation determined a minimum of 15 subjects necessary to detect the specified correlation of 0.6 between intra-tester and inter-tester measurements.

Subjects

A convenience adult women sample around the broader university campus area was invited to participate in the study through advertisements and university announcements. Inclusion criteria were healthy women, nulliparous and parous, aged 18-60. Exclusion criteria were previous diagnosis of severe chronic respiratory conditions (*i.e.*, COPD, chronic bronchitis etc.), Body Mass Index (BMI) > 30 kg/m^2 that could challenge US imaging, previous major abdominal surgery, pregnancy, and neuromuscular disorders that could alter diaphragm behavior and function.

Testers and training procedure

Two physiotherapists, both experienced clinicians (over 7 years of clinical musculoskeletal and respiratory experience), but novices in diaphragmatic RUSI, participated in a two-day training session led by a consultant radiologist, specializing in musculoskeletal sonographic imaging. Training was followed by several hours of independent and joint practice between the radiologist and physiotherapists. Throughout the study assessments, the physiotherapists had access to the radiologist for consultation whenever they were uncertain about a measurement.

Diaphragm measurement procedure

All subjects received US assessment using the Versana Active™ 2-D ultrasound system (GE HealthCare, Chicago, Illinois, US) in the semi-recumbent position (trunk elevated by 35°), hands positioned behind the head, and knees bent (60º of flexion).

Diaphragm Thickness (DT) was assessed by the examiners at the zone of apposition of the right hemi-diaphragm, using a high-resolution linear transducer at 10 MHz in B-mode. The scan was performed over a marked area (Figure 1A), as described in Boussuges et al.[36]. The subject was asked to fully exhale and then inhale as deeply as possible, aiming at full chest and abdominal expansion. DT was measured in frozen B-mode in dual split screen (capturing two images), at the end of full inspiration and full expiration (Figure 1B). Three consecutive measurements were taken during each session. After the assessment, an independent researcher calculated and documented the thickening fraction according to the following formula:

Thickening Fraction (TF) = (Thickness at End Inspiration - Thickness at End Expiration) / Thickness at End Expiration

TF refers to the percentage of change in diaphragm thickness between inspiration and expiration, indicating the extent of diaphragmatic contraction and relaxation during the respiratory cycle. TF can either be expressed as a percentage (multiplied by 100%) or as a quotient from the given fraction, which was employed in this case.

Diaphragm Excursion (DE) was assessed at the right hemidiaphragm, through the liver acoustic window, using a high-resolution curvilinear transducer at 4 MHz in M-mode. To locate the best position for optimal visualization of the posterior hemidiaphragm displacement, B-mode was utilized to scan over the marked skin area (Figure 2A), as described in Boussuges *et al.*[36] Once visualization of the right hemi-

Eur J Transl Myol 35 (4) 13907, 2025 doi: 10.4081/ejtm.2025.13907

diaphragm with B-mode was achieved, M-mode was utilized to display the craniocaudal displacement of the diaphragm through the respiratory cycle, appearing as a hyperechoic line. The subject was asked to inhale and exhale naturally, allowing free movement of the chest and abdomen and the image was captured and measured after 1-3 respiratory cycles (Figure 2B). Three consecutive measurements were taken during each session.

Reliability testing

For intra-tester reliability, anatomical guide points were marked with a skin marker independently by each tester according to literature recommendations.[36] These markings were erased after each tester completed their measurements. Each tester performed six consecutive measurements, three for DT and three for DE with a two-minute resting interval between each measurement. Tester 1 initiated the measurements for half of the sample, while Tester 2 measured the other half first.

For inter-tester reliability, the other physiotherapist repeated the measurements during the same session; three DT and three DE measurements as previously described were measured again in random order. Examiners and subjects were blinded to all measurements taken, with post-it

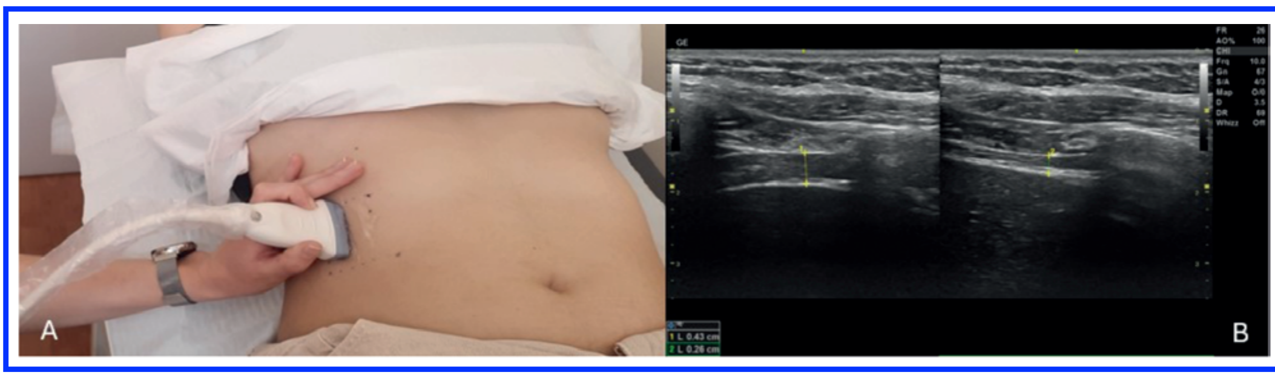

Figure 1. Diaphragm thickness assessment. A) Position of the probe over the marked area, delineated by anatomical guide points. These points indicate the region 0.5 to 2.0 cm beneath the phrenico-costal sinus, positioned between the 8th and 9th intercostal spaces and extending between the anterior and mid-axillary line. Probe position is perpendicular to the chest wall. B) Caliper placement. The diaphragm is identified as a hypoechoic area enclosed between two parallel echogenic lines, the diaphragmatic pleura and the peritoneal fascia. Calipers are placed on the two hyperechoic lines (diaphragmatic pleura and peritoneal fascia) and the distance is calculated automatically by the machine's built-in digital measurement system. Two measurements were taken per respiratory cycle (end of full expiration and full inspiration) using a dual split-screen in B-mode. The image displays diaphragm thickness at full inspiration (left) and full expiration (right). Measurements were repeated across three consecutive cycles, with a two-minute interval between efforts.

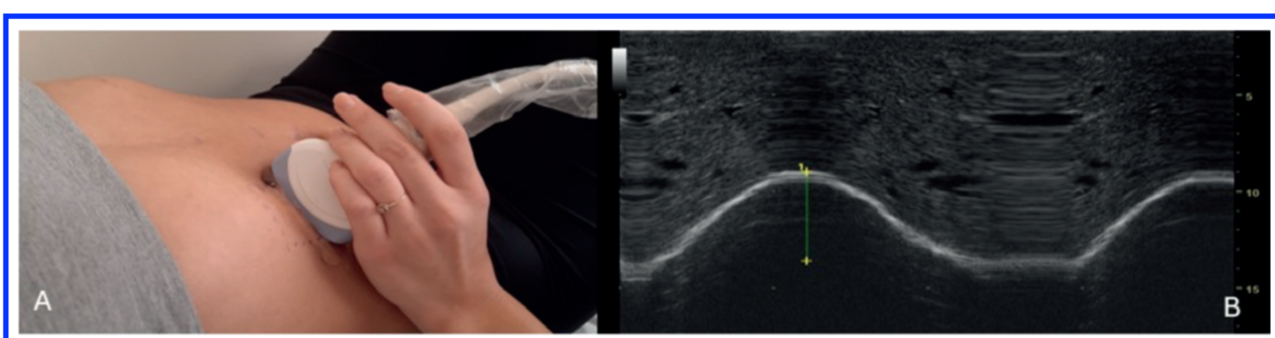

Figure 2. Diaphragm excursion assessment. A) Position of the probe over the marked area, delineated by anatomical guide points. These points indicate the low intercostal area/ area below the subcostal margin, between the mid-clavicular and the mid-axillary line. The probe is pointed medially, dorsally and in a cephalad direction. B) Excursion measurement in M-mode. Calipers are placed at the foot of the slope and at the apex of the hyperechoic diaphragmatic line and their distance is calculated automatically by the machine's built-in digital measurement system. Each image was captured during quiet breathing, typically after 1 to 3 full respiratory cycles, depending on the subject's respiratory rate and how many cycles fit within the image frame. Measurements were taken based on the clearest excursion trace. Image B shows one complete respiratory cycle and a second cycle that ends at inspiration, representing an incomplete trace.

Eur J Transl Myol 35 (4) 13907, 2025 doi: 10.4081/ejtm.2025.13907

stickers placed over the tables displaying distance values at the upper and lower left edges of the US screen. Following examination, an independent researcher blinded to the subject and examination procedure extracted and documented all variables for each subject in a spreadsheet software. The mean of 3 repeated measurements for each testing variable was calculated.

Statistical analysis

Descriptive statistics, including measures of central tendency (range, mean, median, and standard deviation values) as well as frequencies and percentages, were employed for interval and categorical variables, respectively. A Kolmogorov–Smirnov test was used to assess normality. Differences between testers were reported utilizing paired samples t-tests, while differences across parous and nulliparous women were reported utilizing independent samples t-tests. Intra-tester and inter-tester reliability were estimated utilizing an Intraclass Correlation Coefficient (ICC) and 95% Confidence Intervals (CIs) according to literature recommendations.[37] Intra-tester reliability was assessed with a two-way mixed effects model ($ICC_{3,1}$) and inter-tester reliability with a two-way random effects model ($ICC_{2,1}$). The interpretation of ICC values followed established guidelines;[38] thus, values <0.40 were classified as poor, those between 0.40 and 0.59 as weak, between 0.60 to 0.74 as good, and values ranging from 0.75 to 1.00 were considered excellent. The standard error of measurement (SEM), indicating average amount of error of individual measurements, was calculated using the formula $SEM = SD \times \sqrt{(1 - ICC)}$. The smallest detectable difference (SDD), representing the threshold beyond which a real change in measurement is considered, was calculated using the formula $SDD = \sqrt{2} \times 1.96 \times SEM$ for a 95% CI. Statistical analysis was carried out utilizing SPSS (version 28.0).

Results

Twenty women aged 19-59 participated in the study; 11 nulliparous (55%) and 9 parous (45%). The sample's sociodemographic characteristics are presented in Table 1. Statistically significant differences in age across groups were found, as parous women were older than nulliparous women. Each tester captured and measured 180 images. DT, DE and TF descriptives are presented in Table 2

Table 1. *Sociodemographic characteristics of the sample.*

	Women sample (n=20)	**Nulliparous (n=11)** **Mean (SD)**	**Parous (n=9)**	**Group differences** ***P*-value**
Age (years)	33.1 (14.29)	24.3 (11.32)	43.8 (9.4)	<0.001*
Height (cm)	165.0 (6.6)	165.1 (5.7)	164.8 (8.0)	0.92
Weight (kg)	63.7 (9.8)	63.8 (10.1)	63.5 (10.0)	0.95
BMI	23.4 (3.3)	23.5 (4.1)	23.3 (2.2)	0.88

*SD, standard deviations. * Statistically significant differences between groups.*

Table 2. *Diaphragm parameter measurements across the two testers.*

Parameters	**Tester 1**	**Tester 2**	***P*-value**
	Mean (SD)		
Inspiratory Thickness (cm)	0.38 (0. 06)	0.40 (0.09)	0.224
Expiratory Thickness (cm)	0.24 (0.05)	0.23 (0.07)	0.706
Thickening Fraction	0.62 (0.31)	0.78 (0.42)	0.106
Excursion (cm)	4.66 (1.23)	4.28 (1.16)	0.058

SD, standard deviations.

whereas, Table 3 summarizes descriptive statistics across nulliparous and parous women. No statistically significant differences were found in DT, TF, or DE between testers across the sample ($p>0.05$). No statistically significant differences in diaphragm parameters were found across nulliparous and parous women except for Tester 2's inspiratory thickness measurements, where a statistically significant difference was evident ($p<0.01$).

Intra-tester reliability

Intra-tester reliability results for each tester are presented in Table 4. Tester 1 demonstrated a moderate to good agreement for DT measurements, both in inspiratory (ICC=0.511) and expiratory phases (ICC= 0.617). Tester 2 exhibited good intra-tester agreement, with inspiratory and expiratory DT yielding slightly higher agreement coefficients (ICC=0.685, and ICC=0.691, respectively), compared to Tester 1. SEM calculations revealed relatively low variability in inspiratory and expiratory thickness measurements for both, Tester 1 (with SEM values of 0.045 and 0.032, respectively) and Tester 2 (with SEM values of 0.055 and 0.045, respectively). SDD, representing the threshold beyond which a genuine change in measurement is considered, was acceptable, falling within the expected range of measurements (between 0.12cm and 0.15cm for inspiratory, and 0.09cm and 0.12cm for expiratory thickness).

For DE, both testers demonstrated very good intra-tester agreements (ICCs>0.700). For Tester 2, US DE measurements demonstrated low variability, with small SEM and SDD (0.045cm and 0.12cm respectively). Tester 1 exhibited a moderate level of precision, with slightly larger SEM and SDD values (0.66cm and 1.83cm, respectively). TF revealed relatively low intra-tester agreement for both Tester 1 (ICC=0.481) and Tester 2 (ICC=0.501), larger SEM (0.276cm and 0.367cm, respectively), and SDD values (0.764 cm and 1.018 cm, respectively).

Inter-tester reliability

Inter-tester reliability results across the two testers are presented in Table 5. Inter-tester agreement on DT was deemed good for inspiratory thickness (ICC=0.706), moderate for expiratory thickness (ICC=0.573) whereas SEM and SDD values were relatively small across testers. Inter-

Table 3. *Diaphragm parameter measurements across nulliparous and parous women.*

Parameters	Tester 1 Mean (SD)		Tester 1 *P*-value	Tester 2 Mean (SD)		Tester 2 *P*-value
	Nulliparous	Parous		Nulliparous	Parous	
Inspiratory thickness	0.36 (0.04)	0.40 (0.07)	0.061	0.35 (0.06)	0.45 (0.09)	0.01*
Expiratory thickness	0.23 (0.05)	0.25 (0.06)	0.484	0.21 (0.05)	0.26 (0.08)	0.067
Thickening fraction	0.56 (0.22)	0.70 (0.39)	0.324	0.77 (0.49)	0.80 (0.34)	0.895
Excursion	4.93 (1.26)	4.33 (1.18)	0.288	4.55 (1.05)	3.95 (1.27)	0.265

*SD, standard deviations. * Statistically significant differences between groups.*

Table 4. *Intra-tester reliability results (n=20).*

Parameters	Tester 1 ICC	Tester 1 95% CI Lower - Upper	Tester 1 SEM	Tester 1 SDD	Tester 2 ICC	Tester 2 95% CI Lower - Upper	Tester 2 SEM	Tester 2 SDD
Inspiratory thickness	0.511	0.100 - 0.773	0.045	0.124	0.685	0.461 - 0.847	0.055	0.152
Expiratory thickness	0.617	0.251 - 0.828	0.032	0.088	0.691	0.470 - 0.851	0.045	0.124
Thickening fraction	0.481	0.210 - 0.724	0.276	0.764	0.501	0.232 - 0.737	0.367	1.018
Excursion	0.756	0.558 - 0.888	0.662	1.834	0.700	0.482- 0.855	0.045	0.124

ICC, intraclass correlation coefficients; CI, confidence interval; SEM, standard error of measurement; SDD, smallest detectable difference.

Table 5. Inter-tester reliability results.

Parameter	ICC	95% CI		SEM	SDD
		Lower	Upper		
Inspiratory thickness	0.706	0.393	0.872	0.045	0.124
Expiratory thickness	0.573	0.186	0.806	0.045	0.124
Thickening fraction	0.346	-0.102	0.677	0.297	0.822
Excursion	0.748	0.466	0.892	0.600	1.665

ICC, intraclass correlation coefficients; CI, confidence interval; SEM, standard error of measurement; SDD, smallest detectable difference.

tester agreement on DE was very good (ICC=0.748); however, SEM and SDD values yielded some measurement variability across testers (0.6cm and 1.67cm, respectively). Inter-tester agreement for TF measurements was poor (ICC=0.346), revealing substantial variability between testers. Larger SEM and SDD values (0.3cm and 0.82cm, respectively) were also reported across testers.

Discussion

This study investigated intra-tester and inter-tester reliability of measuring DT and DE using RUSI among two novice examiners, following a standardized protocol, in a cohort of healthy women including both nulliparous and parous individuals. Actual findings by both testers were more or less in line with values reported in previous studies for DT and DE in healthy populations.[25,27,30] The incorporation of nulliparous and parous women was intended to provide a foundational assessment for future research in postpartum populations and explore potential differences in diaphragm function related to childbirth. However, the sample lacked diversity in terms of BMI, as participants with a lower BMI (<30 kg/m^2) were intentionally enrolled, aiming to minimize the challenges associated with obtaining less clear US images which are often linked to higher BMI levels. This practice aligns with previous reliability studies.[30,31,33]

Both intra- and inter-tester reliability yielded overall satisfactory results across all parameters except for TF. Intra-tester reliability for DT measurements in both inspiratory and expiratory phases yielded good agreement for repeated measurements by the same tester. However, DT measurements appeared to be somewhat lower than those reported in previous studies.[25,26,28,30,31] SEM and SDD for intra-tester measurements aligned with previous studies,[30] indicating a similar level of precision and error. For DE, both testers yielded very good agreements, consistent with prior research.[29,30]

Similar agreement patterns were reported in inter-tester reliability, with a slightly lower agreement for expiratory thickness. Nevertheless, despite lower agreement compared to previous studies,[25,26,31] SEM and SDD values remained relatively low, indicating minimal errors in individual measurements. Additionally, between-tester agreement for DE was close to excellent (ICC=0.748), similar to that reported by Scarlata *et al.*[28,29] for quiet breathing excursion. The observed low agreement in TF, both within repeated measurements and between testers, aligns with previous findings,[33] indicating difficulties in achieving consistent measurements.

Regarding parous and nulliparous women, no significant differences were observed in all but one measurement. Interestingly, inspiratory DT was marginally increased in parous women, reaching statistical significance with one tester, and yielding a close but non-significant *p*-value with the other one. In view of the small sample size, however, (only 9 parous women), this finding should be interpreted with caution. Additionally, as the study was not powered for between-group comparisons, these findings should be considered exploratory. Previous research has reported reduced DE in postpartum women with lumbopelvic pain, but no differences in DT, during a challenging postural task. Hence, future work should further explore these associations between parity and diaphragm function through larger, adequately powered samples.

While this study provides valuable insights, it is essential to acknowledge certain issues that may influence the interpretation and generalizability of our findings. Firstly, the relatively small sample size and the use of only two testers represent a notable limitation. Despite sample size calculations and the reliability achieved in this study, future studies should incorporate larger samples and more examiners to strengthen the evidence. Another limitation is the lack of validity testing against an expert examiner, as this was not within this study's scope. Future studies should consider including expert validation to provide a more comprehensive understanding of the accuracy and consistency of diaphragm measurements.

The sample's relatively lower BMI as previously mentioned (due to difficulties in visualizing the diaphragm in individuals with higher BMI, particularly among novice sonographers) could also limit the study's generalizability.

Consequently, these findings may not directly apply to higher BMI populations. However, the focus of this preliminary study was on healthy women within a specific demographic aimed at developing initial insights into the reliability of diaphragm measurements amongst novice clinicians.

Our reliability may be slightly lower than studies involving experienced sonographers,[25,28,30,31] making our findings especially relevant however, to novice users employing US imaging for rehabilitation purposes. Additional training is certainly essential to refine testers' skills in diaphragmatic US imaging, improving accuracy and reliability in DT and DE measurements. However, TF may not be a suitable assessment outcome among novice practitioners. This could be attributed to several factors like measurement variability, where small errors in static thickness measurements contribute to the variable dynamic calculation of TF, and breathing variability of the subject, where slight inconsistencies in breath depth affect DT. These findings suggest that TF values require more advanced skills, being better suited for use by experienced sonographers or across highly standardized research settings. For novice practitioners, prioritizing simpler and more direct indicators, such as DE or static DT may provide more reliable and clinically valuable information during training and early clinical application.

Subject fatigue and the need for standardized subject effort also require consideration. Subject fatigue was observed during the deep breathing phase, requiring three maximal inspiratory efforts (for DT measurement), as well as during quiet breathing (for DE measurements), requiring 1-3 respiratory cycles per captured image. This might have introduced some measurement variability, and it may have contributed to the slightly lower reliability in diaphragm measurements performed by the testers. Although verbal instructions were consistently repeated across trials, further standardization of the method, perhaps by training subjects with an incentive spirometer as biofeedback to generate the same volume during each respiratory effort, as described by Orde *et al*.[27,] could enhance between-measurement and between-tester agreement.

These findings hold significant implications for respiratory muscle rehabilitation across diverse settings, including pregnant and postpartum women, athletes, and those with trunk dysfunctions, supporting the use of diaphragmatic RUSI by physiotherapists, who typically have less professional experience in US imaging techniques compared to other clinicians.[39] They show that diaphragmatic RUSI can be a quick and feasible assessment method that can also be effectively applied to patient populations. It should be noted that initially, capturing all six images required approximately 30 minutes per tester. There was an observable learning curve; however, with increased experience and practice, the time required decreased to approximately 20 minutes. This progression indicates that the method is both feasible and efficient for quick clinical assessments. Additionally, the involvement of a consultant radiologist was initially necessary to ensure image quality, but her assistance became less needed as testers gained proficiency. This finding supports the potential for broader implementation in clinical and research settings, particularly for novice users.

Thus, healthcare professionals unfamiliar with diaphragmatic US measurements can become proficient through structured hands-on training and standardized protocols provided by experienced sonographers. Incorporating RUSI educational modules into formal academic curricula, whether in physiotherapy programs, Physical Rehabilitation Medicine (PRM) residency training, or through continuing professional education, with supervised hands-on practice and performance evaluation under the supervision of US experts, could help bridge the gap between theoretical knowledge and clinical application, expanding diaphragmatic RUSI accessibility across various clinical settings.[11] Adherence to clearly defined protocols can also ensure consistency, reduce variability, and accelerate skill acquisition. Furthermore, standardized procedures, especially during tidal breathing or full inspiration/expiration maneuvers, can improve inter-tester reliability. As this study showed, even brief but well-structured training programs can yield reliable results when standardized protocols are consistently applied.

Understanding the reliability of DT and DE measurements, can shape therapeutic interventions and guide clinical decision-making. Research suggests that training respiratory muscles using tapered flow resistive or threshold loading devices may benefit spinal stabilization and improve symptoms such as low back pain.[17,40,41] Stabilization exercises focused on inner unit muscles may also impact diaphragm thickness and mobility.[42,43] DT and DE assessment using RUSI may thus contribute to an improved understanding of trunk muscle synergies and enable the development of tailored rehabilitation interventions. Its potential applications may also contribute to treating a wide range of conditions, including respiratory, musculoskeletal, and urogynecological dysfunctions, providing clinicians and researchers with precise information about intervention efficacy. In postpartum care, where diaphragmatic breathing exercises are common, our study provides insights into the diaphragm's role in stability. Thus, incorporating US assessments into postpartum rehabilitation can tailor interventions based on reliable measurements, enhancing overall trunk function.

Conclusions

This study provides valuable insights into the good intra-tester and inter-tester reliability obtained by novice examiners in assessing diaphragm thickness and excursion with US across both nulliparous and parous healthy women. These findings add to the growing body of literature supporting the potential application of RUSI through measurements of diaphragm thickness and excursion in respiratory or musculoskeletal/postpartum rehabilitation and broader clinical contexts. They also contribute to establishing optimal practices and emphasize the significance of standardized protocols for meaningful and reliable results. Further research with larger and more diverse populations is needed to explore RUSI's applicability in clinical conditions beyond this study's scope.

List of abbreviations

COPD, Chronic Obstructive Pulmonary Disease
DE, Diaphragm Excursion
DT, Diaphragm Thickness
RUSI, Rehabilitative Ultrasound Imaging
SDD, Smallest Detectable Difference
SEM, Standard Error of Measurement
TF, Thickening Fraction
US, Ultrasound/Ultrasonography

Contributions

AS, EB, conception, writing - original draft preparation; AS, EB, MA, ED, EK, and IM, design, writing - review and editing; AS, EB, DTP and MT, Formal analysis and investigation; MT, EK and IM, resources; EB, ED and IM, supervision.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript. Publication charges were covered by the department of Physiotherapy, University of Patras.

Conflict of interest

The authors have no relevant financial or non-financial interests to disclose.

Ethics approval and consent to participate

The study was approved by the Ethics Committee of the University of Patras (number 15543/12-04-2023). All participants signed an informed consent form before participating in the study.

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

Conference presentation

Some preliminary data of the study were presented in a poster at the 31st Annual Panhellenic Physiotherapy Congress, held at the War Museum, Athens on Dec 1-3, 2023.

Corresponding author

Anastasia Skoura, Laboratory of Clinical Physiotherapy and Research, Department of Physiotherapy, School of Health Rehabilitation Sciences, University of Patras, Building B, Central Campus, 26504, Rio-Patras, Achaia, Greece.
Tel.: +30.2610.962406
ORCID ID: 0009-0007-4446-3970
E-mail: a.skoura@upatras.gr

Co-authors

Evdokia Billis
ORCID ID: 0000-0002-5828-809X
E-mail: billis@upatras.gr

Maria Andriopoulou
ORCID ID: 0000-0001-9570-3742
E-mail: mariandriop@yahoo.gr

Elena Drakonaki
ORCID ID: 0000-0002-3629-3277
E-mail: drakonaki@yahoo.gr

Dimitra Tania Papanikolaou
ORCID ID: 0000-0002-9453-5238
E-mail: dpapanikolaou@upatras.gr

Maria Tsekoura
ORCID ID: 0000-0003-2649-0522
E-mail: mariatsekoura@upatras.gr

Eleni Kortianou
ORCID ID: 0000-0003-2214-5545
E-mail: ekortianou@uth.gr

Ioannis Maroulis
ORCID ID: 0000-0002-6221-6099
E-mail: ioan.maroulis@gmail.com

References

1. Hodges PW, Butler JE, McKenzie DK, Gandevia SC. Contraction of the human diaphragm during rapid postural adjustments. J Physiol 1997;505:539–48.
2. Hodges PW. Is there a role for transversus abdominis in lumbo-pelvic stability? Man Ther 1999;4:74-86.
3. Hodges PW, Gandevia SC. Activation of the human diaphragm during a repetitive postural task. J Physiol 2000;522:165-75.
4. Hodges PW, Cresswell AG, Daggfeldt K, Thorstensson A. In vivo measurement of the effect of intra-abdominal pressure on the human spine. J Biomech 2001;34:347-53.
5. Hodges PW, Sapsford R, Pengel LH. Postural and respiratory functions of the pelvic floor muscles. Neurourol Urodyn 2007;26:362-71.
6. Shirley D, Hodges PW, Eriksson AE, Gandevia SC. Spinal stiffness changes throughout the respiratory cycle. J Appl Physiol (1985) 2003;95:1467-75.
7. Hodges PW, Eriksson AE, Shirley D, Gandevia SC. Intra-abdominal pressure increases stiffness of the lumbar spine. J Biomech 2005;38:1873-80.
8. Prentice CLS, Milanese S, Massy-Westropp N, Maranna S. The reliability of rehabilitative ultrasound to measure lateral abdominal muscle thickness: A systematic review and meta-analysis. Musculoskelet Sci Pract 2021;53:102357.
9. Fernández-Carnero S, Martin-Saborido C, Achalandabaso Ochoa-Ruiz de Mendoza A, et al. The role of re-

habilitative ultrasound imaging technique in the lumbopelvic region as a diagnosis and treatment tool in physiotherapy: systematic review, meta-analysis and meta-regression. J Clin Med 2021;10:5699.
10. Sferrazza Papa GF, Pellegrino GM, Di Marco F, et al. A review of the ultrasound assessment of diaphragmatic function in clinical practice. Respiration 2016; 91:403-11.
11. Papathanasiou J, Panayotov K, eds. Essentials of physical and rehabilitation medicine for undergraduate medical students. 2nd ed. Ruse, Bulgaria: Avangard Print Ltd; 2023.
12. Takahashi Y, Morisawa T, Okamoto H, et al. Diaphragm dysfunction and ICU-acquired weakness in septic shock patients with or without mechanical ventilation: a pilot prospective observational study. J Clin Med 2023;12:5191.
13. Andriopoulou M, Dimaki N, Kallistratos MS, et al. Skeletal muscle alterations and exercise intolerance in heart failure with preserved ejection fraction patients: ultrasonography assessment of diaphragm and quadriceps. Eur J Heart Fail 2022;24:729-31.
14. Maranta F, Cianfanelli L, Rizza V, et al. Diaphragm dysfunction after cardiac surgery: insights from ultrasound imaging during cardiac rehabilitation. Ultrasound Med Biol 2022;48:1179-89.
15. Shiraishi M, Higashimoto Y, Sugiya R, et al. Diaphragmatic excursion is correlated with the improvement in exercise tolerance after pulmonary rehabilitation in patients with chronic obstructive pulmonary disease. Respir Res 2021;22:271.
16. Altmann CH, Zvonova E, Richter L, Schüller PO. Pulmonary recovery directly after COVID-19 and in Long-COVID. Respir Physiol Neurobiol 2023;315: 104112.
17. Finta R, Nagy E, Bender T. The effect of diaphragm training on lumbar stabilizer muscles: a new concept for improving segmental stability in the case of low back pain. J Pain Res 2018;11:3031-45.
18. Vicente-Campos D, Sanchez-Jorge S, Terrón-Manrique P, et al. The main role of diaphragm muscle as a mechanism of hypopressive abdominal gymnastics to improve non-specific chronic low back pain: a randomized controlled trial. J Clin Med 2021;10:4983.
19. Ziaeifar M, Sarrafzadeh J, Noorizadeh Dehkordi S, Arab AM, Haghighatkhah H, Zendehdel Jadehkenari A. Diaphragm thickness, thickness change, and excursion in subjects with and without nonspecific low back pain using B-mode and M-mode ultrasonography. Physiother Theory Pract 2022;38:2441-51.
20. Mohan V, Paungmali A, Sitilerpisan P, et al. Respiratory characteristics of individuals with non-specific low back pain: A cross-sectional study. Nurs Health Sci 2018;20:224-30.
21. Perez AM, Fernández-Carnero S, Sicilia-Gomez-de-Parada C, et al. Diaphragmatic activation correlated with lumbar multifidus muscles and thoracolumbar fascia by b-mode and m-mode ultrasonography in subjects with and without non-specific low back pain: a pilot study. Medicina (Kaunas) 2023;59:315.
22. Calvo-Lobo C, Almazán-Polo J, Becerro-de-Bengoa-Vallejo R, et al. Ultrasonography comparison of diaphragm thickness and excursion between athletes with and without lumbopelvic pain. Phys Ther Sport 2019;37:128-37.
23. Kharaji G, ShahAli S, Ebrahimi Takamjani I, et al. Ultrasound assessment of the abdominal, diaphragm, and pelvic floor muscles during the respiratory and postural tasks in women with and without postpartum lumbopelvic pain: a case-control study. Int Urogynecol J 2023;34:2909-17.
24. Baldwin CE, Paratz JD, Bersten AD. Diaphragm and peripheral muscle thickness on ultrasound: intra-rater reliability and variability of a methodology using non-standard recumbent positions. Respirology 2011;16: 1136-43.
25. Harper CJ, Shahgholi L, Cieslak K, et al. Variability in diaphragm motion during normal breathing, assessed with B-mode ultrasound. J Orthop Sports Phys Ther 2013;43:927-31.
26. Boon AJ, Harper CJ, Ghahfarokhi LS, et al. Two-dimensional ultrasound imaging of the diaphragm: quantitative values in normal subjects. Muscle Nerve 2013;47:884-89.
27. Orde SR, Boon AJ, Firth DG, Villarraga HR, Sekiguchi H. Use of angle-independent m-mode sonography for assessment of diaphragm displacement. J Ultrasound Med 2016;35:2615-21.
28. Scarlata S, Mancini D, Laudisio A, Raffaele AI. Reproducibility of diaphragmatic thickness measured by M-mode ultrasonography in healthy volunteers. Respir Physiol Neurobiol 2019;260:58-62.
29. Scarlata S, Mancini D, Laudisio A, Benigni A, Antonelli Incalzi R. Reproducibility and clinical correlates of supine diaphragmatic motion measured by m-mode ultrasonography in healthy volunteers. Respiration 2018;96:259-66.
30. Nassiri K, Abedi M, Manshadi FD, et al. Comparison of the reliability of sonographic measurements of diaphragm thickness and mobility in individuals with and without pelvic girdle pain. Iran Red Crescent Med J 2019; 21:e97449.
31. Marugán-Rubio D, Chicharro JL, Becerro-de-Bengoa-Vallejo R, et al. Concurrent validity and reliability of manual versus specific device transcostal measurements for breathing diaphragm thickness by ultrasonography in lumbopelvic pain athletes. Sensors (Basel) 2021;21:4329.
32. Cappellini I, Picciafuochi F, Bartolucci M, et al. Evaluation of diaphragm thickening by diaphragm ultrasonography: a reproducibility and a repeatability study. J Ultrasound 2021;24:411-16.
33. Wilches-Luna EC, Pabón-Muñoz FE, Arias-Campo JM, Caballero-Lozada AF. Inter-rater reliability of the measurement of diaphragmatic excursion and fraction of diaphragmatic thickening by ultrasonography in healthy volunteers. Rev Esp Anestesiol Reanim (Engl Ed) 2022;69:536-43.
34. Kottner J, Audigé L, Brorson S, et al. Guidelines for reporting reliability and agreement studies (GRRAS)

were proposed. J Clin Epidemiol 2011;64:96-106.
35. Holt, GR. Declaration of Helsinki-the world's document of conscience and responsibility. South Med J 2014;107:407.
36. Boussuges A, Rives S, Finance J, Brégeon F. Assessment of diaphragmatic function by ultrasonography: Current approach and perspectives. World J Clin Cases 2020;8:2408-24.
37. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15:155-63. Erratum in: J Chiropr Med 2017;16:346
38. Hallgren KA. computing inter-rater reliability for observational data: an overview and tutorial. Tutor Quant Methods Psychol 2012;8:23-34.
39. Whittaker JL, Ellis R, Hodges PW, et al. Imaging with ultrasound in physical therapy: What is the PT's scope of practice? A competency-based educational model and training recommendations. Br J Sports Med 2019;53:1447-53.
40. Janssens L, McConnell AK, Pijnenburg M, et al. Inspiratory muscle training affects proprioceptive use and low back pain. Med Sci Sports Exerc 2015;47:12-19.
41. Ahmadnezhad L, Yalfani A, Gholami Borujeni B. Inspiratory muscle training in rehabilitation of low back pain: a randomized controlled trial. J Sport Rehabil 2020;29:1151-58.
42. O'Sullivan PB, Beales DJ. Changes in pelvic floor and diaphragm kinematics and respiratory patterns in subjects with sacroiliac joint pain following a motor learning intervention: a case series. Man Ther 2007;12: 209-18.
43. Dülger E, Bilgin S, Bulut E, et al. The effect of stabilization exercises on diaphragm muscle thickness and movement in women with low back pain. J Back Musculoskelet Rehabil 2018;31:323-29.

Disclaimer

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article or claim that may be made by its manufacturer is not guaranteed or endorsed by the publisher.

Submitted: 7 April 2025.
Accepted: 25 July 2025.
Early access: 27 October 2025.