INTRARATER AND INTERRATER RELIABILITY OF ABDOMINAL DRAWING-IN TEST IN ASYMPTOMATIC INDIVIDUALS

EVALUAREA GRADULUI DE ÎNCREDERE INTERGRUP ȘI INTRAGRUP A TESTULUI DE VACUUM ABDOMINAL, LA PACIENȚII ASIMPTOMATICI

Jogani Khushali Dilipbhai[1], Bid Dibyendunarayan[2], Thangamani Ramalingam A.[3]

Key words: transversus abdominis (TrA), pressure biofeedback unit (PBU), reliability.

Introduction
The TrA stabilizes the back and back pain adversely affects the activation and endurance capacity of TrA. TrA is local and deepest muscle and abdominal drawing in test by PBU provides an indirect way of evaluating endurance capacity of TrA muscle activity which is often used by clinicians and researchers.
Aim. This is across-sectional study to investigate the intra-rater and inter-rater reliability of Abdominal drawing-in test (ADIT) in asymptomatic individual.
Methods. Sixty asymptomatic subjects were randomly selected for the study. The ADIT was measured for each subject with PBU by the two raters for inter-rater reliability and by one of the rater after a gap of seven days for intra-rater reliability. All the subjects were previously taught and compensations were corrected.
Results. The study demonstrated intra-class correlation coefficient (ICC) with standard error of mean (SEM) of 0.944 and 0.69725 for inter-rater reliability and 0.910 and 0.85814 for intra-rater reliability. A Bland-Altman limit of agreement has also confirmed that inter-rater and intra-rater were within the limits of agreement in 95% of occasions.
Conclusions. ADIT has high inter-rater and intra-rater reliability in asymptomatic individuals.

Cuvinte cheie: transvers abdominal (TrA), unitate de presiune cu biofeedback (PBU), grad de încredere.

Introducere
Transversul abdominal stabilizează spatele iar durerea lombară inversează activarea și rezistența transversului abdominal. Transversul este un mușchi profund iar testul de vacuum abdominal cu PBU oferă o modalitate indirectă de evaluare a rezistenței transversului abdominal, fiind adeseori folosit de clinicieni și cercetători.
Scop. Este un studiu transversal având ca scop investigarea gradului de încredere al testului de vacuum abdominal (ADIT) la indivizii asimptomatici.
Metode. Șaizeci de subiecți asimptomatici au fost selectați randomizat pentru acest studiu. ADIT s-a măsurat pentru fiecare pacient cu PBU de către doi evaluatori pentru gradul de încredere intergroup și după o săptămână s-a efectuat reevaluarea de către un singur evaluator pentru tastarea intragrup. Toți subiecții au fost învățați anterior manevra corectă și compensațiile corectate.
Rezultate. Studiul a demonstrat un coeficient de corelație intra-clasă (ICC) cu o eroare standard (SEM) 0.944 și 0.69725 pentru gradul de încredere intergroup și 0.910 și 0.85814 pentru gradul de încredere intragrup. Limita Bland-Altman a confirmat că valorile intergroup și inatragrup sunt în gradul de confidență în 95% din situații.
Concluzii. ADIT are un grad mare de încredere intergroup și intragrup la pacienții asimptomatici.

[1] *MPT (Musculoskeletal Sciences)*
[2] *MPT(Ortho), PGDSPT*
Corresponding Author; *dnbid71@gmail.com*; Mobile: 09427139711.
Department of Physiotherapy, Sarvajanik College of Physiotherapy, Veer Narmad South Gujarat University, Surat-395003, Gujarat, INDIA
[3] *BPT, MSc(Psy), PGDRM,* Department of Physiotherapy, Sarvajanik College of Physiotherapy, Veer Narmad South Gujarat University, Surat-395003, Gujarat, INDIA

Introduction

Low back pain is one of the most common health problems and creates a substantial personal, community, and financial burden globally. [1-3] Low back pain was defined as pain limited to the region between lower margins of 12th rib and gluteal folds with or without leg pain (sciatica). [4] Lumbar spine is more mobile than the thoracic spine but stability is also a very important feature of the lumbar spine. For load transfer stability is required throughout the entire range of motion and this is provided by the active system i.e. muscles. [5].

Bergmark has categorized the trunk muscles into local and global muscle systems based on their main mechanical roles in stabilization. Anatomically, the deep muscles of the local system are capable of making a major contribution to spinal stability, being closer to the center of rotation of the spinal segments and, with their shorter muscle lengths, they are ideal for controlling intersegmental motion. [6] The smaller intersegmental muscles, such as the intertransversarii and interspinales, may not predominate as mechanical stabilizers but have a proprioceptive role instead. Overlapping multisegmental muscles linking adjacent lumbar vertebrae and the sacrum, such as the lumbar multifidus, have the capacity to, and have been shown to be efficient in stabilizing the spinal segment. [7,8] The TrA has also shown to contribute to this function of segmental stability. [9]

These muscles could be dysfunctional in back pain patients. The local muscles may not be able to maintain prolonged or sustained muscle contraction in order to protect continuously any unstable spinal segments, which could leave the low back pain patient vulnerable to persistent strain and pain. [9]

Cholewicki & McGill's modelnot only highlighted the prime role of local muscles in spinal stabilization at high loads, it also pointed to the importance of the local system in providing spinal support during low-load activities requiring only low muscle forces. [10] When the TrA contracts bilaterally it produces a drawing-in of the abdominal wall, resulting in an increased pressure within the abdominal cavity [11] and an increase in tension in the thoracolumbar fascia. [12] The concept behind the strengthening of local system is to create stiffness in the spine before load is placed on it, thus controlling mid-range or neutral zone of the inter-vertebral joints. Control of this mid-range helps decrease shear forces and compression during movement and spinal loading. When working properly, the local intrinsic musculature fires before the actual motion of an extremity or trunk. Weaknesses of these muscles decrease the person's ability to control joint neutral position during movement or under load and hence can lead to spinal instability.

Tools have been designed to measure a person’s ability for recruitment of the TrA muscle. It has been divided into clinical test and laboratory test. Clinical test involves the recruitment by palpation[13] and by PBU. [14] But palpation test will be subjective so it requires skill of physiotherapist. Moreover with PBU, objective measurement can be done.

Laboratory test includes ultrasound imaging measures from a pressure sensor, EMG and surface electromyography. [15] Most of the studies that have measured the activity of the deep abdominal wall muscles used fine-wire electromyography. However, this type of assessment is invasive, painful, uncomfortable, and expensive and may present the risk of infection. [16]

Test should be done with teaching the patients in four point kneeling and then test should be conducted in prone lying with PBU (Stabilizer, Chattanooga, USA). [17] It is a reliable and valid clinical instrument for assessing deep abdominal muscle function, and has been used to develop a method for the careful monitoring of lumbar stabilization. [18,19]

Once the patient has contracted the TrA than endurance can be checked by maintaining the contraction and holding it for 10s upto maximum of 10 repetitions[20]. The outcome measure used is **Performance index (PI)**. Performance index can be defined as activation score (pressure level the subject is able to achieve)*number of successful repetitions. The outcome measure was developed by Jull [19] in which endurance of deep cervical muscle was measured using PBU.

Hence it is used to measure the endurance of TrA.

Among all of the reliability studies, two studies. [21,22] were available evaluating the reliability in asymptomatic individuals. These studies have evaluated only intra-rater reliability and even sample size was small. And among all the studies [21-24] available on asymptomatic individuals and chronic back pain patients, they have measured only recruitment of TrA and there is lack of literature measuring the endurance of TrA. The systematic review done by Lima *et al.* [25] has said that the measurement properties of PBU for TrA activity are yet to be answered.

Purpose

The purpose of this study is to test the intra-rater and inter-rater reliability of ADITin asymptomatic individual by using Performance Index as outcome measure.

Matherial and Methods

In this study, 60 asymptomatic subjects were studied. Sample size was calculated based on test-retest designs, and agreement between the raters. According to that if assumptions kept as the observed R will be 0.80 or greater with a lower 1-sided 95% confidence interval i.e. CI=0.10 (i.e., R acceptable ≥ 0.70). Therefore 55 subjects are required. And with 5% drop out rate, total of 59 subjects are required. Thus, total 60 subjects were evaluated with no drop outs[26, 27].

Inclusion criteria included: (a) Age: 18 to 25 years; (b) Both males and females; & (c) Body mass index ≤ 24. Exclusion criteria included: (a) History of back pain or current back pain; (b) Pregnancy; (c) Menstruation on the test days; (d) Any trauma to lower back; (e) Any abdominal wall or spinal surgeries; (d) Confirmed serious pathologies; (e) Inability to contract the abdominal muscles; (f) Pressure reduction of less than 3 mmHg; (g) Inability to lie in prone; & (h) Cardiovascular or respiratory problem.

Figure 1: Pressure Biofeedback Unit

Figure 2: Stop Watch

The PBU is a reliable and valid clinical instrument for assessing deep abdominal muscle function, and has been used to develop a method for the careful monitoring of lumbar stabilization[18, 19]It is also utilized in previous studies for measuring the activation of TrA.

The PBU is a simple pressure transducer consisting of a three-chamber air-filled pressure bag, a catheter and a sphygmomanometer gauge. The pressure bag has 16.7×24 cm in size and made from non-elastic material. The sphygmomanometer scale ranges from 0 mmHg to 200 mmHg, with 2 mmHg intervals on the scale. The accuracy of the apparatus is described as ± 3 mmHg. Movement or change in position causes volume changes in the pressure bag, which is registered by this device(28). The outcome measure used isPerformance Index (PI)[29].

Performance index can be defined as activation score (pressure level the subject is able to achieve)*number of successful repetitions. Successful repetitionsmeans maintaining the activation score by 10s hold.

Procedure

Subjects were selected from the one nursing and two physiotherapy colleges of Surat city. According to the inclusion and exclusion criteria, 20 subjects were obtained from each college by using systematic random sampling.

Figure 3: Raters scoring the Performance Index for Inter-Rater Reliability.

Figure 4: Rater scoring the Performance Index for Intra-Rater Reliability

Measurements were obtained by the two physiotherapists in order to test the inter-rater reliability of ADIT. Measurements were taken by the same physiotherapist on two different days with seven days interval for intra-rater reliability[30]. Both the raters (who were pursuing Master of Physiotherapy program) had practiced sufficiently beforedoing it on the subjects. All the subjects were examined through the screening test and signed the written informed consent form. The study was approved by the Institutional Ethical Committee of the SarvajanikCollege of Physiotherapy.

The process was as follows:

- All the participants received basic information about the function of TrA, as well as about the procedure of testing and training the TrA muscle contraction and both raters were present during the actual test was conducted.
- All subjects were instructed to fast for 2 hours prior to testing (including water), empty the bladder immediately before the tests and not perform abdominal exercises prior to the tests[21] well in advance.
- For both inter- and intra- rater scoring, participants and raters had adopted the same clinical, temporal and environmental conditions to avoid external influences or internal errors during the period of data collection.
- First of all, the subjects were taught in four point kneeling position, standing and sitting position. Four-point kneeling, sitting and standing positions were used in order both to identify substitution strategies and to start the learning of correct TrA contraction from positions easier than the prone position to be used in the test as in relaxed abdomen the TrA is more in its lengthened position during the contraction. After the patient has learned enough and is able to do than the test was conducted in prone.
- The patient lies prone with the arms by the side, head fully relaxed in the designated mould so that the neck was straight and relaxed with the head in the midline with the lower limbs positioned with the feet off the plinth and the PBU is placed under the abdomen with the navel in the centre and the distal edge of the pad in line with the right and left anterior superior iliac spines.
- The pressure pad is inflated to 70 mmHg and allowed to stabilize. This pressure has been identified to be that which inflates the pad sufficiently to detect changes in position of the abdominal wall but is comfortable and does not press into the abdominal contents. According to Richardson and Jull [9], this tool was designed to monitor movement of the

abdominal wall by measuring a change in pressure during abdominal hollowing. At rest, small deviations of the indicator on the pressure dial will be evident with abdominal movement during normal respiration, and thus it is essential to identify the point about which the level fluctuates.

- Before the actual data collection was commenced in this study, pilot study with ten subjects was conducted.
- Participants were instructed to breathe mainly using the abdominal wall and then inflatable bag was adjusted to 70 mmHg again.
- The patients were instructed to breathe in and out and then without breathing in, to perform the test with the verbal instruction given by the rater as follows 'Draw in your abdominal wall without moving your spine or pelvis and hold for 10 s while breathing normally'.
- Deep inspiration was avoided. And after the contraction was achieved, patient had breath normally between the contractions. The ability to contract the muscle results in pressure reduction from 4 to 10 mmHg which was recorded by the pressure gauge of PBU. [17]
- After one successful completion of one episode of contraction, participants were instructed to relax their whole body fully, especially the abdomen, before each contraction and sufficient brief period of rest was given to the subject before the procedure is repeated up to 10 times to test the endurance of TrA. The amount of holding time was measured using the stop watch.
- Possible compensation to be avoided were identified as: (a) Contraction with visible co-contraction of other muscles for example: gluteus, quadriceps, back muscles; (b) Tilting of pelvis or flexing of spine; (c) Pressure reduction of 0 mmHg; & (d) Increase in pressure from baseline.

The above procedure was developed by Richardson *et al.* [9] The test was terminated when the subjects were not able to contract further and if the subjects experienced fatigue during the succeeding contractions and that score was recorded by the raters. The data was calculated using the performance index [29] (activation score*number of successful repetition). Activation score is the amount of pressure level the subject is able to achieve. For each of the pressure level the subjects achieved, 10 s hold was to be maintained for the successful repetition. Both the raters recorded the score on the scoring sheet. Both raters were prohibited from exchanging information to remain blinded to the score taken by each other. This procedure was followed for inter-rater reliability.

The subjects were not told the scores that they achieved during first test so as avoid bias on the results of performance level of the subjects, and the procedure was repeated after seven days and data thus obtained was used to calculate for intra-rater reliability. The same testing procedure and equipment was used for all the subjects.

Results

In this study, total 60 asymptomatic subjects (females=57 and males=3) were studied. Table 1 shows the demographic data of all the subjects.

Table 1: Demographic data of subjects

Subjects	Mean	SD
Age (years)	20.40	1.669
Height (meters)	1.53	0.068
Weight (kg)	49.77	5.973

Table 2 shows the descriptive statistics as mean and standard deviation with minimum and maximum values for performance index of ADIT.

Table 2: Descriptive statistics

Rater	N	Minimum	Maximum	Mean	SD
Rater 1	60	4	46	13.58	10.887
Rater 2	60	4	76	14.08	12.473
Retest (Rater 1)	60	4	60	13.63	12.094

Table 3 shows the intra-class correlation coefficient (ICC) for the inter-rater reliability taken by the rater 1 and rater 2 along with confidence interval (CI) with a p value < 0.05. The ICC value shows very high reliability.

Table 3: ICC (Inter-rater reliability) with CI

ICC (inter-rater)	CI(lower)	CI(upper)
0.944	0.906	0.966

Table 4 shows the intra-class correlation coefficient (ICC) for the intra-rater reliability taken by the rater 1 twice along with confidence interval (CI) with a p value < 0.05. The ICC value showsvery high reliability.

Table 4: ICC (intra-rater reliability) with CI

ICC (intra-rater)	CI (lower)	CI(upper)
0.910	0.850	0.946

Figure 9 shows the Bland Altman limits of agreement between the two raters (raters).

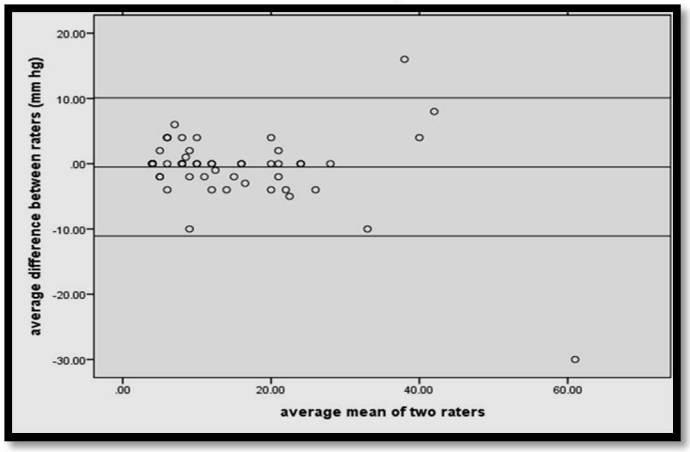

Figure 9: Bland-Altman limits of agreement analysis between two raters

The Bland-Altman chart is a scatter plot with the difference of the two measurements for each sample on the vertical axis and the average of the two measurements on the horizontal axis.

Three horizontal reference lines are superimposed on the scatter plot - one line at the average difference between the measurements, along with lines to mark the upper and lower control limits of plus and minus 1.96*sigma, respectively, where sigma is the standard deviation of the measurement differences. If the two methods are comparable, then differences should be small, with the mean of the differences close to 0 (31).It shows reasonable agreement between the raters as most of the values fall in M ± 2SD($p<0.05$). It indicates high reliability.

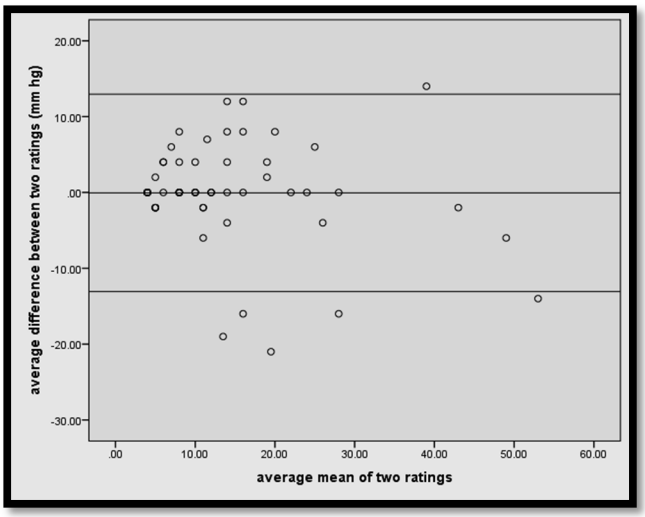

Figure 10: Bland-Altman limits of agreement analysis between scores taken by the same rater twice

It shows reasonable agreement as most of the values fall in M ± 2SD ($p < 0.05$).

The standard error of measurement(SEM) is a measure of absolute reliability; the smaller the SEM the more reliable the measurements(32, 33).The SEM value calculated for variability in measurements between the two raters is 0.69725 which is very small; whereas the variability in measurements of same raters is 0.85814 which is very small. Thus these measurements are reliable.

Table 5: Standard error of measurement (SEM) values

	Variability in measurements between two raters	**Variability in measurements of same raters**
Standard error of measurement	0.69725	0.85814

The true SEM value for variability in measurements between two raters (0.69725*1.96=1.36661) suggests that any individual value lies within the range of ±1.36661 PI from their measured value.The true SEM value for variability in measurements of the same raters (0.85814*1.96=1.6819544) suggests that any individual value lies within the range of ±1.6819544 PI from their measured value.

Table 6: TrueStandard error of measurement (SEM) values

	Measurements between two raters	**Measurements of same raters**
True Standard error of measurement	1.36661	1.6819544

The smallest real difference (SRD) value for variability of measurements between the two raters ($1.96*\sqrt{2}*SEM = 1.932$) and between the measurements taken by the same rater ($1.96*\sqrt{2}*SEM = 2.378$) is claimed to be capable of representing the “real” change but these values cannot simply be generalised to symptomatic populations.

Table 7: Smallest real difference (SRD) values

	Measurements between two raters	**Measurements taken by same raters**
Smallest real difference	1.932	2.378

Discussion

In this cross-sectional study, which aimed at measuring the intra and inter-rater reliability of ADIT in asymptomatic individuals by use of pressure biofeedback, the reliability estimates ranged from satisfactory to excellent for both intra-rater and inter-rater conditions.The use of pressure biofeedback for the evaluation of subjects with and without low back pain or for providing the feedback for the rehabilitation of patients with low back pain has been increased. In clinical practice, it is common for patients to be evaluated several times by the same or by different examiners. Therefore, it is important to know the reproducibility of measures and instruments used by the same examiner on different occasions as well as by different examiners. [34]

Is found in this study, the intra class correlation coefficient for the inter-rater reliability between the two raters is 0.944 and for intra-rater reliability is 0.910. The study by Lima *et al* [23] found the intra class correlation coefficient of 0.76 for inter-examiner reliability and 0.74 for intra-rater reliability. The study done by Von Garnier*et al* [24] reported low inter-observer reliability of 0.47 and ICC of 0.81 for intra-rater reliability. Costa *et al* [21] in their study reported moderate reliability with ICC of 0.58. The study done by Storheim*et al* [22] reported low intra-observer reliability. The ICC of present study cannot be compared to any other studies as the outcome measure used was different than the other studies.

The discrepancy of values existing among the studies may be due to methodological differences between studies, such as sample sizes, study participants, different criteria for the test, and standardization of breathing during the tests and different methods of statistical analysis. The difference between the values can be due to the different population taken i.e. study by Lima *et al*[23] recruited chronic non specific low back pain patients and study by Von Garnier*et al*[24] recruited the subjects who had with and without low back pain. The present study targeted asymptomatic individuals, and even the studies by Storheim*et al*[22] and Costa *et al*[21] recruited asymptomatic individuals.

Storhiem *et al.* [22] used coefficient of variation for reliability analysis whereas in the present study and the studies by Lima *et al.* [23], Von Garnier *et al.* [24], Costa *et al.* [21] used ICC for reliability analysis. This might be the one possible reason for low intra-observer reliability.

In this study, both the examiners had practiced sufficiently before application of test; and equipment and testing conditions had been used were same throughout for all the subjects.

Similarly the studies conducted by Lima *et al.* [23], Storheim *et al.* [22], and Costa *et al.* [21], also maintained uniform testing conditions but then too conflicting results were found between the previous studies. This was likely due to different criteria was adopted for the each study as to how the pressure data was collected. While Richardson *et al.* [9] collected a pressure reduction of 4-10 mmHg for 10 s, Costa *et al.* [21] and Storheim *et al.* [22] recorded the maximum pressure reduction of at least 2 s within a period of 8-10 s. In contrast, Von Garnier *et al.* [24] performed their data collection using a set of four criteria that participants would have to fulfill for the correct TrA muscle contraction: continuous breathing, absence of muscle substitution maneuvers, appropriate muscle contraction checked by palpation test and a pressure reduction of at least 1 mmHg for 4 s within a period of 10 s.

This study is in accordance with the criteria by Richardson *et al.* [9] who collected a pressure reduction of 4-10 mmHg for 10 s and repeating the procedure for 10 times. This was the main target of the study which focused on testing the reliability of the endurance of TrA while the studies mentioned above targeted the reliability of activation of TrA.

There was also conflicting results in all the studies because some studies evaluated only peak of contraction in certain period of time while some studies targeted specific pressure reduction within stable period of time. The outcome measure used in this study was **performance index** which is not being used in any other studies.

Though the reliability of abdominal drawing-in test reported high reliability for intra and inter rater, there was poor contractile capacity of TrA for some individuals as performance index

for some subjects were as low as 4 mmHg while the contractile capacity for some subjects were good but only few of the subjects were able to complete the test for 10 times which indicated good endurance capacity of that individuals. The possible reason could be as this study targeted the normal healthy populations and populations which consist of allied health professionals who may have higher degree of body awareness and coordination skills than sedentary populations.

It can also be said that pressure reduction is different in subjects with low back pain and in asymptomatic individuals because individuals with back pain have difficulty in performing correct recruitment of TrA so therefore this emphasizes that the study to be done in homogenous group of population.

Standardization of breathing was utmost important for the proper recruitment of TrA because this muscle is most active towards end of expiration and due to its anatomical location.

Lafound *et al.* [35] found that there are significant differences between pressure measurements collected during breathing and apnea, with higher values observed during normal breathing. Participant without guidance with regards to normal breathing have a tendency to contracts TrA with apnea. [18] Thus to minimize the error, pressure measurements should be collected at the end of expiration which was maintained in the present study. While Storhiem *et al.* [22] did not standardized breathing during the study.

Standardization of protocol is also very much necessary for the proper result of inter- and intra-rater reliability. In the present study the subjects were positioned in same way for all subjects, on hard surface. The studies by Costa *et al*[21], Storheim *et al.*[22] had small sample size and only one study by Von Garnier *et al.* [24] conducted a pilot study. To analyze the clinometric properties of assessment tools, it is recommended that samples should include at least 50 individuals, or a pilot study should be performed prior to the sample size calculation which was done in the present study [36].

In all test situations there is a learning effect that may improve test results of the re-test. [37] The choice of seven days between tests was made to limit the learning effect. A time interval between tests of 7 days was mentioned in studies of Lima *et al.* [23], Storheim *et al.* [22], Costa *et al.* [21]. The time period between repetitions of the measures should be long enough to avoid memorization of data by examiners, but short enough to ensure that there were no clinical changes in the participants. It is recommended that 1 or 2 weeks would be ideal, but there may be reasons for the choice of another interval. [36] Subjects were told not to exercise the TrA muscle during the seven-day period between tests.

In this study, the findings of Bland-Altman limits of agreement showed excellent inter-rater agreement between the raters (limits of agreement (LOA) = 10.08 to -11.08 mmHg) indicating that measures related to the rater 1 were in agreement with the rater 2 in 95% of occasions. Similarly, we found excellent intra-rater agreement (LOA) = 12.9 to -13.06 mmHg), which means that measures relating to first test were in agreement with the second test in 95% of occasions. Similar results were found by Lima *et al.*[23] who also reported excellent agreement between the raters; and same rate on two separate occasions.

As this study targeted the reliability of endurance of TrA in asymptomatic individual this result cannot be generalized to back pain patients. Moreover Rothstein (38) claimed that measurement errors may be higher in patient groups than in healthy people owing to pain and dysfunction. Richardson et al[9] claim that many patients need a long period of practice to learn an effective contraction of the TrA muscle, and the studies of Hodges *et al*[14] and Cairns *et al*[18]conclude that subjects with low back pain have severe problems with conducting the abdominal drawing-in action and reduce the pressure measured by the PBU.

This might indicate even lower reproducibility in patients than in normal subjects. As the study by Lima *et al.* [23] have established a successful result of pressure reduction of 4 mmHg in chronic non specific low back pain, taking this into account the reliability of endurance of TrA can be studied before using it in intervention strategies.

This study also found SEM of 0.69725 mmHg for inter-rater and 0.85814 mmHg for intra-rater reliability. The true SEM for inter rate is 1.36 mm of Hg and intra rate is 1.68 mm of

Hg which suggest the absolute measurement error of PBU. The SRD for inter-rater and intra-rate is 1.932 and 2.378 which suggest that there should be a small difference of these values so as to say that "real" change has occurred.

The study by Lima *et al.* [23] has found the SEM and SRD value but that values are for the activation of TrA. There is no normative data in literature available for the endurance of TrA so the result of the present study cannot be compared.

The scoring of inter-rater reliability was taken by both the raters together so that duration of contraction or fatigue has homogenous effect on all subjects and moreover to avoid the effect of fatigue on the performance level of the subjects. If the scores were taken at different times, than it would have been difficult to decide that scores were result of true performance of the subject; or, had fatigue affected the level of performance of subject.

The accuracy of PBU device is ± 3 mmHg[17] which can cause random error in subjects and to avoid that same; contact of abdomen and inflatable bags should be maintained identical during both test and retest. There were much compensation that could have occurred but this was minimized during the practice sessions of all subjects. This study supports the use of ADIT as an objective measure to assess the TrA endurance.

Conclusion

The inter-rater and intra-rater reliability of ADIT is 'very high'[39] in asymptomatic individuals. Thus it can be used as an objective measure to assess the endurance of TrA.

However the studies should be conducted on patient populations to generalize the results.

If the results show low endurance capacity of TrA, than appropriate rehabilitative measures can be implemented.

Limitations

It is possible to monitor the activity of global muscles by observation, but it is less accurate. So EMG analysis would have been more appropriate. Absolute blinding of the raters was not possible. To minimize the error on performance results of subjects, both the raters didn't discuss anything during the recording of the scores.

Conflict of Interest: None declared.

Authors' Contribution:

JKD performed review of literature and collection of data; BD drafted the manuscript, designed and coordinated the study; TRA performed the statistical analysis, interpretation of data and review of manuscript.

Acknowledgements

We deeply thank Dr. M. T. Rangwala (Director, Sarvajanik College of Physiotherapy, Surat), Ms.Vijaya Ghaskadhvi, Director of T&TV Nursing College, Surat and Mr. Anjan Desai, Principal I/c of SPB College of Physiotherapy, Surat for their support and all the participants of the study for their cooperation. Also we acknowledge the help extended by Ms. Nidhi Chhabra [PT].

References

[1] Andersson GB. Epidemiology of low back pain. Acta orthopaedica Scandinavica Supplementum. 1998; 281:28-31.

[2] Dionne CE, Dunn KM, Croft PR. Does back pain prevalence really decrease with increasing age? A systematic review. Age and ageing. 2006; 35(3):229-34.

[3] Deyo RA, Cherkin D, Conrad D, Volinn E. Cost, controversy, crisis: low back pain and the health of the public. Annual review of public health. 1991;12:141-56.

[4] Manek NJ, MacGregor AJ. Epidemiology of back disorders: prevalence, risk factors, and prognosis. Current opinion in rheumatology. 2005;17(2):134-40.

[5] Winkel D. Diagnosis and Treatment of spine: Nonoperative Orthopaedic Medicine and Manual therapy. Gaithersburg: Aspen; 1996.
[6] Panjabi M, Abwni K, Duranceau J, Oxland T. Spinal stability and intersegmental muscle forces. A biomechanical model. Spine1989.
[7] Bogduk N. Clinical anatomy of lumbar spine and sacrum. Edinburgh: Churchill Livingstone; 1997.
[8] Crisco JJ, Panjabi MM. The intersegmental and multisegmental muscles of the spine: a biomechanical model comparing lateral stabilising potential. Spine. 1991;7:793-9.
[9] Richardson C, Jull G, Hodges P, Hides J, Panjabi M. Therapeutic Exercise for spinal segmental stabilization in Low back pain. Scientific Basis and Clinical Approach. 1 ed: Churchill Livingstone; 1999.
[10] Cholewicki J, McGill SM. Mechanical stability of the in vivo lumbar spine: implications for injury and chronic low back pain. Clinical biomechanics. 1996;11(1):1-15.
[11] Cresswell AG, Grundstrom A, A T. Observations on intra-abd ominal pressure and patterns of abdominal intra-muscular activity in man. Acta Physiologica Scandinavica. 1992;144:409-18.
[12] Tesh KM, ShawDunn L, Evans JH. The abdominal muscles and vertebral stability. Spine. 1987;12:501-8.
[13] Hodges PW, Richardson CA. Altered trunk muscle recruitment in people with low back pain with upper limb movement at different speeds. Archives of physical medicine and rehabilitation. 1999;80(9):1005-12.
[14] Hodges P, Richardson C, Jull G. Evaluation of the relationship between laboratory and clinical tests of transversus abdominis function. Physiotherapy research international : the journal for researchers and clinicians in physical therapy. 1996;1(1):30-40.
[15] Ferreira ML, Ferreira PH, Latimer J, Herbert RD, Hodges PW, Jennings MD, et al. Comparison of general exercise, motor control exercise and spinal manipulative therapy for chronic low back pain: A randomized trial. Pain. 2007;131(1-2):31-7.
[16] Costa LO, Maher CG, Latimer J, Hodges PW, Shirley D. An investigation of the reproducibility of ultrasound measures of abdominal muscle activation in patients with chronic non-specific low back pain. European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society. 2009;18(7):1059-65.
[17] Chattanooga-Group-Inc. Stabilizer Pressure Biofeedback. Operating instructions. Hixson2005.
[18] Cairns MC, Harrison K, Wright C. Pressure Biofeedback: A useful tool in the quantification of abdominal muscular dysfunction? Physiotherapy. 2000;86(3):127-38.
[19] Richardson CA, Jull GA. Muscle control-pain control. What exercises would you prescribe? Manual therapy. 1995;1(1):2-10.
[20] Richardson C, Jull G, Hodges P, Hides J, Panjabi MM. Therapeutic Exercise for spinal segmental stabilization in Low back pain.Scientific Basis and Clinical Approach: Churchill Livingstone; 1999.
[21] Costa LO, Costa Lda C, Cancado RL, Oliveira Wde M, Ferreira PH. Short report: intra-tester reliability of two clinical tests of transversus abdominis muscle recruitment. Physiotherapy research international : the journal for researchers and clinicians in physical therapy. 2006;11(1):48-50.
[22] Storheim K, Bo K, Pederstad O, Jahnsen R. Intra-tester reproducibility of pressure biofeedback in measurement of transversus abdominis function. Physiotherapy research international : the journal for researchers and clinicians in physical therapy. 2002;7(4):239-49.
[23] Lima PO, de Oliveira RR, de Moura Filho AG, Raposo MC, Costa LO, Laurentino GE. Reproducibility of the pressure biofeedback unit in measuring transversus abdominis muscle activity in patients with chronic nonspecific low back pain. Journal of bodywork and movement therapies. 2012;16(2):251-7.
[24] Von Garnier K, Koveker K, Rackwitz B, Kober U, Wilke S, Ewert T, et al. Reliability of a test measuring transversus abdominis muscle recruitment with a pressure biofeedback unit. Physiotherapy. 2009;95(1):8-14.
[25] De Paula Lima PO, de Oliveira RR, Costa LO, Laurentino GE. Measurement properties of the pressure biofeedback unit in the evaluation of transversus abdominis muscle activity: a systematic review. Physiotherapy. 2011;97(2):100-6.
[26] Zou GY. Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. Statistics in medicine. 2012;31(29):3972-81.

[27] Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. Statistics in medicine. 1998;17(1):101-10.
[28] Chattanooga G. Stabilizer pressure bio-feedback. Operating instructions. Hixson: Chattanooga Group Inc. 2005.
[29] Jull G, Barrett C, Magee R, Ho P. Further clinical clarification of the muscle dysfunction in cervical headache. Cephalalgia : an international journal of headache. 1999;19(3):179-85.
[30] Currier DP. Elements of Research in Physical Therapy. 3 ed: Lippincott Williams and Wilkins; 1990.
[31] Martin Bland J, Altman D. Statistical methods for assessing agreement Between two methods of clinical measurement. The Lancet.327(8476):307-10.
[32] Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports medicine (Auckland, NZ). 1998;26(4):217-38.
[33] De Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. Journal of clinical epidemiology. 2006;59(10):1033-9.
[34] Portney LG, Watkins MP. Foundations of Clinical Research: Applications to Practice: Prentice Hall; 1999.
[35] Lafond D, Dimmock M, Champagne A, Descarreaux M. Intrasession reliability and influence of breathing during clinical assessment of lumbar spine postural control. Physiotherapy theory and practice. 2009;25(3):218-27.
[36] Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. Journal of clinical epidemiology. 2007;60(1):34-42.
[37] Thomas J, Nelson J. Research Methods in Physical Activity. IL: Human Kinetics Books; 1990.
[38] Rothstein J. Measurement in Physical Therapy. New York: Churchill Livingstone; 1985.
[39] Venturni C, Andre A, prates A, Giacomelli B. Reliability of two evaluation methods of active range of motion in the ankle of healthy individuals. Acta Fisiatr. 2006; 13(1):39-43.