Reliability study

Intra-rater and inter-rater reliability of total faulty breathing scale using visual observation and videogrammetry methods

Vikram Mohan [a, b], Maria Perri [c], Aatit Paungmali [a, *], Patraporn Sitilertpisan [a], Leonard Henry Joseph [a], Romizan Jathin [b], Munirah Bt Mustafa [b], Siti Hawa Bt Mohd Nasir [b]

[a] Neuro-Musculoskeletal and Pain Research Unit, Department of Physical Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai, Thailand
[b] Department of Physiotherapy, Faculty of Health Sciences, Universiti Teknologi MARA, Puncak Alam, Malaysia
[c] Wellness Springs, 489 Route 32, Highland Mills, NY, 10930, USA

A R T I C L E   I N F O

A B S T R A C T

Faulty breathing is an aspect of alteration in the normal fundamental pattern of breathing. The available existence of scales in assessing faulty breathing has not frequently been used. Measurement errors in assessing and quantifying breathing patterns may originate from unclear directions and variation between observers. This study determined the measure reliability of the Total Faulty Breathing Scale (TFBS) for quantifying breathing patterns. Twenty seven participants were recruited comprising healthy and unhealthy subjects. Two examiners assessed their breathing patterns using the TFBS on two different occasions with visual observation and a videogrammetry method. Evaluation of the observational breathing pattern method for intra-rater and inter-rater showed agreement of 96.30% and a kappa score of greater than 0.78, which indicated substantial agreements. The videogrammetry method showed a percent agreement of (100%) with a kappa score of (1.00). This study indicates that the TFBS is a considerably reliable tool for evaluating breathing patterns with both visual observation and a videogrammetry method.

© 2016 Published by Elsevier Ltd.

1. Introduction

Breathing is an essential fundamental mechanical, physiological and psychological process which is required throughout our lifespan (CliftonSmith and Rowley, 2011). An alteration in these fundamental processes could be the first sign of faulty breathing (FB) or breathing pattern disorders (BPD) or dysfunctional breathing (DB) (Barker and Everard, 2015; CliftonSmith and Rowley, 2011; Perri and Halford, 2004). According to earlier evidence, expiration is faulty if rib motion is reduced, the breath is held and not fully exhaled or paradoxical breathing takes place (Perri and Halford, 2004). This signifies that rib motion cannot be normal if there is no abdominal initiation during inhalation, which is the key criteria for normal respiration (Lewit, 1999). Other than that, an altered breathing pattern occurs when upper costal expansion is greater than abdominal and lateral costal expansion during inhalation (Ha et al., 2014). As a result, upper costal breathing requires accessory muscles to work more in respiration and causes muscle strain which could lead to respiratory dysfunction and musculoskeletal disorder. Abnormal movement during respiration can also be aggravated by diseases and injuries to lung tissue, rib cage, respiratory muscles and nerves (Gunnesson and Olsén, 2011). In addition, the clinical picture of faulty breathing patterns may be present among unhealthy individuals who are ailing with pain-related musculoskeletal problems and cardiorespiratory illness. It must be noted that faulty breathing patterns could also occur among healthy individuals as a result of underlying abnormal patterns that have become habitual in the motor program. Hence, in clinical practice, quantification of breathing patterns is crucial in rehabilitation of patients with respiratory, neurological and musculoskeletal disorders. However, it is claimed that the prevailing effect of faulty breathing is not always documented in clinical settings (Chaitow, 2014).

In routine clinical settings, the techniques used to assess breathing patterns are usually through visual inspection and manual palpation methods (Hammer and Newth, 2009; Pryor and Prasad, 2002). However, these two methods of assessment have not been standardized, as there are differences in hand placement on the chest wall for assessing breathing patterns. In addition, there is no gold standard for assessing normal and faulty breathing patterns. Hence, the assessment of normal and faulty breathing patterns has never become a routine part of standard clinical examination protocols (Hammer and Newth, 2009).

In general, the objective method of assessing breathing patterns is measured through magnetometer, Respiratory Inductive Plethysmography (RIP), Respiratory Movement Measuring Instrument (RMMI) and the ELITE system. This equipment is costly and rarely available in routine clinical set up and practice (Gunnesson and Olsén, 2011; Hammer and Newth, 2009; Kaneko and Horie, 2012). Additional tools which can be used to assess breathing patterns are the Manual Assessment Respiratory Motion (MARM), Hi Lo breathing assess-

ment and The Nijmegen questionnaire (Courtney et al., 2009; Dixhoorn and Folgering, 2015). Although several methods of assessment exist, these tools are not widely used due to expense, complexity of use, and lack of awareness and training. It is evident that a simple, reliable and easily available method for evaluating breathing patterns is necessary.

An earlier study, conducted by Perri and Halford, investigated both relaxed and deep breathing to measure the occurrence of normal and faulty breathing (Perri and Halford, 2004). In order to interpret the data, the authors created a simple rating scale which is easily understandable and requires minimal training to score. They named it the Total Faulty Breathing Scale (TFBS).

Despite the usefulness of the TFBS, however, its reliability has not yet been tested. It is claimed that this score can be derived through a visual observation method in clinical practice (Perri and Halford, 2004). In addition, earlier studies showed a technique of measuring breathing patterns and the orientation of the thoracic wall with photogrammetric and videogrammetry methods (Cihak et al., 2006; Herráez et al., 2013). Photogrammetry is the science of making analysis of two or more photographs, whereas videogrammetry is the science of measuring two or more videos with single or multiple cameras. However, these methods of measurement used surface markers which require specialized software such as Corel R.A.V.E and a video kinematic analysis system for image processing that also requires a trained person in that field to interpret the data. (Cihak et al., 2006; Herráez et al., 2013). It is evident that a simpler and more cost effective assessment method is necessary for grading breathing patterns. Hence, in this study, the concept of the videogrammetry method was utilized in three different directions to estimate breathing patterns using the TFBS and the reliability measures were tested. The aim of the study was to investigate the reliability of assessing breathing patterns using the TFBS for visual observation and videogrammetry methods.

2. Materials and methods

2.1. Study design and subjects

This study was a test-retest reliability design in determining the breathing score for visual observation and video recording methods. The selection criteria for the study were as follows: being a volunteer, male gender, and between 18 and 24 years of age. A convenience sampling from the physiotherapy department at a public university was used. A total of 27 male healthy and unhealthy adults were screened and participated in the study. The total of twenty seven subjects was required to establish the significant $\alpha = 0.05$ and $\beta = 0.20$, when the one-way random effects model is used for estimating reliability as described by earlier statistical guidelines (Shoukri et al., 2004). Measurements of breathing score were obtained through a visual observation method in the physiotherapy clinic, followed by a video recording method. The study protocol was approved by the university research ethics committee. Prior to data collection, informed consent and health evaluation forms were obtained from each individual participant.

2.2. Experimental procedures

Two physiotherapists with 3 years of clinical experience were instructed in assessing the breathing patterns and scoring methods of the TFBS. The two therapists scored each participant simultaneously using the TFBS. Both visual observation and video recordings were carried out simultaneously during the same sessions. To improve reli-

ability, a 24 h period passed between assessments, and the participants were observed in random order. To insure the randomness of the observations, video recordings of the subjects were taken from the neck to the anterior superior iliac spine (ASIS) and each was given the same colored trousers.

Initially, the subjects rested for a period of 5 min in a seated position to restore optimal vital parameters. Then, they were given an appropriate explanation regarding the study procedures. The subjects were requested to stand with their shirts off during the procedures and were not aware that they were being assessed for breathing. The breathing assessment was carried out in an upright standing position against a white background without any additional support (Cihak et al., 2006).

2.3. Total faulty breathing scale

The scale was created based on accepted research for both quiet and deep breathing as suggested by earlier literature (Perri and Halford, 2004). The criteria employed for a normal breathing pattern is that it initiates in the abdomen, which expands outward during inhalation and inward during exhalation, has some degree of horizontal lower rib motion, and presents no lifting motion in the upper ribs and no presence of clavicular grooves. (Perri and Halford, 2004).

The scoring was made based on three main criteria during inhalation; absence of outward lateral rib motion, lifting of the clavicle and paradoxical breathing in both relaxed and deep breathing. The present study adapted a scoring system for grading normal and abnormal breathing patterns. In this scoring system, a range of values is given to differentiate between normal, mild, moderate and severe breathing patterns.

The scoring system as presented in Table 1, is as follows: For example, if a participant presented with no outward lateral rib motion (√), he would score a 1. If he lifted his clavicle (√), he would score a 2 on this scale which is different from earlier methods (Perri and Halford, 2004). If the above findings presented during quiet breathing assessment, the TFBS score for quiet breathing would be a 3. The same criteria would be applied during deep breathing assessment and a separate TFBS score would be tallied.

To test for intra-examiner reliability, the participants were evaluated twice by each examiner for both quiet and deep breathing during the visual observation and the examiners reviewed each participant's

Table 1
Total faulty breathing scale form.

| Date of assessment: | | |
| ID: | | |

Score	Criteria	Observation (√/X)
Normal/relaxed breathing		
1	Absence of outward lateral rib motion	
2	Lifting of the clavicle	
3	Paradoxical breathing	
Deep breathing		
1	Absence of outward lateral rib motion	
2	Lifting of the clavicle	
3	Paradoxical breathing	

Symbols: √ = 1, 2 or 3 depending on the criteria, X = 0.
Total score:
Grading of Breathing Pattern Dysfunction:
Normal: 0.
Mild: 1–4.
Moderate: 5–8.
Severe: 9–12.

video of both quiet and deep breathing twice on two different days during the video method.

If a participant had none of the 3 faulty findings (X), he would score a zero. The higher the score, the greater the severity of faulty breathing present. The grading of dysfunction in the scoring system is as follows: Normal: 0, Mild: 1–4, Moderate: 5–8 and Severe: 9–12.

The scoring sheet Table 1 shows each observation for both assessment methods - visual observation and video recording. The readings were then pooled to a single spreadsheet by the principal investigator.

2.4. Procedure for visual observation and video recording methods

During the visual observations, the examiners evaluated quiet breathing without giving instruction to the participants. In fact, the participants were unaware that their breathing was being assessed. In contrast, during evaluation of deep breathing, the examiners asked the participants to take deep and slow breaths.

Simultaneously, video recordings were made using a digital video camera model (Canon EOS 600D, lens EFS 18-15 mm macro 0.25 m/0.8 ft) at the exact time of the visual observation evaluations. A tripod stand was used to hold the camera in an upright position so that the height of the camera could be adjusted to the height of each subject. In order to adjust the distance between the camera and the subject, a reference point was kept between the ASIS and clavicle. This enabled the same area to be examined from clavicle to ASIS in evaluating breathing in each participant. The videos were recorded from 3 directions: anterior, lateral and posterior views. Each examiner assessed and graded both quiet and deep breathing from all three views; anterior, lateral and posterior.

2.5. Image processing

The video selected was exported to a Lenovo computer for analysis (Guerra et al., 2011). Both examiners analyzed and scored the same videos for a second time with a 24 h time lapse between evaluations of both relaxed and deep breathing. The videos were shuffled to keep the study random.

2.6. Statistical analysis

The data was analyzed using SPSS version 21 (IBM Corporation; Armonk, New York). The SPSS data sheet imported all averaged data from a Microsoft excel spreadsheet. Distribution of variables such as age, height, and weight and body mass index is presented as mean and standard deviations (SD). The health status of the participants is presented as frequency and percentages. The focusing outcomes for the present study are breathing patterns, with the dependent variable being normal versus faulty breathing. Evaluation of intra-rater reliability and inter-rater reliability of the assessment of normal and faulty breathing patterns were determined using percent agreement statistics. We also considered kappa statistics together, which is appropriate when reporting percent agreement. The interpretation of kappa value was made based on earlier guidelines as "< 0" indicated less than chance agreement, "0.01–0.20" indicated slight agreement, "0.21–0.40" indicated fair agreement, "0.41–0.60" indicated moderate agreement, "0.61–0.80" indicated substantial agreement, "0.81–0.99" indicated almost perfect agreement and "1.00" indicated perfect agreement (Viera and Garrett, 2005). In addition, comparison of reliability scores between visual observation and video recording techniques was performed using Wilcoxon-signed rank test.

3. Results

A total of 27 male subjects participated in this study. Nineteen healthy individuals (70.4%), 4 with low back pain (14.8%), 2 with neck pain (7.4%), 1 with a chest deformity (3.7%) and 1 with respiratory disease – Bronchial asthma (3.7%) – comprised the studied population, with a mean age (21.15 ± 1.76) years, height (1.67 ± 0.66) meters, weight (65.04 ± 12.12) kilograms and body mass index (65.04 ± 12.12) kilograms/meter^2.

Out of 27 participants, 25 showed faulty breathing scores, which were dispersed between healthy and unhealthy subjects. Only 2 of the healthy subjects exhibited a normal breathing pattern in both the visual observation and the videogrammetry technique.

3.1. Intra-rater and inter-rater reliability of TFBS

The following results were obtained using data from all 27 participants. It must be noted that all 27 had the TFBS scores that ranged from normal to mild abnormality when assessed by a single investigator or two different investigators.

Table 2 and Table 3 show the intra-rater reliability of the independent investigators (examiner I & II) when assessing breathing patterns by different methods (visual observation and videogrammetry). The results of the individual percent agreement performed by both examiners for the TFBS are more than 96.30%. The kappa score for the visual observation method when reported at two different times by the same investigator was greater than 0.78, which indicates acceptable agreement. The videogrammetry method of assessing breathing using the TFBS reported 100 percent agreement and a kappa score of 1.00, which indicates that the individual raters had perfect agreement.

Table 4 shows the inter-rater reliability analysis for both visual observation and the videogrammetry method, which resulted in 100 percent agreement with a kappa score of 1.00. This indicates that the scoring system exhibits perfect agreement.

Table 2
Intra-rater reliability of visual observation and videogrammetry technique using TFBS (1st Examiner).

Techniques	Percent agreement	Kappa score	Standard error	p-value
Visual observation	100	0.839	0.154	0.000
Videogrammetry	100	1	0.000	0.000

Table 3
Intra-rater reliability of visual observation and videogrammetry technique using TFBS (2nd Examiner).

Techniques	Percent agreement	Kappa score	Standard error	p-value
Visual observation	96.3	0.780	0.210	0.000
Videogrammetry	100	1	0.000	0.000

Table 4
Inter-tester reliability of visual observation and videogrammetry technique using TFBS.

Techniques	Percent agreement	Kappa score	Standard error	p-value
Visual observation	100	1	0.000	0.000
Videogrammetry	100	1	0.000	0.000

Comparison of reliability scores of the visual observation and video recording methods revealed that there were no statistically significant differences for percent agreement and Kappa score (p > 0.05) between the two methods. This indicates that either method of assessment can be utilized interchangeably when using the TFBS.

4. Discussion

This goal of this present study was to assess the intra-rater and inter-rater reliability of assessing breathing patterns using the TFBS for two different methods of evaluation: visual observation and videogrammetry. The measurement of breathing patterns using the TFBS was evaluated in an attempt to begin to define an easy way to assess normal and faulty breathing patterns. The findings of this study show the need to explore the impact of faulty breathing patterns in a wide range of populations as well as to know the effectiveness of treatment of faulty breathing in clinical studies. The findings of the current study showed that the TFBS is a reliable tool for assessing normal and faulty breathing patterns in both healthy subjects and subjects with pathologies. The study was successful as it was accurately able to identify normal and faulty breathing patterns using the TFBS system.

Our results showed that two therapists were able to reach between substantial agreement and perfect agreement in identification of normal and faulty breathing pattern using the TFBS system with a visual observation method. Similarly, the examiners reached perfect agreement with a videogrammetry method. These results support using the TFBS to evaluate breathing patterns in both healthy and unhealthy subjects. The findings of this study indicated that twenty five subjects showed a faulty breathing score, which was dispersed between healthy and unhealthy subjects. However, only two of the healthy subjects exhibited a normal breathing pattern in both visual observation and videogrammetry techniques. The most significant finding that could be inferred is that the TFBS has the ability to differentiate between normal and faulty breathing patterns. This observation supports the hypothesis that even healthy subjects can have faulty breathing patterns, which may be a predisposing factor for future illness or pain syndromes.

The results of the reliability measures in this present study cannot be directly compared with other studies, as this is the first study to our knowledge using the TFBS system. Despite this, the results of the study can be compared indirectly with an earlier study which utilized other methods of evaluating breathing patterns. The study utilized MARM and RIP methods of assessing breathing patterns, and the results suggest that MARM can be a reliable clinical tool for assessing breathing patterns (Courtney et al., 2009). However, the findings of the previous studies were limited to inter-examiner reliability measures alone (Courtney et al., 2009). A significant drawback of the MARM is that it requires specialized training to use it. In contrast, the TFBS system of breathing pattern assessment is easy to use and would require minimal training for medical and health care professionals.

The videogrammetry method used in this present study differs from an earlier method that measured three-dimensional rib excursion during breathing by placing cameras in three different directions with external biomarkers on the chest (Sarro et al., 2009). This three dimensional positioning of cameras with fixation of external biomarkers in the chest wall is not cost or time effective in most clinical setups of today and was not used in the present study. In spite of methodological variation, the present study yielded perfect agreement in measuring both quiet and deep breathing patterns and determining if the patterns of breathing were normal or faulty. Visual observation

using the TFBS is a reliable and easy way of assessing and charting breathing patterns.

The decision to use Kappa statistics in this present study was justified because percent agreement is limited as a result of random judgement rather than actual agreement (Hunt, 1986; Viera and Garrett, 2005). In addition, Kappa statistics are well-known and widely accepted for evaluation of agreement between two ratings of categorical data (Hunt, 1986; Viera and Garrett, 2005). In general, the purpose of the scale is to identify normal and faulty breathing patterns. It can therefore be assumed that the TFBS can be used as a preliminary screening tool to identify normal and faulty breathing patterns. Once the faulty breathing pattern is identified, other measures of quantification such as the photogrammetric method proposed earlier can be used for identifying types of breathing (Cihak et al., 2006).

4.1. Limitations and recommendations

Further work is required to establish the reliability measures between the genders as the present findings was limited to studying only male participants. The study was conducted with a male population due to the cultural constraints of exposing the female chest in observation and videography. Hence, these constraints need to be restrained in future studies to generalize the study results. Future studies to test the reliability of this scoring system should be performed in a clinical setting with various levels of abnormal breathing patterns for individuals who are ailing from cardio-respiratory illness or chronic neuromusculoskeletal pain rather than on a mixed population. Studies are necessary to institute optimal cut-off scores by matching the outcome or scoring system to inexperienced persons with breathing or other difficulties.

Apart from that, since the observational method was carried out on two different occasions, the subjects would have been aware of what the examiners were doing. Hence, the mechanism of breathing may have been changed to control breathing and this can be dealt when the assessment is performed only one time.

In addition, the present study did not consider assessing reliability separately for those considered healthy and those considered diseased as the number of subjects recruited for this study was small. Moreover, it has been set in the study protocol to study mixed populations before specialized populations are studied. Also, the present study did not consider assessing quiet and deep breathing separately, as the total scoring range may change.

5. Conclusion

This study indicates an acceptable reliability for the TFBS to evaluate normal and faulty breathing patterns using visual observation and videogrammetry methods. Hence, it suggests that the TFBS appears to detect both normal breathing and faulty breathing patterns in various populations.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgement

References

Barker, N., Everard, M.L., 2015. Getting to grips with "dysfunctional breathing". Paediatr. Respir. Rev. 16 (1), 53–61. http://dx.doi.org/10.1016/j.prrv.2014.10.001.

Chaitow, L., 2014. Functional movement and breathing dysfunction. J. Bodyw. Mov. Ther. 1, 2006–2007. http://dx.doi.org/10.1016/j.jbmt.2014.08.008.

Cihak, J, Cumpelik, J., Krticaka, L., Smutny, V., Strnad, P., Sara, R., et al. Zyka, V., 2006. Recognition of Breathing Pattern by a Photogrammetric Method.

CliftonSmith, T., Rowley, J., 2011. Breathing pattern disorders and physiotherapy: inspiration for our profession. Phys. Ther. Rev. 16 (1), 75–86. http://dx.doi.org/10.1179/1743288X10Y.0000000025.

Courtney, R., Cohen, M., Reece, J., 2009. Comparison of the manual assessment of respiratory motion (MARM) and the hi Lo breathing assessment in determining a simulated breathing pattern. Int. J. Osteopath. Med. 12 (3), 86–91. http://dx.doi.org/10.1016/j.ijosm.2008.10.002.

Dixhoorn, J. Van, Folgering, H., 2015. The Nijmegen Questionnaire and dysfunctional breathing. Eur. Respir. J. 3–6. http://dx.doi.org/10.1183/23120541.00001-2015.

Guerra, L.V., Ochoa, M.C., Tamayo, M.Z., 2011. Effectiveness of a photogrammetric model for the analysis of thoracoabdominal respiratory mechanics in the assessment of isovolume maneuvers in children. J. Bras. Pneumol. 27 (1), 36–48. http://dx.doi.org/10.5144/0256-4947.2012.243.

Gunnesson, I.L., Olsén, M.F., 2011. Validity in measuring breathing movements with the respiratory movement measuring instrument, RMMI. Clin. Physiol. Funct. Imaging 31 (1), 1–4. http://dx.doi.org/10.1111/j.1475-097X.2010.00970.x.

Ha, S., Kwon, O., Kim, S., Choung, S., 2014. The importance of a normal breathing pattern for an effective abdominal-hollowing maneuver in healthy people: an experimental study. J. Sport Rehabil. 23 (1), 12–17. http://dx.doi.org/10.1123/jsr.2012-0059.

Hammer, J., Newth, C.J.L., 2009. Assessment of thoraco-abdominal asynchrony. Paediatr. Respir. Rev. 10 (2), 75–80. http://dx.doi.org/10.1016/j.prrv.2009.02.004.

Herráez, J., Martínez-Llario, J., Coll, E., Rodríguez, J., Martin, M.T., 2013. Design and calibration of a 3D modeling system by videogrammetry. Meas. Sci. Technol. 24 (3), 035001. http://dx.doi.org/10.1088/0957-0233/24/3/035001.

Hunt, R.J., 1986. Percent agreement, Pearson's correlation, and kappa as measures of inter-examiner reliability. J. Dent. Res. 65 (2), 128–130. http://dx.doi.org/10.1177/00220345860650020701.

Kaneko, H., Horie, J., 2012. Breathing movements of the chest and abdominal wall in healthy subjects. Respir. Care 57 (9), 1442–1451. http://dx.doi.org/10.4187/respcare.01655.

Lewit, Karel, 1999. Manipulative Therapy in Rehabilitation of Locomotor System (Third). Reed Educational and Professional Publishing Ltd, Butterworth-Heinemann.

Perri, M., Halford, E., 2004. Pain and faulty breathing: a pilot study. J. Bodyw. Mov. Ther. 8 (4), 297–306. http://dx.doi.org/10.1016/S1360-8592(03)00085-8.

Pryor, J.A., Prasad, S.A., 2002. Physiotherapy for Respiratory and Cardiac Problems: Adults and Paediatrics, third ed. Churchill Livingstone, Edinburgh.

Sarro, K.J., Silvatti, A.P., Aliverti, A., Barros, R.M.L., 2009. Proposition and evaluation of a novel method based on videogrammetry to measure three-dimensional rib motion during breathing. J. Appl. Biomech. 25 (3), 247–252.

Shoukri, M.M., Asyali, M.H., Donner, A., 2004. Sample size requirements for the design of reliability study: review and new results. Stat. Methods Med. Res. 13 (4), 251–271. http://dx.doi.org/10.1191/0962280204sm365ra.

Viera, A.J., Garrett, J.M., 2005. Understanding interobserver agreement: the kappa statistic. Fam. Med. 37 (5), 360–363.