

Reliability of ultrasonic diaphragm thickness measurement in mechanically ventilated infants and children: A pilot study

M N Terhart,¹ BSc (Physiotherapy); S Hanekom,¹ PhD; A Lupton-Smith,¹ PhD (Paediatrics); B Morrow,² PhD

¹ Division of Physiotherapy, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa

² Department of Paediatrics and Child Health, Faculty of Health Sciences, University of Cape Town, South Africa

Corresponding author: M N Terhart (maxineterhart@gmail.com)

Background. Diaphragmatic atrophy in mechanically ventilated infants and children may be due to ventilator-induced diaphragmatic dysfunction, which could lead to extubation failure. Ultrasound may be used as a means by which diaphragmatic atrophy can be reliably identified. There are currently no data reporting on the use of ultrasound to monitor diaphragm atrophy in the paediatric population.

Objectives. To assess the inter- and intra-rater reliability of using ultrasound to measure diaphragm thickness in mechanically ventilated infants and children.

Method. Diaphragm thickness measurements were compared between two individual researchers for inter-rater reliability and between multiple measurements from a single researcher for intra-rater reliability. Measurements were compared using Intraclass correlation coefficients and Bland-Altman plots.

Results. Results indicated excellent reliability between measurements for both inter- and intra-rater reliability, with slightly better reliability for intra-rater compared with inter-rater reliability. Intraclass correlation coefficients for inter-rater reliability were between 0.77 and 0.98, and 0.94 for intra-rater reliability.

Conclusion. Ultrasound measurements of diaphragm thickness can be used to reliably measure diaphragm thickness in mechanically ventilated infants and children. This modality could therefore be used as a reliable outcome measure for future clinical research studies to establish the relationship between ventilator-induced diaphragmatic atrophy and children who are at risk for extubation failure.

S Afr J Crit Care 2018;34(1):22-27. DOI:10.7196/SAJCC.201.v34i1.344

Ultrasound (US) is a non-invasive tool which can be used to visualise and evaluate the structure and function of the diaphragm.^[1] The non-invasive and risk-free nature of US also makes it appropriate to use with critically ill infants and children.^[2] The diaphragm is the main muscle of respiration and infants rely even more so on the diaphragm for respiration than adults owing to physiological and anatomical differences in the respiratory system.^[3,4] Research has shown that even brief periods of mechanical ventilation (MV) can lead to diaphragmatic atrophy in adults.^[5,6] Diaphragmatic atrophy is associated with diaphragm dysfunction,^[7] which has been linked to weaning difficulties and extubation failure (EF) in adults.

Ventilator-induced diaphragmatic dysfunction (VIDD), originally defined by Vassilakopoulos and Petrof,^[8] refers to the loss of force-generating capacity of the diaphragm as a result of unloading during MV. Depending on the mode of ventilation, unloading leads to diaphragmatic inactivity which has been shown to be a direct cause of VIDD.^[8] Signs of VIDD include a decrease in diaphragm thickness and activity.^[7] Although many other factors may affect the function of the diaphragm, MV has an independent negative impact on the diaphragm muscle.^[7] Other contributing factors to VIDD include underlying disease, medication, sepsis and metabolic imbalances.^[8,10] Diaphragmatic dysfunction may lead to weaning difficulty or failure to wean from MV, which in turn predisposes to poorer clinical outcomes, including prolonged duration of MV and intensive care unit (ICU) stay, and increased healthcare costs.^[8,9]

Diaphragmatic atrophy can occur even during brief periods of MV and can start within 18 - 69 hours of initiation of MV.^[5,6] Schepens *et al.*^[7] studied the progression of diaphragm atrophy in adult ventilated patients using US and reported that after just one day of MV there was already a substantial decrease in diaphragm thickness, with the biggest decrease occurring within the first 72 hours of MV. Although diaphragm atrophy during full MV support has been well described in adults, paediatric data are lacking and research in this field is warranted.^[9]

US is a portable and non-invasive tool which can be used to visualise and evaluate the structure and function of the diaphragm with no risk of ionising radiation. It is relatively easy to use, making it an appealing modality to use for children and critically ill patients.^[1,2] The costal part of the diaphragm is relatively close to the surface of the skin; therefore a high resolution and low penetration US probe (between 7.5 - 15 MHz) can be used to effectively visualise and measure the thickness of the diaphragm.^[11] The diaphragm can be measured using US at different lung volumes, during tidal breathing and during static inspiratory efforts.^[11] The measurement is rapid and produces accurate and reproducible results in most adult individuals.^[11] To the best of our knowledge, there are no studies using US to monitor VIDD in the paediatric population.

The present study aimed to evaluate the inter- and intra-rater reliability of using ultrasound to measure diaphragm thickness in mechanically ventilated infants and children in the paediatric ICU (PICU) for future use in research and clinical practice.

Methods

A prospective observational study was conducted in the PICU at Red Cross War Memorial Children's Hospital (RCWMCH), Cape Town, South Africa – a 22-bed multidisciplinary PICU in a tertiary academic hospital. The PICU admits approximately 1 400 children per annum, with an estimated 9% mortality rate. Ethical approval was obtained from the Human Research Ethics Committees of the University of Cape Town and the University of Stellenbosch (ref. no. 179/2013). Written informed consent was obtained from the parent or legal guardian of the child.

All intubated and mechanically ventilated patients between 0 and 2 years of age who were admitted to the PICU of RCWMCH were eligible for inclusion if they had received MV for 24 hours. Patients were excluded from this study if they: were chronically ventilated; were likely to be extubated within 24 hours of intubation; were haemodynamically unstable; were prematurely born neonates; had thoracic burns and/or dressings covering the thoracic area; had current or previous diaphragmatic abnormalities, hernias or known diaphragmatic paralysis; were diagnosed with neuromuscular diseases or spinal paralysis; had marked abdominal distension; or if were unable to be turned from prone into the supine position for therapeutic or oxygenation purposes.

Procedure

Researchers underwent a short training session on the use of diaphragmatic US by a qualified professional, prior to starting this study.

After obtaining informed consent from the patients' parents or legal guardians, baseline data and US measurements were recorded using standardised data capture sheets. Diaphragm thickness measurements were obtained once a day for two days at roughly the same time of the morning by the same researcher (MNT) and co-researcher (ALS); provided the child did not have any new condition which excluded him/her from the study (e.g. new development of ascites or abdominal distension).

Measurements were stopped if the patient became uncomfortable or distressed, or if any other urgent medical care was required and resumed afterward if the patient was deemed stable by the attending clinician.

Ultrasound methodology

We used bedside US methodology as described previously by Gibson *et al.*^[11] A high-resolution, low-penetration probe (between 6 and 13 MHz) was placed over one of the lower intercostal spaces, between the 10th -12th intercostal space, in the mid-axillary line on the right side of the chest and perpendicular to the chest wall in the zone of apposition (ZAP). These measurements were taken using the Sonosite EDGE Ultrasound System (Sonosite Fujifilm Inc., USA) in the supine position with the head up at ~30 degrees, as per standard PICU practice. US was used in two-dimensional B-mode for identification of the diaphragm, followed by an M-mode recording of consecutive breaths in order to accurately identify the exact point of end-expiration and end-inspiration as depicted in Figs 1 A - C. The diaphragm is identified in B-mode as two outer echogenic layers of parietal pleura and peritoneum. A less bright echo may also be visible, which arises from a layer of connective tissue and vessels which runs through the middle of the muscle.^[11] The freeze-frame function was then used to measure the thickness of the diaphragm at end-expiration and end-inspiration in centimetres (cm) using the on-screen callipers, as displayed in Figs 1 D and E). For consistency between researchers, when measurements were taken, the callipers were placed in the middle of the upper and lower boundaries of the diaphragm muscle by each researcher. Measurements

were taken at the end of inspiration and expiration, respectively, and the average thickness over three breaths was used for analysis.

The thickness at end-inspiration and end-expiration was documented on the daily data capture sheet for each participant. After completion of measurements, the researcher ensured that the patient's chest was clean from US gel, the patient was returned to the nursing position and was comfortable and stable.

Inter-rater reliability

Measurements of diaphragm thickness using US were taken on the same patient for two consecutive days by the primary researcher and co-researcher. The area of measurement on the chest of the patient was cleaned between measures, i.e. there were no markings or US gel residue on the patient's skin to indicate the placement of the US probe by the first researcher.

Intra-rater reliability

The primary researcher took measurements of diaphragm thickness using the same ultrasound machine at two different times (~30 minutes apart) on the same participant. The area of measurement on the chest of the patient was cleaned between measures (i.e. there were no markings or US gel residue on the patient's skin to indicate the previous placement of the US probe).

Statistical analysis

Intraclass correlation coefficients (ICCs) were calculated using SPSS version 24 (IBM Corp., USA) and Microsoft Excel software was used to create Bland-Altman plots using the measurements of end-expiration and end-inspiration. As described previously, an ICC <0.4 indicates poor agreement, between 0.40 and 0.59 indicates fair agreement, between 0.60 and 0.74 indicates good agreement, and between 0.75 and 1.00 indicates excellent agreement.^[12]

Results

A total of 145 PICU patients were screened by the primary researcher for inclusion into the study from 18 July 2016 - 31 August 2016. Twelve patients met the inclusion criteria and, owing to unobtained consent, only 5 (3 male) patients were ultimately included in this reliability study. Baseline demographic data of included participants are shown in Table 1.

Inter-rater reliability

The ICCs for the average of end-expiration on day 1 and day 2 showed an excellent correlation, with the coefficients being 0.77 and 0.98, respectively. The ICC for the end-inspiration readings for day 1 and day 2 also showed excellent agreement, with the ICCs being 0.80 and 0.91 (Table 2).

The Bland-Altman plots showed broader limits of agreement (LOA) for day 1 end-expiration compared with day 2; however, all the points lie within the upper and lower LOA on both days. For day 1 end-expiration, the points lie slightly further from the mean line, but are distributed evenly above and below the mean line, indicating little to no bias; whereas on day 2 all the points lie much closer to the mean line but are not as evenly distributed above and below the mean line (Fig. 2). For end-inspiration on both day 1 and day 2, the LOA are narrow and all the points lie between the upper and lower LOA. On day 2 the points are more evenly distributed above and below the mean line than day 1. All the points for day 1 and day 2 lie relatively close to the mean line besides one outlier that lies closer to the lower LOA on day 1 (Fig. 2).

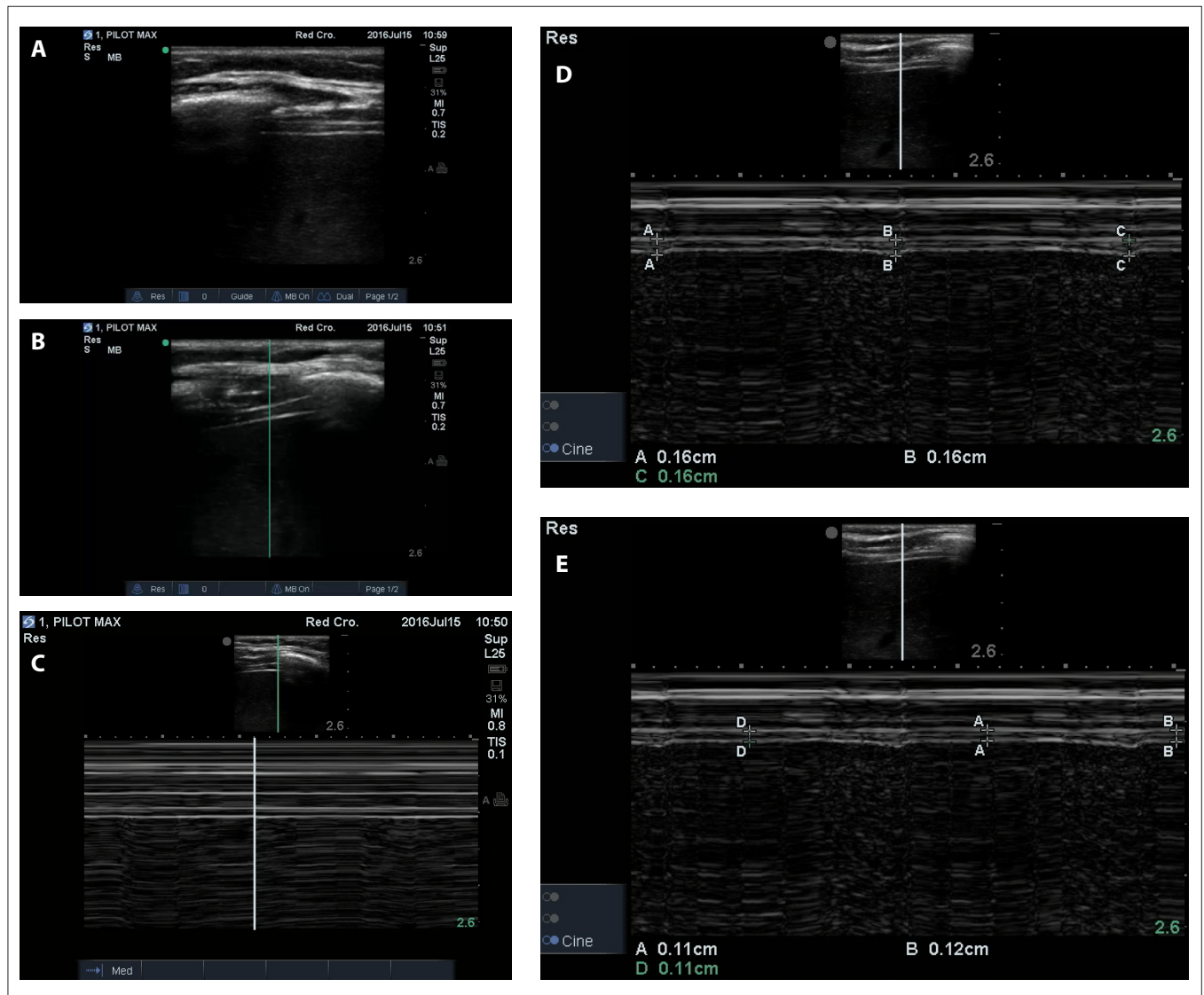


Fig. 1. Ultrasound methodology used to measure diaphragm thickness. (A) Identification of diaphragm using B-mode. (B) M-mode line (green) used to create a cineloop. (C) M-mode cineloop recording of consecutive breaths. (D) Freeze-frame of M-mode cineloop measurements of end-inspiration. (E) Freeze-frame of M-mode cineloop and measurements of end-expiration.

Table 1. Included participant characteristics

Participant	Age (months)	Sex	Diagnostic category	Duration of MV (days)
1	7	M	GIT (laparotomy)	3
2	9	F	Pulmonary (pneumonia)	4
3	6	M	Cardiac (thoracotomy)	1
4	16	M	Cardiac (thoracotomy)	1
5	1	F	Cardiac (cardiomegaly)	3
Mean	7.8			2.4

M = male; F = female; MV = mechanical ventilation; GIT = gastrointestinal.

Intra-rater reliability

The ICCs for end-expiration and end-inspiration both showed an excellent correlation at 0.94 between measures 1 and 2 (Table 3).

The Bland-Altman plots for end-inspiration had slightly narrower LOA than end-expiration. All the points fall between the upper and lower LOA for both end-expiration

and end-inspiration; and all the points are distributed equally above and below the mean line which indicates little to no bias (Fig. 3).

Discussion

The aim of this study was to assess the reliability of obtaining diaphragm thickness measurements using US in a mechanically ventilated paediatric population. The inter-

and intra-rater reliability of US was evaluated by comparing measurements of diaphragm thickness between researchers and by comparing one researcher's measurements on two separate occasions, respectively. Ultimately, it is hoped that this tool could be incorporated into clinical practice to identify diaphragm atrophy in mechanically ventilated infants and children, which could be due to

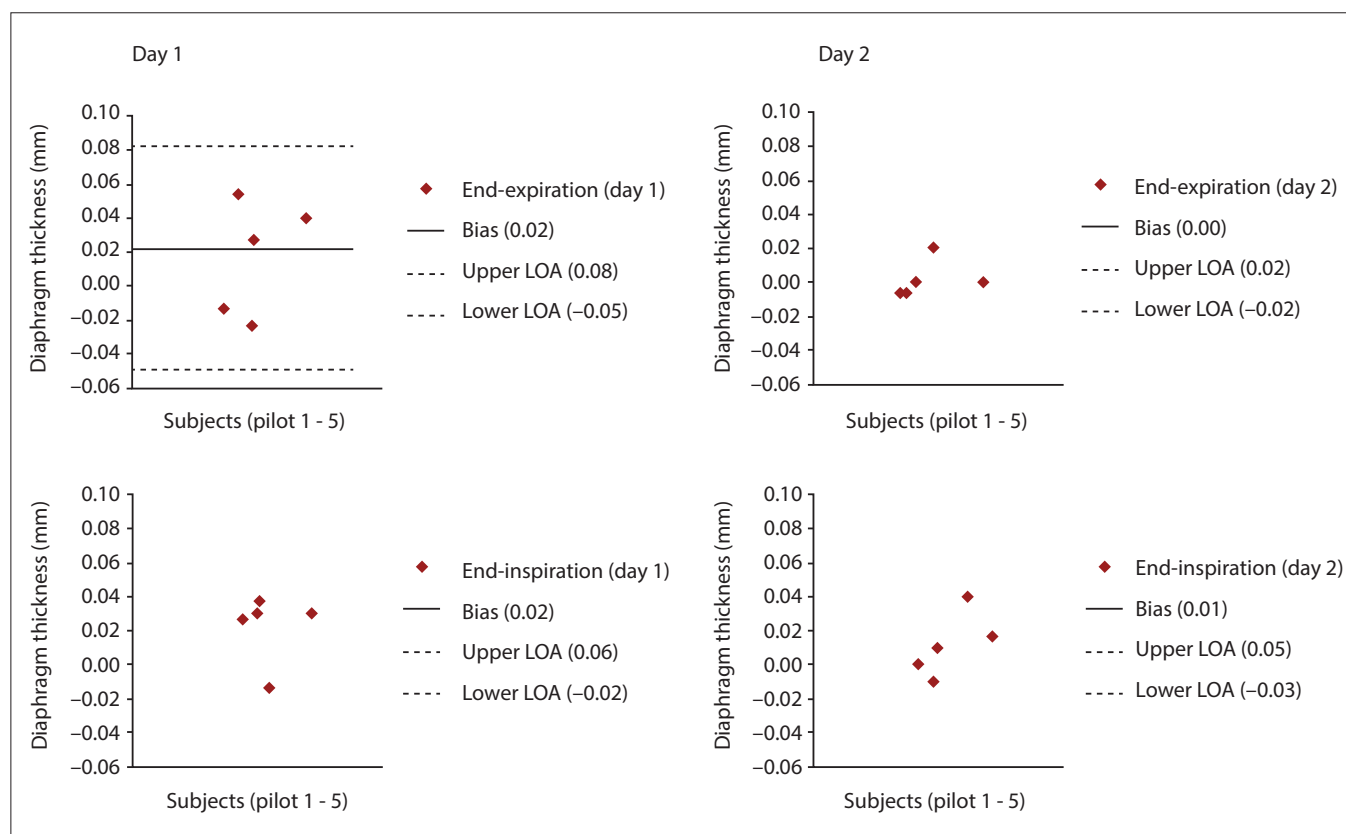


Fig. 2. Bland-Altman plots for inter-rater reliability.

Table 2. Inter-rater reliability intraclass correlation coefficients

Average measures	Day	Intraclass correlation*	95% CI
End-expiration	1	0.766 [†]	-0.620 - 0.975
	2	0.983 [†]	0.836 - 0.998
End-inspiration	1	0.798 [†]	-0.272 - 0.978
	2	0.912 [†]	0.367 - 0.990

CI = confidence interval.

*Two-way mixed effects model where people effects are random and measures effects are fixed.

[†]Type A intraclass correlation coefficients using an absolute agreement definition. This estimate is computed assuming the interaction effect is absent, because it cannot be estimated otherwise.

Table 3. Intra-rater reliability intraclass correlation coefficients

Average measures	Intraclass correlation*	95% CI
End-expiration	0.942 [†]	0.464 - 0.994
End-inspiration	0.936 [†]	0.405 - 0.993

CI = confidence interval.

*Two-way mixed effects model where people effects are random and measures effects are fixed.

[†]Type A intraclass correlation coefficients using an absolute agreement definition. This estimate is computed assuming the interaction effect is absent, because it cannot be estimated otherwise.

VIDD, or other mechanisms of muscle thinning such as muscle catabolism during MV.

Based on the results of the tests of inter-rater reliability, it is evident that there is excellent agreement between researchers. The observed improvement in agreement between researchers from day 1 to 2 two could indicate that the accuracy of measurements

will improve over time with practice and experience, as positioning of the US probe is patient-specific owing to differences in chest size, age, etc. Measurements of end-inspiration showed slightly better agreement than measurements of end-expiration, which could be due to better visualisation of the diaphragm during end-inspiration as the diaphragm

thickens, than during end-expiration; although, on day 1, the measurements of end-inspiration showed better agreement than measures of end-expiration, while end-expiration showed better correlation than end-inspiration on day 2.

The intra-rater reliability also showed excellent correlation, with both end-inspiration and end-expiration having very high ICC values. There was slightly better agreement between the first and second readings for end-expiration than for end-inspiration, which was consistent with what was found between researchers on day 2. This could be due to the variability of diaphragm thickening at end-inspiration between spontaneous and non-spontaneous breaths (as a consequence of active v. passive diaphragm thickening), while there is unlikely to be a difference between end-expiratory thickness, as the diaphragm will return to resting thickness after both spontaneous and non-spontaneous breaths. Another reason for better reliability with end-expiration measurements could be that expiration is longer than inspiration and is therefore easier to identify and measure the point of end-expiration than end-inspiration.

A possible reason for the observed difference in measures between researchers may relate to individual differences in calliper positioning during the process of measuring the diaphragm muscle. Ideally, for both inter- and intra-rater reliability, the measurements

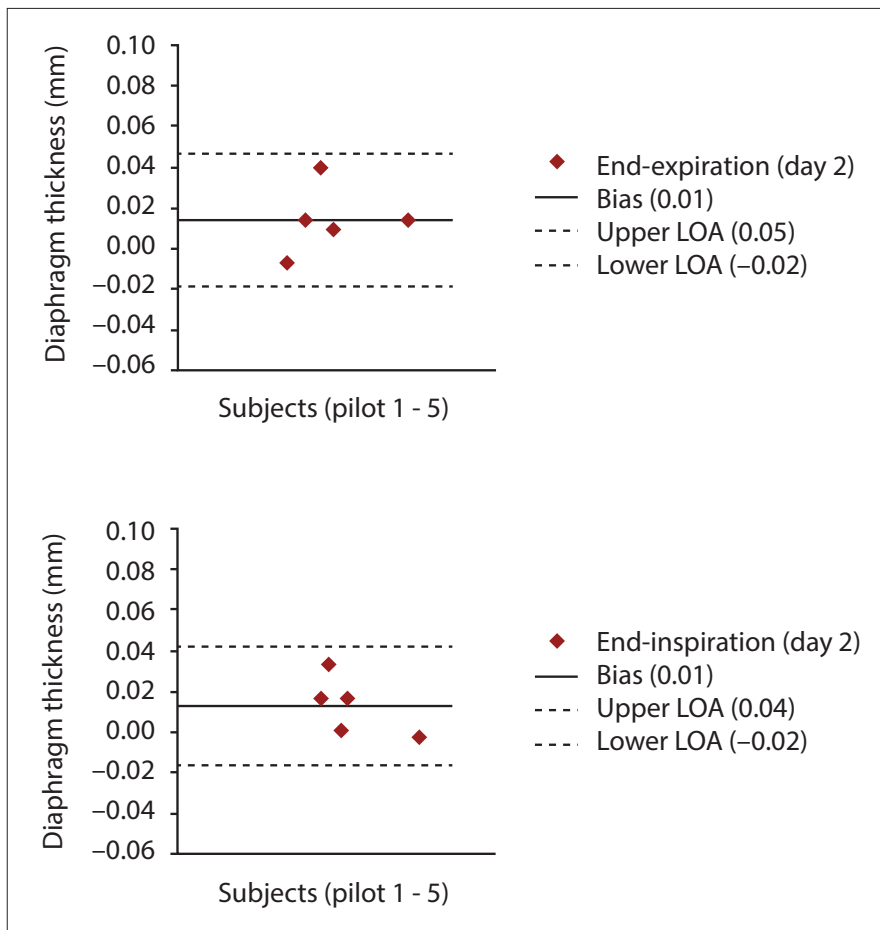


Fig. 3. Bland-Altman plots for intra-rater reliability.

of diaphragm thickness would be done simultaneously. However, given the specificity of the area and placement of the US probe on the chest wall, this was not possible. Goligher *et al.*^[2] reported that there is a direct relationship between the level of contractile activity of the diaphragm, i.e. activity during spontaneous and non-spontaneous breaths, and changes in diaphragm thickness in adults. Thus, given breath-by-breath variability, and differences in spontaneous v. non-spontaneous breaths over time, it was expected that measurements obtained would differ even for the same patient between breaths, which probably affected the results obtained for both inter- and intra-rater reliability.

Study limitations

This study was performed on a small sample of five infants and children, using only two raters. Future research should aim to increase the participant sample size and rater numbers for improved reliability of the results and to establish the generalisability of these measurements to a larger sample of patients. Researchers could not accurately assess whether the measured breath was spontaneous or non-spontaneous

during measurements, which is a limitation of this study. This would be important to assess owing to the suspected difference in diaphragm thickness during a spontaneous (patient-driven) breath v. a non-spontaneous (ventilator-driven) breath. A suggestion for future research or practice of this method would be to assess and measure diaphragm thickness only during spontaneous breaths, which may also provide an indication of weaning and extubation readiness. All measurements were manually recorded by the primary researcher while the co-researcher was blind to all measurements, thus there was a possibility of measurement bias by one researcher. Given the relatively short time between inter-rater measurements, it is possible that there was some recall bias; however, the results of this study demonstrate that the methodology and the repeatability of these measurements were reliable and can be used effectively in the clinical research setting.

Conclusion

The results of this study indicate an excellent inter-rater and intra-rater reliability when using US as a measure for diaphragm thickness in mechanically ventilated infants and children.

Using US to measure diaphragm thickness was not only an appropriate and safe modality, due to the nature of US being portable, non-invasive and radiation-free, but it is also feasible to use in this population. Given the negative effect of MV on the diaphragm, measures of diaphragm thickness over a period of MV may be used to identify patients who have developed VIDD. Using these measures and plotting a trend may make it possible for clinicians to identify VIDD and the degree of diaphragm dysfunction. Thus, these measures may be useful as a predictor for infants and children who may be at risk for weaning difficulties and EF. Ultimately, it is hoped that this measure could be implemented in the clinical setting to identify mechanically ventilated paediatric patients with diaphragmatic atrophy or dysfunction

Acknowledgements. MNT would like to thank the PICU personnel for allowing her to conduct the research in their facility, as well as for allowing her to use the US machine, specifically Dr Beyra Rossouw, who also took the time to train her on the US machine.

Author contributions. MNT: primary researcher and author. SH and BM: supervised the project, made valuable contributions to the study and edited drafts, as well as the final version of manuscript. AL-S: co-researcher, assisted with data collection and analysis, as well as editing of the drafts and final manuscript.

Funding. The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF. MNT would like to acknowledge the South African Society of Physiotherapy (SASP) for providing funding which supported this research.

Conflicts of interest. None.

1. Sarwal A, Walker FO, Cartwright MS. Neuromuscular ultrasound for evaluation of the diaphragm. *Muscle Nerve* 2013;47(3):319-329. <https://doi.org/10.1002/mus.23671>
2. Goligher EC, Fan E, Herridge MS, et al. Evolution of diaphragm thickness during mechanical ventilation: Impact of inspiratory effort. *Am J Respir Crit Care Med* 2015;192(9):1080-1088. <https://doi.org/10.1164/rccm.201503-0620OC>
3. Anraku M, Shargall Y. Surgical conditions of the diaphragm: anatomy and physiology. *Thorac Surg Clin* 2009;19(4):419-429. <https://doi.org/10.1016/j.thorsurg.2009.08.002>
4. Siren PMA, Siren MJ. Critical diaphragm failure in sudden infant death syndrome. *Ups J Med Sci* 2011;116(2):115-123. <https://doi.org/10.3109/03009734.2010.548011>
5. Ferguson LP, Walsh BK, Munhall D, Arnold JH. A spontaneous breathing trail with pressure support overestimates readiness for extubation in children. *Pediatr Crit Care Med* 2011;12(6):e330-e335. <https://doi.org/10.1097/pcc.0b013e3182231220>
6. Levine S, Nguyen T, Taylor N, et al. Rapid disuse atrophy of diaphragm fibres in mechanically ventilated humans. *N Engl J Med* 2008;358(13):1327-1335. <https://doi.org/10.1056/nejmoa070447>
7. Schepens T, Verbrugge W, Dams K, Corthouts B, Parizel PM, Jorens PG. The course of diaphragm atrophy in ventilated patients assessed with ultrasound: A longitudinal cohort study. *Crit Care* 2015;19(1):422. <https://doi.org/10.1186/s13054-015-1141-0>

8. Vassilakopoulos T, Petrof BJ. A stimulating approach to ventilator-induced diaphragmatic dysfunction. *Am J Respir Crit Care Med* 2017;195(3):281-282. <https://doi.org/10.1164/rccm.201608-1619ED>
9. Emeriaud G, Larouche A, Ducharme-Crevier L, et al. Evolution of inspiratory diaphragm activity in children over the course of the PICU stay. *Intensive Care Med* 2014;40(11):1718-1726. <https://doi.org/10.1007/s00134-014-3431-4>
10. Petrof BJ, Jaber S, Matecki S. Ventilator-induced diaphragmatic dysfunction. *Curr Opin Crit Care* 2010;16(1):19-25. <https://doi.org/10.1097/mcc.0b013e328334b166>
11. Gibson GJ, Whitelaw W, Sifakas N, et al. ATS/ERS statement on respiratory muscle testing. *Am J Respir Crit Care Med* 2002;166(4):518-624. <https://doi.org/10.1164/rccm.166.4.518>
12. Cicchetti D V. The precision of reliability and validity estimates re-visited: Distinguishing between clinical and statistical significance of sample size requirements. *J Clin Exp Neuropsychol* 2011;23(5):695-700. <https://doi.org/10.1076/jcen.23.5.695.1249>

Accepted 19 June 2018.