

## Reply to Reviewer's comments

**Reviewer's code:**00502907

1. Measurement of diaphragm thickness is relevant at the bedside as its atrophy and dysfunction is an important predictor of extubation success. Previous studies have compared the diaphragm thickness at end expiration and fraction shortening with other parameters used to assess fitness for extubation and found it to be useful (Ref: DiNino E, Gartman EJ, Sethi JM, McCool FD. Diaphragm ultrasound as a predictor of successful extubation from mechanical ventilation. *Thorax*. 2014 May;69(5):423–7.) However, the reproducibility of these measurements was not independently assessed. The aim of our study was to assess whether the measurement of the diaphragm thickness by critical care physician is reliable after adequate training so that it could be done at the bedside and used in future to assess fitness for extubation.
2. Reviewer's comment that these measurements need to be compared to extubation success and time on ventilator is appropriate and needs to be evaluated in further studies. We are also planning to design a future study looking into this outcome.
3. The aim of our study was to find out whether these measurements taken by the critical care physicians are reproducible. We have conducted similar study in ICU for measurement of arm and quadriceps muscle thickness for assessment of nutritional status. (Ref: Hadda V, Kumar R, Dhungana A, Khan MA, Madan K, Khilnani GC. Inter- and intra-observer variability of ultrasonographic arm muscle thickness measurement by critical care physicians. *J Postgrad Med*. 2017 Jul-Sep;63(3):157-161.) (Ref: Hadda V, Khilnani GC, Kumar R, Dhungana A, Mittal S, Khan MA, et al. Intra- and inter-observer reliability of quadriceps muscle thickness measured with bedside ultrasonography by critical care physicians. *Indian J Crit Care Med* 2017;21:448-52. )

**Reviewer's code:**02454185

**Ques 1.** It is unclear from the description on how many measurements were taken by the same observer for the same patient. And what is the interval between these measurements?

**Ans:** Three observations were taken by the same observer for the same patient and these three measurements were used to assess intraobserver agreement. There was no time gap between the three measurements. The necessary changes have been made and highlighted in the 'abstract' and 'material and methods' sections of the revised manuscript.

**Ques 2:** The measurements were taken only at end-expiration, which is not enough. In the assessment of the diaphragm function for the evaluation of weaning outcome, we are interested in the change of the diaphragm thickness during respiratory cycle, thus it is also interesting to measure diaphragm in other stages of the respiratory cycle.

**Ans:** It is rightly pointed out that the measurement of diaphragm thickness at both end expiration and end inspiration and the fraction shortening during the respiratory cycle helps in true assessment of diaphragm function for evaluation of weaning outcome. For the measurement of baseline thickness, measurements at end expiration were chosen as it is least affected by the ventilator parameters and easiest to compare. Thickness at end inspiration is also dependent on the baseline thickness of the diaphragm, which is affected by various factors such as body weight and time on mechanical ventilation.

**Ques 3.** I also suggest to use kappa statistics, which is a standard in the report of agreement between observers.

**Ans:** The interobserver and intraobserver agreement can be measured using kappa value (categorical data), weighted kappa value (ordinal data) and intraclass correlation coefficient (continuous data). (Ref: Mandrekar JN. Measures of Interrater Agreement. J Thorac Oncol.

2011;6: 6–7 ). We have used the intraclass correlation coefficient for comparison of both intrarater and interrater agreements as diaphragm thickness is a continuous measurement. We have also compared the observations using the bland and altman plots for visual reflection of high agreements both between observers and measurements by the same observer.

**Ques 4:** The software used for the statistical analysis should be explicitly reported

**Ans:** The necessary changes have been made in the revised manuscript and highlighted in the ‘Material and methods’ section.

**Ques 5:** Factors that may influence the reproducibility of the measurements were not reported. these factors included, but not limited to, the BMI, the presence of pleural effusion, abdominal hypertension, and lung volume (COPD vs. NON-COPD)

**Ans:** We acknowledge the fact that we did not specifically look into the factors that influence the thickness of diaphragm like BMI and presence of obstructive airway disease. It would be interesting to see how the diaphragm would be affected in these different conditions during mechanical ventilation. As we only aimed to evaluate the reproducibility of the method, we did not compare the results in different subgroups, which would require a larger sample size too.

**Ques 6.** Baseline characteristics of your cohort should be presented in a separate table.

**Ans:** The necessary changes has been made in the ‘results’ section of the revised manuscript and highlighted

**Reviewer's code:**03345330

**Ques 1.**The method used to measure diaphragm thickness might be too simple to assess the diaphragm function in this study. I suggest the authors should provide the related reference literature and the validation of the method used in the study. Importantly, diaphragm thickness always is measured by M-mode in clinical practice, and M-mode could assess the function of the diaphragm [1-4]. Sometimes, the line of pleural and peritoneal membranes is difficult for a physician to identify without the indicator of Liver or spleen. Moreover, the end of expiration also is difficult to determine without the M-mode. I suggest the method of diaphragm thickness need further clarification. This is a potential cofounder.

**Ans:** USG can be used for the assessment of diaphragm function by both B-Mode and M-Mode. B mode can be used to measure diaphragm thickness at end inspiration and expiration and the fraction shortening, which is a surrogate of its function (Ref: DiNino E, Gartman EJ, Sethi JM, McCool FD. Diaphragm ultrasound as a predictor of successful extubation from mechanical ventilation. *Thorax*. 2014 May;69(5):423–7. ) Motion mode ultrasonography is used to measure the excursion of the diaphragm and its function, and to assess whether paradoxical excursion is present. We have used the B mode to measure the diaphragm thickness at end expiration and compared the intraobserver and interobserver reproducibility. This limitation of not having used the M Mode USG has been acknowledged and included in the discussion section with relevant references and highlighted.

**Ques 2.** The authors just included 10 patients for the comparison of the agreement between two observers with Bland-Altman test. It is unreasonable for a small sample to use Bland and Altman test. I suggest the pair-t test or correlation analysis might be better.

**Ans:** We only assessed the interobserver reproducibility in 10 randomly selected patients which is a limitation of our study and has been mentioned in the discussion on limitations. The interobserver agreement was compared with intraclass correlation coefficient which was found to be high. Bland and altman analysis was used as a secondary measure of agreement and the graphs were depicted for visual reflection of high agreements both between observers and measurements by the same observer.

**Ques 3.** Since studies have supported reproducibility of diaphragm thickness, I think the aim of the study should focus on the ICU doctor could be trained and make a good measurement of diaphragm thickness in a short time. This points might be interesting. So, the golden standards measurement by the teacher should be included in the study.

**Ans:** We agree to the reviewer's comment that the golden standard measurement by the radiologist who trained us in measurement of diaphragm thickness would have added more weightage to the present study. We work in a setup where there is a dearth of radiologist to be available in the ICU at all points of time. The aim of this study was to assess whether the diaphragm measurement by critical care physicians is reproducible after an adequate training.