Journal of Critical Care 34 (2016) 111–115

Contents lists available at ScienceDirect

Journal of Critical Care

journal homepage: www.jccjournal.org

The validity and reliability of the clinical assessment of increased work of breathing in acutely ill patients☆,☆☆

Aiman Tulaimat, MD [a,*], Aiyub Patel, MD [b], Mary Wisniewski, RN [c], Renaud Gueret, MD [d]

[a] Cook County Health and Hospitals System, Pulmonary, Critical Care, and Sleep Medicine, 1900 West Polk St, Room 1404, Chicago, IL 60612
[b] Unitypoint Methodist Hospital, Pulmonary, Critical Care, and Sleep Medicine, Peoria, IL 61602
[c] John H. Stroger Hospital of Cook County, Chicago, IL 60612
[d] John H. Stroger Hospital of Cook County, Pulmonary, Critical Care, and Sleep Medicine, Chicago, IL 60612

ARTICLE INFO

Keywords:
Respiratory insufficiency
Physical examination
Reproducibility of results
Sensitivity and specificity
Dyspnea
Distress

ABSTRACT

Background: Mechanical ventilation is frequently indicated to reduce the work of breathing. Because it cannot be measured easily at the bedside, physicians rely on surrogate measurements such as patient appearance of distress and increased breathing effort.
Objective: We determined the validity and reliability of subjectively rating the appearance of respiratory distress and the reliability of 11 signs of increased breathing effort.
Subjects: The study included consecutive, acutely ill patients requiring various levels of respiratory support.
Methods: Blinded to each other's observations, a fellow and a critical care consultant rated the severity of distress (absent, slight, moderate, severe) after observing subjects for 10 seconds and then determined the presence of the signs of increased breathing effort.
Results: A total of 149 paired examinations occurred 6 ± 6 minutes apart. The rating of respiratory distress correlated with oxygenation, respiratory rate, and 9 signs of increased work of breathing. It had the highest intraclass correlation coefficient (0.69; 95% confidence interval, 0.59-0.78). Rating distress as moderate to severe had a sensitivity of 70%, specificity of 92%, and positive likelihood ratio of 8 for the presence of 3 or more of hypoxia, tachypnea, and any sign of increased breathing effort. Agreement was moderate ($\kappa = 0.53$-0.47) for rating of distress, nasal flaring, scalene contraction, gasping, and abdominal muscle contraction, and fair ($\kappa = 0.36$-0.23) for sternomastoid contraction, tracheal tug, and thoracoabdominal paradox.
Conclusion: Assessing the increased work of breathing by rating the severity of respiratory distress based on subject appearance is a valid and moderately reliable sign that predicts the presence of serious respiratory dysfunction. The reliability of the individual signs of increased breathing effort is moderate at best.

1. Introduction

In acutely ill patients, mechanical ventilation is usually instituted when the dysfunction in respiratory system leads to serious abnormalities in blood gases or increased work of breathing [1]. We interpret measurements of arterial blood gases according to normative data and physiologic principles. We cannot as objectively measure the work of breathing. Instead, we have to form a gestalt after we assess a patient's appearance of distress, measure the vital signs, and examine for the physical signs of increased breathing effort (eFigure) [2–7].

A patient's appearance of distress reflects an increase in the work of breathing in response to worsening respiratory mechanics [8]. Although this appearance alerts us to patients that need urgent attention [9], very little is known about its validity. Gilston [10] considered obvious dyspnea as one of the signs of acute respiratory failure, and Campbell [11] incorporated the look of fear in a scale validated to measure respiratory distress in terminally ill patients.

The reliability of assessing distress has received limited but recent attention. Two recent studies have evaluated the agreement on the presence of respiratory distress during weaning from mechanical ventilation [12,13]. A third study assessed agreement between patients and their nurses and physicians on the severity of dyspnea during weaning [14]. Our knowledge about the reliability of assessing distress at other phases of acute illness is very limited [15].

☆ Funding: none.
☆☆ Author contribution: concept and design: AT, MW, and AP; data collection: AT, AP, and RG; analysis: AT; drafting manuscript: AT; critical review of the manuscript: AT, AP and RG.
Conflict of interest: The authors have no potential conflicts of interest involving the work under consideration for publication (during the time involving the work, from initial conception and planning to present), any "relevant financial activities outside the submitted work" (over the 3 years before submission), and any "other relationships or activities that readers could perceive to have influenced or that give the appearance of potentially influencing" what is written in the submitted work (based on all relationships that were present during the 3 years before submission)
* Corresponding author: Tel.: +1 312 864 7387 (office); fax: +1 312 864 7394.
E-mail addresses: atulaimat@cookcountyhhs.org (A. Tulaimat), Aiyub.Patel@unitypoint.org (A. Patel), mwisniewski@cookcountyhhs.org (M. Wisniewski), rgueret@cookcountyhhs.org (R. Gueret).

http://dx.doi.org/10.1016/j.jcrc.2016.04.013
0883-9441/

Similarly, few studies have evaluated the reliability of the physical signs of increased breathing effort such as nasal flaring, contraction of the sternomastoid, and thoracoabdominal paradoxical movement. These studies have limitations that reduce our ability to generalize their findings to the assessment of all acutely ill patients [6,7,13,16,17].

The limited understanding of the reliability and validity of our assessment of increased work of breathing is concerning for 2 reasons. First, research in critical care where mechanical ventilation is an outcome or an intervention usually includes a list of its indications. Increased effort of breathing and distress are common indications. If these indications are not reliable, the generalizability of the findings becomes questionable. Second, communication between physicians caring for acutely ill patients frequently includes an assessment of the work of breathing. The lack of reliability makes it difficult to determine whether differences between observations are due to measurement error or represent real change in a patient's condition.

Hence, we aim to assess in acutely ill patients the validity and reliability of some elements of the clinical assessment of the work of breathing. Specifically, we sought to determine the agreement between physicians when they subjectively rate the level of respiratory distress based on a patient's appearance and the ability of this rating to predict the presence of other signs of respiratory dysfunction. In addition, we determined the agreement between physicians on the presence of the physical signs of increased breathing effort.

2. Methods

2.1. Setting

The study was conducted in a 22-bed medical intensive care unit of a public hospital. The hospital did not have an intermediate care or respiratory unit. Patients that needed close monitoring were admitted to this unit. The unit's average daily census was 15 patients, and its average number of patients receiving invasive mechanical ventilation was 4. The Institutional Review Board approved the study and waived the requirement of written consent.

2.2. Subjects

All patients in the intensive care unit were screened daily by the study coordinator who was a hospital employee with background in health care. Patients were selected if they were receiving respiratory support with oxygen therapy by oxygen mask or by noninvasive ventilation, if they were intubated for mechanical ventilation without sedation, or if they were breathing spontaneously during a weaning trial. Patients were excluded if they were not on oxygen therapy or if they were on sedation. One of the investigators (AP) confirmed the accuracy of selection.

2.3. Physicians

Four critical care consultants and 4 critical care fellows participated in this study. The consultants had 1-28 years of experience with an average of 12 years. The fellows finished 3 years of training in internal medicine and were in the second or third years of critical care training. A fellow and a consultant were summoned to examine the selected subject. They were not involved in the care of any of the subjects and were blinded to each other's responses. Paired examinations (consultant-fellow) were included in the analysis only when they occurred less than 15 minutes apart.

2.4. Assessment of subjects

We developed a standardized form to guide fellows and consultants through 3 steps of assessing each subject: step 1, subjectively rate the level of respiratory distress; step 2, examine for 11 signs of increased breathing effort from the face, then neck, then chest, and then abdomen [3,18]; and step 3, measure the vital signs.

2.5. Subjective rating of respiratory distress

The physicians (fellows and consultants) were instructed that their rating of distress should reflect the effort of breathing and that they should assess it in a manner similar to their usual practice. Each of them observed each subject appearance, without touching, for 10 seconds and then rated the level of distress on a 4-level scale: absent, slight, moderate, and severe. For the analysis, these levels were coded 1, 2, 3, and 4, respectively. The monitor screens in the subject rooms were turned off during the assessment.

2.6. Signs of increased breathing effort

The physicians examined each subject in a specified sequence (eTable 1) [3,18]. They observed each sign for 5 breath cycles and then rated it as definitely absent, probably absent, probably present, or definitely present. These ratings were coded 1, 2, 3, and 4, respectively.

2.7. Vital signs

The physicians counted breaths over 1 or 2 minutes. The study coordinator collected blood pressure, heart rate, fraction of inspired oxygen (F_{IO_2}), and oxygen saturation by pulse oximetry (S_{PO_2}) from the monitoring devices immediately after the assessment. The S_{PO_2} to F_{IO_2} ratio was calculated to assess oxygenation [19].

2.8. Analysis

We reported continuous variables as means ± standard deviation and compared groups using the *t* test or 1-way analysis of variance. For correlations, we calculated Spearman coefficient. Statistical significance was present when the *P* was ≤ .05, and adjustment was made for multiple tests.

To validate the rating of distress, we determined the correlation between it and respiratory rate, S_{PO_2} to F_{IO_2} ratio, and the signs of increased breathing effort assessed by the same physician and by the second physician (to correct for observer bias). We also calculated its sensitivity, specificity, and likelihood ratio for predicting the presence of tachypnea (respirator rate ≥29), hypoxia (S_{PO_2} to F_{IO_2} ratio <2.5), and each of the signs of increased breathing effort. We considered a subjects to have serious acute respiratory dysfunction when they had any 3 or more of the 13 signs (hypoxia, tachypnea, 11 increased effort). To validate the signs of increased effort, we compared the rating between the subjects divided into 5 groups based on the level of respiratory support (mechanical ventilation, weaning, spontaneous breathing, NIV (non-invasive ventilation), and just before intubation). This analysis was performed with Kruskal-Wallis test.

We used the paired measurements to assess agreement on the signs of increased work of breathing (distress and effort). We calculated the intraclass correlation coefficient (ICC) using a random 1-way absolute mode. An ICC > 0.7 indicates adequate agreement. We also calculated κ using linear weights. Agreement was considered poor when κ is 0, slight when κ is 0 to 0.2, fair when κ is 0.2 to 0.4, moderate when κ is 0.4 to 0.6, substantial when κ is 0.6 to 0.8, and almost perfect when κ is 0.8 to 1.

3. Results

3.1. Subjects and assessments

The age distribution of the subjects was wide (19-85 years), and there was adequate representation of both sexes. The major diagnoses cover a wide range of systems. At the time of assessments, most subjects