Diagnostic Accuracy of Respiratory Distress Observation Scales as Surrogates of Dyspnea Self-report in Intensive Care Unit Patients

Romain Persichini, M.D., Frédéric Gay, M.D., Ph.D., Matthieu Schmidt, M.D., Julien Mayaux, M.D., Alexandre Demoule, M.D., Ph.D., Capucine Morélot-Panzini, M.D., Ph.D., Thomas Similowski, M.D., Ph.D.

ABSTRACT

Background: Dyspnea, like pain, can cause major suffering in intensive care unit (ICU) patients. Its evaluation relies on self-report; hence, the risk of being overlooked when verbal communication is impaired. Observation scales incorporating respiratory and behavioral signs (respiratory distress observation scales [RDOS]) can provide surrogates of dyspnea self-report in similar clinical contexts (palliative care).

Methods: The authors prospectively studied (single center, 16-bed ICU, large university hospital) 220 communicating ICU patients (derivation cohort, 120 patients; separate validation cohort, 100 patients). Dyspnea was assessed by dyspnea visual analog scale (D-VAS) and RDOS calculated from its eight components (heart rate, respiratory rate, nonpurposeful movements, neck muscle use during inspiration, abdominal paradox, end-expiratory grunting, nasal flaring, and facial expression of fear). An iterative principal component analysis and partial least square regression process aimed at identifying an optimizing D-VAS correlate (intensive care RDOS [IC-RDOS]).

Results: In the derivation cohort, RDOS significantly correlated with D-VAS ($r = 0.43$; 95% CI, 0.29 to 0.58). A five-item IC-RDOS (heart rate, neck muscle use during inspiration, abdominal paradox, facial expression of fear, and supplemental oxygen) significantly better correlated with D-VAS ($r = 0.61$; 95% CI, 0.50 to 0.72). The median area under the receiver operating curve of IC-RDOS to predict D-VAS was 0.83 (interquartile range, 0.81 to 0.84). An IC-RDOS of 2.4 predicted D-VAS of 4 or greater with equal sensitivity and specificity (72%); an IC-RDOS of 6.3 predicted D-VAS of 4 or greater with 100% specificity. Similar results were found in the validation cohort.

Conclusions: Combinations of observable signs correlate with dyspnea in communicating ICU patients. Future studies in noncommunicating patients will be needed to determine the responsiveness to therapeutic interventions and clinical usefulness.

(Anesthesiology 2015; 123:830-7)

This article is featured in “This Month in Anesthesiology,” page 1A. Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are available in both the HTML and PDF versions of this article. Links to the digital files are provided in the HTML text of this article on the Journal’s Web site (www.anesthesiology.org). This study has been presented at the Dyspnea 2015 Symposium, University of New England, Biddeford, Maine, June 12, 2013; at the 41st Annual Congress, Société de Réanimation de Langue Française, La Défense, France, January 17, 2013; and at the American Thoracic Society 2014 International Conference, San Diego, California, May 18, 2014.


YSPNEA is “a subjective experience of breathing discomfort...”' Like pain, dyspnea involves physical suffering and emotional distress, it has neurovegetative and behavioral repercussions, and can be intrinsically prognostic. Intensive care unit (ICU) patients are at high risk of dyspnea at various stages of their stay. In a study of mechanically ventilated patients, dyspnea was frequent (47%), marked (median dyspnea visual analog scale [D-VAS] of 5), often treatment related (more frequent under assist-control ventilation; often related to ventilator settings), strongly associated with anxiety, and negatively associated with weaning outcome.

Dyspnea—suffering associated with breathing—is a major but underappreciated symptom in acutely ill patients. There is no validated, objective approach to assess dyspnea in critically ill patients who cannot easily communicate.
As with pain, addressing dyspnea in ICU patients appears highly clinically relevant. This requires focused awareness from caregivers and patient cooperation. Indeed, because dyspnea involves the sensory identification of afferent signals by the brain and their cognitive and affective processing, its characterization depends on self-report. Clinical signs of “respiratory distress” and self-perceived dyspnea can be disconnected, setting a limitation to identifying dyspnea in many ICU patients whose ability to communicate verbally is impaired. Nevertheless, a link does exist between dyspnea and certain observable signs. This is all the more true if clinical observation goes beyond physical respiratory signs and extends to signs of emotional distress (e.g., facial expression of fear).

A respiratory distress observation scale (RDOS) has been validated as a surrogate for self-reported dyspnea in palliative care (see table 1, Supplemental Digital Content 1, http://links.lww.com/ALN/B183, which provides the eight weighted components of RDOS). RDOS conceptually derives from observational pain scales, some of which have been validated in mechanically ventilated patients. In patients able to communicate with their caregivers and to use a visual analog scale (VAS) (“communicating” patients), RDOS has good internal consistency, good convergent validity (correlates with D-VAS), and good discriminant validity (not perturbed by pain). In patients who are not able to communicate with their caregivers and are unable to use a VAS (“noncommunicating” patients), for example at the end of life, RDOS relates to clinical outcomes and is sensitive to therapeutic interventions. RDOS is therefore likely to be of value to assess dyspnea in the ICU.

We set out to test how RDOS relates to D-VAS in an unselected ICU population. We hypothesized that RDOS and D-VAS would relate to one another in patients able to verbally communicate. We also hypothesized that RDOS would have to be modified for optimal performance in a population different from its original derivation setting. We tested this by statistically deriving an adapted “intensive care RDOS” [IC-RDOS] in one cohort of patients (derivation data) and validating it in another cohort that was constituted reasonably large (each of our cohorts is larger than the largest ICU population examined for dyspnea in the literature).

### Collection of Derivation Data
Clinical data were gathered during the first 24 h of the stay (on week days only), between 8 and 10 AM, by the physician in charge of the patient. Separate data collection was performed by the nurse in charge. Investigators also assessed, at the same time:

- presence of dyspnea (“do you breathe well,” “is your breathing comfortable,” “are you bothered by your breathing,” “is your breathing difficult”): at least two different questions were asked and the answers had to be concordant;
- intensity of dyspnea (10-cm D-VAS, from “no respiratory discomfort” to “intolerable respiratory discomfort”); of note, a D-VAS of 4 or greater was considered of particular interest, first because this corresponds to the lower quartile of dyspnea ratings in the largest published study of dyspnea in intensive care patients and second by analogy with pain where a VAS of 4 is the lower threshold for “moderate pain” requiring combined analgesic prescription.

Six physicians and 11 nurses participated to data collection. They all had received specific information and training on the use of D-VAS and RDOS.

### Statistical Analysis
The statistical analysis was conducted by a certified biostatistician (F.G.) with the use of XLSTAT v2014 (Addinsoft, France), SPAD 8.0 (Cohere Analytics, France), and StatXact-10 (Cytel, USA).

#### General Statistics. Quantitative variables are summarized as median and interquartile range (IQR) and qualitative variables as frequency. D-VAS versus RDOS and D-VAS versus IC-RDOS were tested using Spearman correlation. RDOS and IC-RDOS receiver operating characteristic (ROC) curves were generated for each exact value of D-VAS from 1 to 10. Univariate comparisons were conducted between dyspneic and nondyspneic patients with the use of Fisher exact test for binomial variables and Mann–Whitney U test for quantitative variables. A probability  value of type I error less than 0.05 was considered significant, with Benjamini–Hochberg correction for multiplicity when necessary.

#### IC-RDOS Derivation Procedure. To identify the best correlation between a set of variables and D-VAS, we used iterative
principal component analysis and partial least square (PLS) regression.\textsuperscript{22–24} This approach was chosen because of parsimony in terms of hypotheses, robustness to missing data, and ability to directly model the data with successive regressions (see the detailed description of the IC-RDOS construction process, Supplemental Digital Content 1, http://links.lww.com/ALN/B183). In brief, PLS iterations were performed by stepwise elimination of the weakest explanatory variables to, in the end, propose the linear combination of explanatory variables that was best adjusted to D-VAS.

**Validation Data.** Using the solution estimated in the derivation cohort, IC-RDOS was calculated in a validation cohort of 100 communicating patients who were recruited according to the same criteria as the derivation patients but during a separate period (July 22, 2013 to October 29, 2013) and in whom the data collection was performed by a single investigator (R.P.). As in the derivation cohort, D-VAS versus RDOS and D-VAS versus IC-RDOS were assessed using Spearman correlation, and the strength of these correlations was compared using Fischer Z test. RDOS and IC-RDOS ROC curves were similarly generated.

**Physicians–Nurses Interobserver Agreement.** Interobserver agreement between physicians and nurses was assessed in the derivation cohort, from a subset of patients in whom the data had been gathered by the two categories of observers within a 2-h delay during which no change in ventilatory support or oxygenotherapy had occurred (n = 87). Concordance was assessed using Kendall W for RDOS and IC-RDOS and their quantitative components, whereas Cohen k was used for binomial components. Of note, the “nurses” data were not taken into account to establish the D-VAS versus RDOS and D-VAS versus IC-RDOS relations.

**Results**

**Derivation Data**

Over 4.5 months, 456 patients were admitted and 193 evaluated (fig. 1). Seventy-three (37%) were noncommunicating (sedation, n = 49; confusion, n = 9; not understanding, n = 6; and other, n = 9): RDOS data were gathered in these patients but by nature were not used to study D-VAS versus RDOS relations. Among the 120 remaining communicating patients, 69 (57%) reported dyspnea (and therefore did not leave the D-VAS cursor on the “no respiratory discomfort boundary”; median D-VAS 4.5 [3.2 to 6.0]; table 1). Dyspneic patients were more likely to receive supplemental oxygen (P < 0.0001) and to report anxiety and pain (P < 0.0001) (table 1).

There were no missing D-VAS data. Among items constitutive of RDOS, heart rate was missing in one case and respiratory rate in another case. D-VAS and RDOS were significantly correlated (r = 0.43; 95% CI, 0.29 to 0.58; P < 0.0001), with RDOS explaining 18.8% of the variance of D-VAS (fig. 2A).

**IC-RDOS Derivation and Comparison with RDOS**

The iterative PLS procedure identified five variables, of which the combination and weighting (namely IC-RDOS) allowed optimal correlation with D-VAS (r = 0.61; 95% CI, 0.50 to 0.72; P < 0.0001) (table 2), with IC-RDOS explaining 37.3% of the variance of D-VAS (fig. 2B). The D-VAS versus IC-RDOS correlation was significantly better than the D-VAS versus RDOS one (P = 0.04).

The median area under the ROC curve was 0.74 (IQR, 0.72 to 0.76) for RDOS and 0.83 (IQR, 0.81 to 0.84) for IC-RDOS (P = 0.0003) (table 3).

![Study flow chart (derivation cohort). D-VAS = dyspnea visual analog scale.](image-url)
This study shows that observation scales incorporating respiratory and behavioral clinical signs can relate with self-evaluated dyspnea in unselected patients admitted to a medical ICU. This is reminiscent of observations made in other categories of dyspneic patients and consistent with pain data (see the review by Gélinas et al.25). How these scales can help identify noncommunicating ICU patients with “silent dyspnea” (unrecognized suffering) remains to be determined.

**Discussion**

This study shows that observation scales incorporating respiratory and behavioral clinical signs can relate with self-evaluated dyspnea in unselected patients admitted to a medical ICU. This is reminiscent of observations made in other categories of dyspneic patients and consistent with pain data (see the review by Gélinas et al.25). How these scales can help identify noncommunicating ICU patients with “silent dyspnea” (unrecognized suffering) remains to be determined.

**Study Limitations**

The study was conducted at a single site, with an inherent risk of data overfitting. It pertained to medical patients only.
The patients were evaluated only once (no indication about responsiveness). They were evaluated during the first day of their stay (hence a low proportion of communicating patients under mechanical ventilation; table 1). Patients were included in the derivation study on a consecutive basis, but this was done on week days only and with a “real-life” concern: we insisted on having the physician in charge of the patient collecting the data rather than a dedicated investigator. This resulted in a number of patients being missed because each physician was in charge of overseeing more than one patient. This could have biased the results toward a less severe population. We do not think that this occurred because the 193 communicant and noncommunicant patients in whom RDOS data were gathered (fig. 1) had severity scores identical to those of the admitted population (e.g., median simplified acute physiology score II of 42 and 43, respectively). For all the above reasons, it is not possible to claim generalized validity for RDOS or IC-RDOS: other study designs in other contexts would probably lead to different numerical results. This should, however, not invalidate the notion that the observation scales can be useful to assess dyspnea in the ICU.

D-VAS was chosen as the reference index in this study because it is, to the best of our knowledge, the only psychometric tool that has been used to evaluate dyspnea in ICU patients in the literature. Visual analog scales are widely used for the measurement of symptoms, and their validity to assess dyspnea has been established by several studies in various contexts. Visual analog scales are considered most suited as a “within-subject” tool and less so to compare patients, even though the comparison between a dyspnea VAS and a verbal rating scale has shown quasiredundancy in cancer patients. In the ICU, VAS to measure dyspnea has not been the object of a specific validation, but it has been shown that it was feasible in this context, well correlated with anxiety (which is clinically and physiologically expected), and responsive to therapeutic intervention. Of interest, the mere fact that we did find a relation between D-VAS (purely psychometric) and RDOS/IC-RDOS (purely physical) not only support the putative interest of observation scales to evaluate respiratory suffering but also, reciprocally, lends support to some “between-patient” validity of D-VAS.

Finally on this, dyspnea is multidimensional and imperfectly apprehended by unidimensional tools such as D-VAS, RDOS and IC-RDOS, because they incorporate “facial expression of fear,” which probably should make it sufficient to bring dyspnea to the mind of ICU caregivers. The interest of a composite clinical score such as RDOS or IC-RDOS is underlined by the fact that, in a study of dyspnea in mechanically

---

**Table 2. Calculation of the Modified Intensive Care Respiratory Distress Observation Scale**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>0—Heart rate (beats/min)</td>
<td>3.3</td>
</tr>
<tr>
<td>1—Use of neck muscle during inspiration</td>
<td>+ (heart rate)/65</td>
</tr>
<tr>
<td>If present</td>
<td>+1</td>
</tr>
<tr>
<td>If absent</td>
<td>−1</td>
</tr>
<tr>
<td>2—Abdominal paradox during inspiration</td>
<td>+1</td>
</tr>
<tr>
<td>If present</td>
<td>+1</td>
</tr>
<tr>
<td>If absent</td>
<td>−1</td>
</tr>
<tr>
<td>4—Facial expression of fear</td>
<td>+1</td>
</tr>
<tr>
<td>If present</td>
<td>+1</td>
</tr>
<tr>
<td>If absent</td>
<td>−1</td>
</tr>
<tr>
<td>5—Supplemental oxygen</td>
<td>+0.7</td>
</tr>
<tr>
<td>If present</td>
<td>+0.7</td>
</tr>
<tr>
<td>If absent</td>
<td>−0.7</td>
</tr>
</tbody>
</table>

---

Fig. 2. Spearman rank correlation plot between dyspnea visual analog scale (D-VAS) and the respiratory distress observation scale (RDOS, A) or the modified intensive care RDOS (IC-RDOS, B) in the 120 patients from the derivation cohort who were able to communicate verbally with their physician.
In our study, RDOS and D-VAS were significantly correlated ($r = 0.43$), with a strength similar to that described in pulmonary rehabilitation patients (the correlation coefficient was $r = 0.39$ in the original study by Campbell$^{11}$). We had hypothesized that ICU specificities could interfere with the D-VAS/RDOS relation. For example, tachypnea in the ICU could proceed from metabolic acidosis, anemia, or pain, which stimulate ventilatory drive but might not be intrinsically dyspneogenic. Accordingly, the statistical process that we applied to the data set produced an RDOS variant with a significantly stronger D-VAS correlation ($r = 0.61$) and improved operative characteristics (fig. 3; fig. 1, Supplemental Digital Content 1, http://links.lww.com/ALN/B183). IC-RDOS and RDOS differ by the number of items (5 vs. 8), the nature of the retained items (respiratory rate, restlessness, nasal flaring, expiratory grunting not in IC-RDOS, and supplemental oxygen not in RDOS), and item weighting. Reassuringly, however, IC-RDOS retains major RDOS items (heart rate—strongly associated with dyspnea in mechanically ventilated patients$^6$—use of neck muscle during inspiration, abdominal paradox, and facial expression of fear). “Supplemental oxygen” in IC-RDOS is consistent with the association between RDOS and oxygenotherapy noted in palliative care.$^{12}$ The very IC-RDOS derivation method makes external validation particularly important: reassuringly, IC-RDOS behavior was well confirmed in our validation cohort (fig. 3). The IC-RDOS interobserver agreement was slightly better than that of RDOS, possibly because of the lesser number of items. Of note, agreement was markedly heterogeneous among items (table 2, Supplemental Digital Content 1, http://links.lww.com/ALN/B183, see above): this suggests that the efforts to improve the identification of abdominal paradox and facial expression of fear (the two items with the lowest agreement) by caregivers could further improve the usefulness of RDOS and IC-RDOS. IC-RDOS could be considered slightly less easy to use than RDOS because of the treatment of heart rate as a continuous variable: an online calculator and a downloadable application are available to resolve this.$^{29}$ Importantly, RDOS, IC-RDOS, and all similar scales require appropriate user training to ensure reliability.

**Relevance**

As for pain, self-assessment is the reference method of dyspnea evaluation. The interest of surrogate markers is therefore questionable in patients able to communicate with their caregivers. However, repeating verbal dyspnea evaluation as often as

---

**Table 3. Areas Under the ROC Curves for the RDOS and the Modified IC-RDOS According to the D-VAS Thresholds**

<table>
<thead>
<tr>
<th>D-VAS Threshold</th>
<th>≥1</th>
<th>≥2</th>
<th>≥3</th>
<th>≥4</th>
<th>≥5</th>
<th>≥6</th>
<th>≥7</th>
<th>≥8</th>
<th>≥9</th>
<th>Median</th>
<th>IQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>RDOS ROC curve AUC</td>
<td>0.69</td>
<td>0.73</td>
<td>0.74</td>
<td>0.75</td>
<td>0.78</td>
<td>0.75</td>
<td>0.68</td>
<td>0.73</td>
<td>0.79</td>
<td>0.74</td>
<td>0.72–0.76</td>
</tr>
<tr>
<td>IC-RDOS ROC curve AUC</td>
<td>0.78</td>
<td>0.82</td>
<td>0.82</td>
<td>0.83</td>
<td>0.83</td>
<td>0.83</td>
<td>0.79</td>
<td>0.84</td>
<td>0.85</td>
<td>0.83*</td>
<td>0.81–0.84</td>
</tr>
</tbody>
</table>

AUC = area under the curve; D-VAS = dyspnea visual analog scale; IC-RDOS = intensive care respiratory distress observation scale; IQR = interquartile range; RDOS = respiratory distress observation scale; ROC = receiver operating characteristic.

* $P = 0.0003$ (Mann–Whitney U test).

---

**Fig. 3.** Areas under the receiver operating characteristic curves (AUC) for the modified intensive care respiratory distress observation scale (IC-RDOS) as a predictor of a dyspnea visual analog scale (D-VAS) of 4 or greater in the initial derivation cohort (A, 120 patients) and in the subsequent validation cohort (B, 100 patients).
desirable can be difficult, mostly in clinically labile situations. Routine monitoring a dyspnea-related observation scale could then be useful to trigger targeted verbal interactions.

The true targets of surrogate markers of dyspnea are patients having difficulties expressing themselves for whatever reason (e.g., sedation, coma, cognitive impairment or intubation, among others). Indeed, being unable to report a symptom does not mean not suffering from this symptom or not being at risk of ulterior consequences. There is, by definition, no means to validate a surrogate dyspnea marker against self-reported dyspnea in such patients. This will have to be tested by future ICU studies assessing the responsiveness of IC-RDOS to interventions known to alleviate dyspnea (e.g., adjusting ventilator settings, see the study by Schmidt et al8) and evaluating the corresponding clinical benefits.

Conclusions
Our findings validate RDOS and its derivative IC-RDOS as potential surrogates of dyspnea in the ICU, proving the concept that observation scales can be useful in this context as they are in other contexts. Clinical usefulness in noncommunicating ICU patients will have to be demonstrated, as it has been in the case of palliative care. Until then, we submit that IC-RDOS could readily be used in the ICU, either to prompt a dyspnea-oriented verbal interaction if possible or to launch a “risk of dyspnea” checklist if impossible. It might indeed be better to correct putative dyspugenic factors without knowing for certain that this correction actually alleviates dyspnea than to leave the patient exposed to respiratory-related suffering and particularly so if this correction does not carry a significant iatrogenic risk.

Acknowledgments
The authors gratefully thank Adrien Kerckaert, Pascaleine Dejaune, Christian De Tymowski, Nöelle Messaoudi, and Sophie Moniez (all medical students or residents, Service de Pneumologie et Réanimation Médicale [Département “R3”], Paris, France), for their help in collecting the data. The authors also thank Stéphane Tortajada (Paris, France), for devising the Internet calculator and the smartphone application, and Paul Robinson (Kettering, United Kingdom), for help with English style and grammar.

This study was supported by “Association pour le Développement et l’Organisation de la Recherche en Pneumologie et sur le Sommeil,” Paris, France, and by the program “Investissement d’Avenir ANR-10-AIHU 06” of the French Government.

Competing Interests
The authors declare no competing interests.

Correspondence
Address correspondence to Dr. Similowski: Department of Respiratory and Critical Care Medicine, Pitié-Salpêtrière Hospital, 47-83 Boulevard de l’Hôpital, 75651 Paris Cedex 13, France. thomas.similowski@pslaphp.fr. Information on purchasing reprints may be found at www.anesthesiology.org or on the masthead page at the beginning of this issue. ANESTHESIOLOGY’s articles are made freely accessible to all readers, for personal use only, 6 months from the cover date of the issue.

References
complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. BMJ 2003; 326:41–4

ANESTHESIOLOGY REFLECTIONS FROM THE WOOD LIBRARY-MUSEUM

Fanny Davenport and the Doctors Rugg: 25-cent “Extracting, with Gas”

In August of 1879, the Doctors Rugg registered as dentists practicing at 777 Broadway in Albany, the capital city of the State of New York. In less than a decade the Ruggs were advertising (right) that their Albany City Dental Association (ACDA) was the “Largest Private Dental Establishment in the World.” Part of their successful advertising included distributing trade cards featuring Fanny Lily Gipsy Davenport (1850–1898). London-born and Boston-educated, Fanny Davenport (left) was a celebrated thespian for the final 36 of her abbreviated lifetime of only 48 years. Just 2 years her junior, the younger dentist, Dr. Newton P. Rugg (1852–1907), worked as ACDA’s secretary. His uncle and managing dental partner was Dr. Datus E. Rugg (1819–1898), a former blacksmith. ACDA’s patients were likely relieved to learn that the managing partner had traded in his forging hammer for nitrous oxide—and that “Extracting, with Gas” cost only 25 cents. (Copyright © the American Society of Anesthesiologists, Inc.)

George S. Bause, M.D., M.P.H., Honorary Curator, ASA’s Wood Library-Museum of Anesthesiology, Schaumburg, Illinois, and Clinical Associate Professor, Case Western Reserve University, Cleveland, Ohio. UJYC@aol.com.