ORIGINAL RESEARCH

Dyspnea could be accurately assessed by a caregiver in hospitalized patients with respiratory diseases: Interrater reliability and agreement study

Gregory Reychler[a,b,c,*], Marc Beaumont[d,e], Anne-Claire Latiers[c], Thierry Pieters[a,b], Antoine Fremault[f]

[a] *Institut de Recherche Expérimentale et Clinique (IREC), Pôle de Pneumologie, ORL and Dermatologie, Université Catholique de Louvain, Brussels, Belgium*
[b] *Service de Pneumologie, Cliniques Universitaires Saint-Luc, Avenue Hippocrate 10, Brussels 1200, Belgium*
[c] *Secteur de Kinésithérapie et Ergothérapie, Cliniques Universitaires Saint-Luc, Brussels, Belgium*
[d] *Service de Réhabilitation Respiratoire, Centre Hospitalier des Pays de Morlaix, Morlaix, France*
[e] *EA3878 (GETBO), CHU Brest, Brest, France*
[f] *Service de Pneumologie, Grand Hôpital de Charleroi, Charleroi, Belgium*

Abstract

Background: The perception of dyspnea is a subjective feeling typically self-assessed by the patient. However, the assessment by a caregiver is sometimes required.

Objectives: The primary aim was to compare patient self-assessment and caregiver assessment of dyspnea (interrater reliability) using the modified Borg and visual analog scale (VAS) in hospitalized patients. The secondary aim was to compare dyspnea assessment between the two scales for patients and caregiver (inter-instrument reliability).

Methods: Self-assessment of dyspnea intensity of hospitalized patients with respiratory diseases was compared with caregiver's assessment. Dyspnea intensity was measured using two scales, the modified Borg scale (0−10 scale) and the 10 cm VAS. Mean difference and 95% confidence interval (CI) between assessors (i.e. patient versus caregiver) were calculated for each scale. Inter- and intra-rater reliability was calculated using intraclass correlation coefficients (ICCs).

Results: A total of 254 patients were recruited. The mean differences between patient and caregiver ratings were 0.31 (95% CI: 0.09, 0.53) for the modified Borg scale and 0.36 (95% CI: 0.06, 0.65) for the VAS scale. Interrater reliability was good for both scales with ICC of 0.79 (95% CI: 0.73, 0.84) for VAS and 0.82 (95% CI: 0.77, 0.86) for the modified Borg scale. The mean differences in scores between scales were 0.93 (95% CI 0.69, 1.17) for patients' ratings and 0.88 (95% CI 0.72, 1.04) for caregiver's rating. The inter-instrument reliability was moderate to good and similar for both assessors.

* Corresponding author at: Service de Pneumologie, Cliniques Universitaires Saint-Luc, Avenue Hippocrate 10, Brussels 1200, Belgium.
  E-mail: gregory.reychler@uclouvain.be (G. Reychler).

Conclusion: Dyspnea can be accurately estimated by caregivers when patients with lung diseases cannot self-report. Scores on the VAS to rate dyspnea were higher than the scores on the Borg scale.

Introduction

Dyspnea is defined as "a subjective experience of breathing discomfort that consists of qualitatively distinct sensations."[1] It is a common and complex symptom frequently assessed in clinical practice. Assessment of dyspnea is relevant to patients admitted to the emergency room or the hospital,[2] community-residing individuals,[3] or patients consulting their general practitioner.[4] Due to the high prevalence of dyspnea in a variety of respiratory diseases, such as lung cancer, chronic obstructive pulmonary disease (COPD), and interstitial lung disease, it is considered an important patient-reported outcome in clinical and research settings.[5]

The mechanisms of dyspnea and contributing factors have been investigated for a long time. Nevertheless, knowledge about the neurophysiology and psychology of dyspnea increased considerably in the last decade. Psychological and cultural factors may play an important role in the perception of shortness of breath.[6] Accordingly, the assessment of dyspnea has improved by notably taking into account different domains.[7] But, despite new tools for assessing dyspnea, that specifically evaluate sensory and affective domains, having recently emerged,[8] the visual analog scale (VAS) and modified Borg scales are still largely used.[9] Both of these tools have already shown to be sensitive enough to measure these two dyspnea domains.[8] However, even if they have a similar scoring scale and the tools are conceptually similar, the level of association between these two measures has not been studied.

While the perception of dyspnea is subjective and, therefore, the patient's self-assessment of dyspnea is the preferred assessment method,[6] the assessment by a caregiver is sometimes required. This occurs when the patient is performing an exercise or for those who are unable to communicate their feelings. Fifty-four percent of patients referred for inpatient palliative care consultation are unable to assess their own level of dyspnea.[10] To facilitate the assessment of dyspnea in these patients, the Respiratory Distress Observation Scale has been validated in terminally ill patients.[11] However, there is a lack of specific tools for noncommunicative patients such as those in coma or diagnosed with cerebral palsy.

Pain is a subjective symptom which may share some similarities with dyspnea. Indeed, pain and dyspnea are both associated to chronic diseases, unpleasantness, decline in quality of life, and physical sensations mediated by a common cortical network. Patients' and physicians' assessment of pain intensity has showed poor agreement.[12] Similar results were found with measurements of other symptoms, especially for those considered more subjective.[13,14] Nevertheless, the agreement between caregiver and patient assessment of dyspnea intensity has been poorly studied. The dyspnea duration and its variability over time could influence this agreement and these elements are partly related to the types of respiratory disease. Moreover, the feelings of emotional distress could differently influence this assessment either by the caregiver or by the patient depending on the disease and its prognosis.

The primary objective of the current study was to compare the dyspnea assessment between patients and caregivers (interrater reliability) using both the modified Borg scale and a VAS in hospitalized patients with respiratory diseases. The secondary objective was to compare dyspnea assessment scores between the two scales for the patients and caregivers (inter-instrument reliability).

Methods

Study design

This was an observational cross-sectional study which follows the Guidelines for Reporting Reliability and Agreement Studies.[15]

Participants

We prospectively recruited all participants who were hospitalized between July and September 2013 in the pulmonology and oncology units of the Cliniques Universitaires Saint-Luc (Brussels, Belgium) and the Grand Hôpital de Charleroi (Charleroi, Belgium). To be included, participants had to be aged 18 years or over and diagnosed with an acute or chronic respiratory disease with or without oxygen therapy. Patients were excluded if presenting with mental or cognitive impairments or unable to understand French and if requiring non-invasive ventilation.

All included patients signed a written informed consent form according to the Declaration of Helsinki and the current guidelines for Clinical Good Practice. The study was approved by the Institutional Medical Ethics Committee (Cliniques Universitaires Saint-Luc, Brussels, Belgium - B403201317644) and was registered at ClinicalTrials.gov (NCT03598569).

Procedure

All participants were assessed during hospitalization by the same qualified caregiver. The caregiver was a physical therapist who worked in the pulmonology unit and in the rehabilitation center on a regular basis for the past 5 years and was experienced in assessing dyspnea intensity. The assessment of dyspnea by the caregiver was made through visual observation and was primarily based on breathing frequency and the recruitment of accessory muscles. This assessment was performed as a single measure just before the patient's self-assessment of dyspnea. The assessment of dyspnea intensity by the patient and the caregiver was performed while the

patient was at rest. Patients were unaware of the time the caregiver's assessment took place and they remained blinded to the caregiver's rating throughout the study.

Instruments

The 10 cm VAS[16] and the modified (0−10) Borg scale[17] were used to assess the sensory domain of dyspnea. Both scales were previously validated for this purpose.[18]

Visual analog scale (VAS)

The VAS is a 10 cm line anchored at the ends by statements that define dyspnea, ranging from "I am not at all short of breath" (0 cm) to "The most short of breath I have ever been" (10 cm). Participants were asked to place a vertical mark on the line to indicate how they rate their dyspnea. The distance between the start of the line and the mark placed by the participant was then measured in centimeters.

Modified Borg scale

The modified Borg scale is an 11-point numerical scale with associated verbal descriptors rating the difficulty of breathing from "Your breathing is causing you no difficulty at all" (zero point) to "Your breathing difficulty is maximal" (10 points). The patients were asked to choose the number that best describeg their feeling about breathing difficulties.

Data analysis

The SPSS 25.0 program (IBM software) was used to analyze the data. Descriptive data were expressed as mean ± SD for normally distributed variables or median (min-max) for non-normally distributed variables. Normality of the distribution was investigated by the Kolmogorov-Smirnov test. The comparison between patient's self-assessment and caregiver's assessment of dyspnea using both scales in hospitalized patients with respiratory diseases was performed by calculating the mean difference and 95% confidence interval (CI) between raters. The degree of concordance between the two assessors (interrater reliability for each scale) was calculated using intraclass correlation coefficients (ICCs) using a two-way mixed effects (single measurement) mode with consistency.[19−22]

For the comparison of dyspnea assessment between the two scales, for patients and caregiver, we calculated the mean difference in scores and 95% CI between the two scales. The degree of concordance between both scales (inter-instrument reliability) for both the patients and the caregiver was also calculated using a two-way mixed effects model (single measurement) with absolute agreement ICC type.[22] All ICCs were expressed with their 95% CI. Values less than 0.5 indicate poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.9 indicate good reliability, and values greater than 0.90 indicate excellent reliability.[22] In a secondary analysis, the above was then conducted separately for the subgroup of patients with lung cancer and those with other lung diseases.

Results

Two hundred and seventy one patients were considered eligible for the study. Sixteen patients were excluded because they did not understand French and one patient was excluded due to cognitive impairments. The characteristics and diagnosis of the included patients are shown in Table 1.

Patient versus caregiver

For the total sample, the mean differences between patient and caregiver ratings were 0.31 (95% CI: 0.09, 0.53) for the modified Borg scale and 0.36 (95% CI: 0.06, 0.65) for the VAS scale (Table 2), which indicates higher scores for patient self-assessment compared to caregiver assessment. We found good reliability for both scales with ICC of 0.79 (95% CI: 0.73, 0.84) for the VAS and 0.82 (95% CI: 0.77, 0.86) for the modified Borg scale (Table 3), which indicates good agreement between patients and caregiver for both scales. Our subgroup analyses indicated similar results for the group of patients with lung cancer. In contrast, a significant difference in rating was found between the caregiver and the patients' self-rating for the group with other lung diseases (Table 2). Finally, there was a higher degree of concordance between patients and caregiver in the lung cancer group

---

Table 1 Patients characteristics.

	All	Lung cancer group	Other lung disease group
N	254	121	133
Age	66.5 ± 11.9	64.9 ± 9.9	68.0 ± 13.4
Gender (F/M)		76/45	61 /72
Diagnosis			
COPD			85
Fibrosis			8
Asthma			3
Pneumonia			7
Lung transplant			3
Bronchiectasis			3
Other			24
Oxygen			
n	74	21	53
Flow (L/min)	0.8 ± 1.4	0.5 ± 1.4	1.0 ± 1.5

Values are expressed as mean ± standard deviation or number of patients.

Table 2 Assessment of dyspnea using the modified Borg scale and the visual analogue scale.

Group/instrument	Caregiver	Patient	Mean Difference (95%CI)*
Total sample (n = 254)			
Borg	2.4 ± 2.2	2.7 ± 2.3	0.31 (0.09, 0.53)
VAS	3.2 ± 2.9	3.6 ± 2.8	0.36 (0.06, 0.65)
Mean Difference (95%CI)**	0.88 (0.72, 1.04)	0.93 (0.69, 1.17)	
Lung cancer group (n = 121)			
Borg	2.0 ± 2.3	1.9 ± 2.3	−0.03 (−0.23, 0.17)
VAS	2.2 ± 2.7	2.2 ± 2.5	−0.02 (−0.29, 0.25)
Mean Difference (95%CI)**	0.25 (0.10, 0.40)	0.25 (0.03, 0.47)	
Other lung diseases group (n = 133)			
Borg	2.7 ± 2.0	3.3 ± 2.2	0.62 (0.24, 0.99)
VAS	4.2 ± 2.7	4.9 ± 2.5	0.70 (0.20, 1.20)
Mean Difference (95%CI)**	1.46 (1.22, 1.70)	1.54 (1.15, 1.93)	

VAS: visual analogue scale.
Values are expressed as mean ± standard deviation or mean difference (95% confidence interval). *Between caregiver and patient's self-rating. **Between Borg and VAS scales.

(ICC > 0.90) compared to those with other lung diseases (ICC between 0.50 and 0.75; Table 3).

VAS versus modified Borg scale
For the total sample, the mean differences between the VAS and modified Borg scale were 0.93 (95% CI 0.69, 1.17) for patients' ratings and 0.88 (95% CI 0.72, 1.04) for caregiver's rating (Table 2). These results indicate higher dyspnea rating obtained with the VAS scale compared to the modified Borg scale regardless of the assessor. Our results also showed at least moderate reliability between scales (Table 3). The subgroup analyses showed similar ratings for both scales, regardless of the assessor, but the degree of concordance was higher for the lung cancer group (ICC 0.90 or greater) compared to the other lung diseases group (ICC ranging from 0.62 to 0.82).

Discussion

Our results showed that the caregiver's ratings of dyspnea in hospitalized patients with respiratory diseases were similar to the patients' self-ratings. Although higher scores were obtained with patients' self-assessment compared to caregiver's assessment, the difference of less than 1 point might not be considered clinically relevant because the minimal clinically important difference (MCID) is 1-unit for the modified Borg scale and approximately 1 to 2 cm for the VAS.[27] It is noteworthy, that the level of agreement between the

caregiver and patients' self-rating was better for those with lung cancer who had lower level of dyspnea compared to those with other lung conditions who had more severe dyspnea.

Although dyspnea encompasses two domains, in the present study, we focused on the sensory domain which is less subjective than the emotional domain. Secondly, the caregiver in charge of assessing dyspnea was a qualified physical therapist with 5 years of experience. Future studies are needed to investigate whether the results found in this study could be generalized to less experienced caregivers.

The majority of patients in the other lung disease group were diagnosed with COPD. Patients with COPD report more frequent and longer episodes of dyspnea than patients with lung cancer[23] and can be overperceivers for this symptom.[24] Differences in the medications used to treat the patients from both groups could also explain the differences in the reliability indices. Opioids have been shown to play a role in dyspnea perception[25] and they are more often prescribed in patient with lung cancer. However, we did not have the rate of opioids prescription in this patient population. The patient's past history also contributes to symptom perception and may differ based on diagnosis.[26]

The secondary purpose of the study established that the VAS and modified Borg scales are not interchangeable to quantify dyspnea severity. If the inter-instrument reliability was good, and even excellent in the group with lung cancer as previously observed in healthy subjects and patients with COPD patients,[28,29] such a difference was expected.[30,31] The

Table 3 Reliability of dyspnea rating in hospitalized patients with lung disease.

Reliability	ICCTotal sample	ICCLung cancer group	ICCOther lung diseases group
Caregiver vs *patient self-rating*			
VAS	0.79 (0.73, 0.84)	0.91 (0.87, 0.94)	0.56 (0.38, 0.69)
Borg	0.82 (0.77, 0.86)	0.94 (0.91, 0.96)	0.64 (0.49, 0.74)
VAS vs Borg (*inter-instrument*)			
Caregiver	0.90 (0.74, 0.95)	0.97 (0.95, 0.98)	0.82 (0.20, 0.93)
Patient	0.81 (0.68, 0.87)	0.93 (0.90, 0.95)	0.62 (0.26, 0.78)

ICC: intraclass correlation coefficient, VAS: visual analogue scale.
Values are intraclass correlation coefficients (95% confidence intervals).

high level of reliability suggests a standard offset between the scales.

The strengths of this study include the large cohort of patients with dyspnea and the variety of pulmonary diseases which offered a wide range of dyspneic conditions. Our results should be interpreted with caution. An emotional influence cannot be excluded.[32,33] Data were collected on hospitalized patients only, therefore making extrapolation to outpatient settings difficult in part due to potential anxiety related to hospitalization.[34] No objective measurements of lung function were used to determine the validity of the data.[35,36] Finally, data have limited external validity as a single examiner with five years of experience participated in the study.

Conclusion

We demonstrated that dyspnea can be accurately estimated by an experienced caregiver in hospitalized patients with lung diseases, using either a VAS or modified Borg scale. This is particularly the case for patients with lung cancer with lower level of dyspnea. A larger difference between raters was shown for patients with other lung diseases and more severe dyspnea. However, the VAS and Borg scales are not interchangeable, as the scores on the VAS were consistently higher than the scores on the Borg scale.

Conflicts of interest

None for all authors.

Acknowledgments

References

1. Parshall MB, Schwartzstein RM, Adams L, Banzett RB, Manning HL, Bourbeau J, et al. An official American Thoracic Society statement: update on the mechanisms, assessment, and management of dyspnea. *Am J Respir Crit Care Med*. 2012;185 (4):435—452. Feb 15.
2. Stevens JP, Dechen T, Schwartzstein R, O'Donnell C, Baker K, Howell MD, et al. Prevalence of dyspnea among hospitalized patients at the time of admission. *J Pain Symptom Manage*. 2018. Feb 21.
3. Frostad A, Soyseth V, Andersen A, Gulsvik A. Respiratory symptoms as predictors of all-cause mortality in an urban community: a 30-year follow-up. *J Intern Med*. 2006;259(5):520—529. May.
4. Frese T, Sobeck C, Herrmann K, Sandholzer H. Dyspnea as the reason for encounter in general practice. *J Clin Med Res*. 2011;3(5):239—246. Oct.
5. Ekstrom M, Sundh J, Larsson K. Patient reported outcome measures in chronic obstructive pulmonary disease: which to use? *Expert Rev Respir Med*. 2016;10(3):351—362.
6. De PS, Van D, I Lemaigre V, Verleden G, Demedts M. Van den Bergh O. Dyspnea: the role of psychological processes. *Clin Psychol Rev*. 2004;24(5):557—581. Sep.
7. Banzett RB, O'Donnell CR, Guilfoyle TE, Parshall MB, Schwartzstein RM, Meek PM, et al. Multidimensional dyspnea profile: an instrument for clinical and laboratory research. *Eur Respir J*. 2015;45(6):1681—1691. Jun.
8. von Leupoldt A, Dahme B. Differentiation between the sensory and affective dimension of dyspnea during resistive load breathing in normal subjects. *Chest*. 2005;128(5):3345—3349. Nov.
9. Dorman S, Byrne A, Edwards A. Which measurement scales should we use to measure breathlessness in palliative care? A systematic review. *Palliat Med*. 2007;21(3):177—191. Apr.
10. Campbell ML, Templin T, Walch J. Patients who are near death are frequently unable to self-report dyspnea. *J Palliat Med*. 2009;12(10):881—884. Oct.
11. Campbell ML, Templin T, Walch J. A Respiratory Distress Observation Scale for patients unable to self-report dyspnea. *J Palliat Med*. 2010;13(3):285—290. Mar.
12. Brunelli C, Kaasa S, Knudsen AK, Hjermstad MJ, Pigni A, Caraceni A. Comparisons of patient and physician assessment of pain-related domains in cancer pain classification: results from a large international multicenter study. *J Pain*. 2014;15 (1):59—67. Jan.
13. Higginson IJ, Gao W. Caregiver assessment of patients with advanced cancer: concordance with patients, effect of burden and positivity. *Health Qual Life Outcomes*. 2008;6:42.. Jun 2.
14. Silveira MJ, Given CW, Given B, Rosland AM, Piette JD. Patient-caregiver concordance in symptom assessment and improvement in outcomes for patients undergoing cancer chemotherapy. *Chronic Illn*. 2010;6(1):46—56. Mar.
15. Kottner J, Audige L, Brorson S, Donner A, Gajewski BJ, Hrobjartsson A, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *J Clin Epidemiol*. 2011;64 (1):96—106. Jan.
16. Burdon JG, Juniper EF, Killian KJ, Hargreave FE, Campbell EJ. The perception of breathlessness in asthma. *Am Rev Respir Dis*. 1982;126(5):825—828. Nov.
17. Borg GA. Psychophysical bases of perceived exertion. *Med Sci Sports Exerc*. 1982;14(5):377—381.
18. Borg GA. A category scale with ratio properties for intermodal and interindividual comparisons. In: Geissler H, Petzol P, eds. North Holland Publishing Co; 1980:25—34.
19. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods*. 1996;1(1):30—46. Mar.
20. Nickerson CE. A note on "a concordance correlation coefficient to evaluate reproducibility". *Biometrics*. 1997;53(4):1503—1507. Dec.
21. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420—428. Mar.
22. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155—163. Jun.
23. Weingartner V, Bausewein C, Higginson IJ, Scheve C, Murtagh FE, Voltz R, et al. Characterizing episodic breathlessness in patients with advanced disease. *J Palliat Med*. 2013;16 (10):1275—1279. Oct.
24. Scioscia G, Blanco I, Arismendi E, Burgos F, Gistau C, Foschino Barbaro MP, et al. Different dyspnoea perception in COPD patients with frequent and infrequent exacerbations. *Thorax*. 2017;72(2):117—121. Feb.
25. Gifford AH, Mahler DA, Waterman LA, Ward J, Kraemer WJ, Kupchak BR, et al. Neuromodulatory effect of endogenous opioids on the intensity and unpleasantness of breathlessness during resistive load breathing in COPD. *COPD*. 2011;8 (3):160—166. Jun.

26. Hoyer J, Reusch A, Leibing E. Accuracy of interoception and withdrawal of attention in airway diseases: bronchial asthma versus chronic obstructive bronchitis. *Verhaltenstherapie*. 1999;9:140−147. Ref Type: Journal (Full).

27. Ries AL. Minimally clinically important difference for the UCSD shortness of breath questionnaire, Borg scale, and visual analog scale. *COPD*. 2005;2(1):105−110. Mar.

28. Bogaerts K, Notebaert K, Van D I, Devriese S, De PS, Van den Bergh O. Accuracy of respiratory symptom perception in different affective contexts. *J Psychosom Res*. 2005;58 (6):537−543. Jun.

29. von Leupoldt A, Ambruzsova R, Nordmeyer S, Jeske N, Dahme B. Sensory and affective aspects of dyspnea contribute differentially to the Borg scale's measurement of dyspnea. *Respiration*. 2006;73(6):762−768.

30. Ohbayashi H. Comparison of the rapid effects of single inhalations of formoterol and tiotropium bromide on respiratory function and COPD symptoms in a randomized crossover study. *Respir Investig*. 2017;55(6):348−356. Nov.

31. Boshuizen RC, Vincent AD, van den Heuvel MM. Comparison of modified Borg scale and visual analog scale dyspnea scores in predicting re-intervention after drainage of malignant pleural effusion. *Support Care Cancer*. 2013;21 (11):3109−3116. Nov.

32. Gift AG, Plaut SM, Jacox A. Psychologic and physiologic factors related to dyspnea in subjects with chronic obstructive pulmonary disease. *Heart Lung*. 1986;15(6):595−601. Nov.

33. Simon ST, Bausewein C, Schildmann E, Higginson IJ, Magnussen H, Scheve C, et al. Episodic breathlessness in patients with advanced disease: a systematic review. *J Pain Symptom Manage*. 2013;45(3):561−578. Mar.

34. Al Aqqad SMH, Tangiisuran B, Hyder Ali IA, Md Kassim RMN, Wong JL. Tengku Saifudin TI. Hospitalisation of multiethnic older patients with AECOPD: exploration of the occurrence of anxiety, depression and factors associated with short-term hospital readmission. *Clin Respir J*. 2017;11(6):960−967. Nov.

35. Marin JM, Carrizo SJ, Gascon M, Sanchez A, Gallego B, Celli BR. Inspiratory capacity, dynamic hyperinflation, breathlessness, and exercise performance during the 6-minute-walk test in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2001;163(6):1395−1399. May.

36. Ozalevli S, Ucan ES. The comparison of different dyspnoea scales in patients with COPD. *J Eval Clin Pract*. 2006;12 (5):532−538. Oct.