Title: Use of the modified Borg scale and Numerical Rating Scale to measure chronic breathlessness: a pooled data analysis.

Authors:

Johnson MJ, Professor of Palliative Medicine, Hull York Medical School, University of Hull, Hull, UK

Close L, Consultant in Palliative Medicine, Lindsey Lodge Hospice, Scunthorpe, North Lincolnshire, UK.

Gillon SC, Consultant in Palliative Medicine, Leeds Teaching Hospitals NHS Trust, Leeds, UK.

Molassiotis A, Chair Professor and Head of School, School of Nursing, The Hong Kong Polytechnic University, Hung Hom, HONG KONG

Lee P.H., Research Assistant Professor, School of Nursing, The Hong Kong Polytechnic University, Hung Hom, HONG KONG

Farquhar MC. Senior Research Associate, Primary Care Unit, Department of Public Health and Primary Care, University of Cambridge, Institute of Public Health, Robinson Way, Cambridge, CB2 0SR

Author for correspondence: Johnson MJ, Professor of Palliative Medicine, Hull York Medical School, University of Hull, Hull, UK HU6 7RX Tel: 01482 463309; Email miriam.johnson@hyms.ac.uk

On behalf of the Breathlessness Research Interest Group (BRIG)

Short message: Measuring chronic dyspnoea in clinical studies: NRS better than mBorg.

Words 1196

BACKGROUND

The subjective nature of the experience of chronic breathlessness (dyspnoea) creates challenges for patients who need to communicate its intensity, and for clinicians and researchers who need to measure the symptom in order to plan management and assess the effect of interventions.

The numerical rating scale (NRS) [1] and modified Borg scale (mBorg) [2] are recommended measures for breathlessness. [3] However, their use has extended *beyond* their initial validation. NRS scales using different time frames ("now" and "average") have been validated,[1, 4, 5] but not for the mBorg. Further, participants might have a preference for mBorg scores with associated verbal descriptors.

Objectives: To investigate whether: i) there is a response bias against using mBorg numerical ratings that lack categorical labels , ii) the timeframe (average/24 hours, "worst", "now" or "at rest") of the mBorg or NRS effects participants' assessment, iii) mBorg and NRS scores are correlated

METHODS

This was a secondary analysis of pooled data from 1,048 participants (510 men, 396 women, 142 gender data unavailable: diagnoses: cancer 223 [21.3%], heart failure 200 [19%], non-malignant lung disease 617 [59%]) with breathlessness due to a variety of causes from 10 studies of people where mBorg, at least, was measured. Where both mBorg and NRS were measured, these were concurrent. All studies used the same version of the Borg; a variant of the Borg Category-Ratio scale with a maximum value of 10, and with verbal descriptors missing for values 6 and 8.

Most contributing studies are described more fully elsewhere[7-14] but are summarised here as follows: i) quantifiable data from a primarily qualitative study (study 1. N=47; mean age 69 (range 46-92); measures mBorg (average 24hrs, worst, rest, non-specific now, exertion) with NRS for 7 participants[8] ii) two phase III trials (study 2. N=35; mean age 70 (41-89); measures mBorg and NRS (average 24hrs, worst, rest, non-specific now, exertion)][10]: study 3. N = 154; mean age 71 (28-91); measures mBorg and NRS (rest, exertion)[12]:) iii)

two feasibility studies, (study 4. N=46; mean age 69.5 (62-73) measures mBorg and NRS (rest)[7]: study 5. N=13; mean age 67 (53-80); measures mBorg only (rest, exertion)[13] iv) five observational (study 6. N = 50; mean age 69 (42-83); measures mBorg only (pre and post exertion)][9]: study 7. N = 109; mean age 65 (38-52); measures mBorg only (now)[11]: study 8. N = 99; measures mBorg only (average over previous 24 hours) [personal communication, unpublished data]: study 9. N = 353; mean age 65 (24-90); measures mBorg only (average and worst over previous 24 hours)[14]: study 10. N = 142; mean age 69 (34-91); measures mBorg and NRS (average, worst over past 24 hours, now) [personal communication, unpublished data]) Proxy scores were excluded.

The individual distributions of mBorgs and NRSs (average, worst, now, rest, exertion) were visualized with predicted values using truncated Poisson distribution with their corresponding mean plotted as a reference. Descriptive statistics including mean, standard deviation, and frequency were examined. The strength of association between mBorgs and corresponding NRSs was examined using two-way mixed intraclass correlation (consistency).

RESULTS

The following measures were available for analysis: i) mBorg; The frequency of scores for numbers 6 and 8 (no verbal descriptor) was less than expected. There were also fewer than expected measures for 0.5 (verbal descriptor of "Very very weak [just noticeable)]").
In general, scores for average/24 hours were normally distributed for mBorg (other than the pattern noted above) and NRS. However, no NRS "average" scores exceeded 8. Although an NRS score of 7 is considered "severe", equivalent to an mBorg of 5, mBorg "average" scores included the maximum of 10. (see Figure 1)

The pattern of scores for NRS and mBorg "worst"/24 hours was similar although, as expected given the equivalent severity scores, there were more high NRS scores.

For point in time measures the patterns for mBorg and NRS "exertion" were similar, with few mild scores. Conversely mBorg "at rest", "now", and NRS "at rest" scores shared a similar pattern but with very few severe scores. However the NRS "now" had measures across the response spectrum, including very severe scores.

The strongest association between NRS (N=21; mean 7.23 [1.80]) and mBorg (N=261; mean 5.55 [2.18]) was for "on exertion" (ICC=0.66, 95% CI 0.33, 0.85); the weakest for "now" (NRS=106, mean 4.51[2.72]; MBorg N=368, mean 2.36[1.79]; ICC= 0.14, ─0.05, 0.33). Others were: "average" 0.51, (0.15, 0.75); "worst" 0.55, (0.34, 0.71)and "rest" 0.33 (─0.09, 0.66).

DISCUSSION

Our data indicate preferential reporting of mBorg scores with descriptors. This may be due to the mBorg's stem question: "Choose a number *whose words* best describe…". A less than predicted use of 0.5, despite a descriptor, suggests that "very very weak" is either indistinguishable in the context of chronic breathlessness or "0.5" is not understood; the VAS may be more sensitive at reporting breathlessness due to light intensity work.[15]

Apart from the observed reduction in non-descriptor mBorg scores, the pattern of mBorg and NRS scores in relation to the previous 24 hours appears to be as expected apart from a possible ceiling with the NRS.

The observed pattern of responses for NRS "now" likely reflects the contemporaneous context. For example a patient waiting in the clinic room for some time will respond differently to one who has hurried into clinic. Thus, unless the measure is taken with close definition of the circumstances of "now", responses will be difficult to interpret.

Despite the numerical discrepancy between the two scales, the intracluster correlations were moderate for "on exertion", and "average", albeit with wide confidence intervals, suggesting that the mBorg might be used to assess intensity of breathlessness on average/past 24 hours. The mBorg and NRS "now" were poorly correlated, presumably for the reasons above. It should be noted that with some ICC calculations there is a large discrepancy between the smaller and larger n. Therefore, the ICCs should be interpreted with caution because the missing data cannot be assumed to be missing at random.

Implications for clinical studies

These data suggest that there is a participant response bias against using numerical ratings that lack categorical labels, in which case, the scale would 'lose' the ratio properties that

Borg wanted to preserve. Therefore we recommend that given the non-controlled conditions in chronic breathlessness clinical studies, the NRS is used. Reported mBorg values may differ if the stem is simplified to "*choose a number* to describe…".

The NRS "at rest" and "on exertion", appear useful as "point in time" measures. However, the circumstances of "now" should be stipulated. Given the possible ceiling for "average" NRS scores, notwithstanding the issue above, the mBorg (average/24 hours) may be preferable in populations with severe daily breathlessness.

CONCLUSIONS

The analysis of this pooled data from people with chronic breathlessness suggests that there is a response bias in favour of mBorg responses with a verbal descriptor. The theoretical advantages of the mBorg scale under known and scalable stimulus conditions (e.g., in pulmonary rehabilitation programs or cardiopulmonary exercise testing) therefore are not necessarily maintained in less-controlled clinical studies. A change in mBorg stem question should be considered and tested. The NRS scale should be used in preference, except for people with very severe breathlessness. The context of "point in time" measures should be clearly stated on measure-completion.

ACKNOWLEDGEMENTS

With thanks to members of BRIG who provided data:

Claudia Bausewein, University of Munich, Germany; Saskie Dorman, Poole Hospital NHS Foundation Trust, UK; Morag Farquhar and Sara Booth, University of Cambridge, UK; Alex Molassiotis, The Hong Kong Polytechnic University. Hong Kong; Stephen Oxberry, Calderdale & Huddersfield NHS Foundation Trust; Farida Malik, Eastbourne and East Sussex Healthcare NHS Trust, UK; Steffen Simon, University of Cologne, Germany; Kyle Pattinson, Associate Professor, Senior Clinical Research Fellow, University of Oxford; Janelle Yorke, The University of Manchester, UK; Patrick White, King's College London, UK.

AUTHOR DISCLOSURE STATEMENT No competing financial interests exist

Reference List

(1) Gift AG, Narsavage G. Validity of the numeric rating scale as a measure of dyspnea. *Am J Crit Care* 1998 May;**7**(3):200-4.

(2) Muza SR, Silverman MT, Gilmore GC, Hellerstein HK, Kelsen SG. Comparison of scales used to quantitate the sense of effort to breathe in patients with chronic obstructive pulmonary disease. *Am Rev Respir Dis* 1990 April;**141**(4 Pt 1):909-13.

(3) Bausewein C, Farquhar M, Booth S, Gysels M, Higginson IJ. Measurement of breathlessness in advanced disease: a systematic review. *Respir Med* 2007 March;**101**(3):399-410.

(4) Gift AG. Validation of a vertical visual analogue scale as a measure of clinical dyspnea. *Rehabil Nurs* 1989 November;**14**(6):323-5.

(5) Wilcock A, Crosby V, Clarke D, Tattersfield A. Repeatability of breathlessness measurements in cancer patients. *Thorax* 1999 April;**54**(4):375.

(6) Morris NR, Sabapathy S, Adams L, Kingsley RA, Schneider DA, Stulbarg MS. Verbal numerical scales are as reliable and sensitive as visual analog scales for rating dyspnea in young and older subjects. *Respir Physiol Neurobiol* 2007 August 1;**157**(2-3):360-5.

(7) Molassiotis A, Charalambous A, Taylor P, Stamataki Z, Summers Y. The effect of resistance inspiratory muscle training in the management of breathlessness in patients with thoracic malignancies: a feasibility randomised trial. *Support Care Cancer* 2015 June;**23**(6):1637-45.

(8) Simon ST, Higginson IJ, Benalia H et al. Episodic and continuous breathlessness: a new categorization of breathlessness. *J Pain Symptom Manage* 2013 June;**45**(6):1019-29.

(9) Herigstad M, Hayen A, Evans E et al. Dyspnea-related cues engage the prefrontal cortex: Evidence from functional brain imaging in COPD. *Chest* 2015 July 2.

(10) Oxberry SG, Torgerson DJ, Bland JM, Clark AL, Cleland JG, Johnson MJ. Short-term opioids for breathlessness in stable chronic heart failure: a randomized controlled trial. *Eur J Heart Fail* 2011 September;**13**(9):1006-12.

(11) Bausewein C, Booth S, Gysels M, Kuhnbach R, Haberland B, Higginson IJ. Understanding breathlessness: cross-sectional comparison of symptom burden and palliative care needs in chronic obstructive pulmonary disease and cancer. *J Palliat Med* 2010 September;**13**(9):1109-18.

(12) Farquhar MC, Prevost A, McCrone P et al. Is a specialist breathlessness service more effective and cost-effective for patients with advanced cancer and their carers than standard care? Findings of a mixed-method randomised controlled trial. *BMC Med* 2014 October 31;**12**(1):194.

(13) Farquhar MC, Higginson IJ, Fagan P, Booth S. The feasibility of a single-blinded fast-track pragmatic randomised controlled trial of a complex intervention for breathlessness in advanced disease. *BMC Palliat Care* 2009;**8**:9.

(14) Yorke J, Moosavi SH, Shuldham C, Jones PW. Quantification of dyspnoea using descriptors: development and initial testing of the Dyspnoea-12. *Thorax* 2010 January;**65**(1):21-6.

(15) Adams L, Lane R, Shea SA, Cockcroft A, Guz A. Breathlessness during different forms of ventilatory stimulation: a study of mechanisms in normal subjects and respiratory patients. *Clin Sci (Lond)* 1985 December;**69**(6):663-72.

Figure 1. MBorg and NRS scores in relation to perception of breathlessness over the past 24 hours and scores measured now, at rest or on exertion. Histograms with predicted truncated Poisson probabilities. Average scores (MBorg N = 498; NRS N = 86); Worst scores (MBorg N = 559; NRS N = 106); Now (MBorg N = 368; NRS N= 108); rest (MBorg N = 261; NRS N = 60); exertion (MBorg N = 261 , NRS N =23).