Validation of the Dyspnea Exertion Scale of Breathlessness in people with life-limiting illness.

Jacob Sandberg, MD[1]; Miriam J Johnson[2]; David C Currow[2,3]; Magnus Ekström, MD, PhD[1,3]

[1] Dept of Clinical Sciences, Division of Respiratory Medicine & Allergology, Lund University, Lund, Sweden.

[2] Wolfson Palliative Care Research Centre, Hull York Medical School, University of Hull, Hull, UK .

[3] IMPACCT, Faculty of Health, University of Technology Sydney, Sydney, New South Wales, Australia.

Corresponding author: Jacob Sandberg, Jämjö Vårdcentral, Hammarbyvägen 6 37300, Jämjö, Sweden; jacob.sandberg@gmail.com

Funding: JS was funded by an unrestricted grant from the Scientific Committee of Blekinge County Council.

Conflict of interest: No conflicts of interest exist for the authors.

Running head: Measurement of chronic breathlessness in advanced disease

Number of tables: 4**; Number of figures:** 5**; Number of references:** 25

Word count for body of text: 2007

ABSTRACT

Background: Although chronic breathlessness is common in life-limiting illnesses, validated, feasible instruments to measure *functional impact* of the symptom in this population are scarce. We aimed to validate the Dyspnea Exertion Scale (DES) compared with the modified Medical Research Council (mMRC) breathlessness scale for test-retest reliability, construct validity and responsiveness in people with life-limiting illness.

Methods: A total of 188 participants, 66% male, with chronic breathlessness mostly (70%) due to chronic pulmonary disease (COPD) self-reported evening scores of mMRC, DES, Numerical rating scale (NRS) and Eastern Cooperative Oncology Group (ECOG) over 9 days.

Results: 44% (n=81) scored the highest score on mMRC indicating a ceiling effect not seen with DES. Both scales had moderate to good test-retest agreement (89% DES; 84% mMRC; $p<0.001$ for both). Analyses for construct validity showed that higher DES and mMRC scores were correlated with higher NRS breathlessness intensity scores and ECOG scores throughout the 9 days. In longitudinal analyses, DES (DES $r=0.30$, $p<0001$) was more responsive to change in NRS score over nine days than the mMRC (mMRC $r=0.16$, $p=0.03$).

Conclusion: Compared to mMRC, DES had comparable or better measurement properties in terms of test-retest reliability and construct validity and could be used as a discriminative tool in this population, but both scales are to insensitive to change to be used as an outcome in clinical trials.

Keywords: Breathlessness, measurement, Medical Research Council breathlessness scale, Dyspnea Exertion Scale (DES), mMRC, DES, validation.

INTRODUCTION

Chronic breathlessness is common and causes major suffering in patients with life-limiting illness.[1] It is associated with increased morbidity and mortality, including worse quality of life and increased dependency on health services.[2, 3] Nearly all people with life-limiting cardiac or respiratory disease will experience chronic breathlessness late in their disease trajectory, becoming persistent and triggered by minimal exertion, or present even at rest.[1, 3-5]

There are few validated simple unidimensional instruments which measure the *functional impact* in people with chronic breathlessness useful and feasible for categorizing patients and for prognosis purposes.[1, 6] Whilst measures of exercise-induced breathlessness may be applicable in the early, more stable phases of pulmonary disease, arguably these are less relevant in later stages when breathlessness is triggered by minimal movement or even at rest without an obvious precipitant. [6] Existing measures are mostly disease-specific (cancer, COPD or motor neuron disease) and designed for use in a research, rather than clinical, setting. [7] Routine clinical assessment of chronic breathlessness is important to identify and manage this often-neglected symptom.[7, 8]

The modified Medical Research Council (mMRC) breathlessness scale is a widely used measure of the presence and severity of breathlessness in relation to physical activities (Table 1).[2, 9-12] It was developed in the 1950's with the main purpose to categorize disability due to breathlessness in research and is still the most used instrument across both clinical and research settings. The mMRC scale is discriminative but to insensitive to change to be used as outcome in clinical trials, it is recommended by international guidelines for categorizing the severity of COPD [13] and strongly predicts increased hospitalization, reduced quality of life and mortality, being a better predictor of death than the degree of airflow limitation.[14, 15]

However, the mMRC scale is only weakly associated with physiological and functional measures of impairment, and is poorly responsive to change over time and in relation to therapy, perhaps because it only has five categories.[16, 17] In patients with severe illness there might be a risk of a ceiling effect (defined as >15% of respondents selecting the highest score category [18]) as many patients, despite varying levels of symptom and disability, are likely to be in the worst category.[19]

The Dyspnea Exertion Scale (DES; Table 1) was developed from the mMRC scale for use in people with advanced cancer.[20] DES may offer better face validity than mMRC for people with severe COPD who have breathlessness at rest or with minimal exertion.[19] DES was presented within an MD-thesis [20] and has never been published as a peer reviewed paper or compared formally against the mMRC. The relative merits of mMRC and DES for measuring exertion-related chronic breathlessness due to life-limiting illness is unknown.[6, 19]

The aim of this study was to compare DES with mMRC in terms of test-retest reliability, construct validity and responsiveness for measuring chronic breathlessness in people with life-limiting illnesses.

METHODS

Study design and population

This was a secondary analysis of a multi-center, double-blind, randomized controlled trial of ambulatory oxygen compared with medical air for one week in people with chronic breathlessness.[21]

Participants (n=239) were recruited between April 2006 and March 2008 from outpatient pulmonary, palliative care, and primary care clinics in Australia (five sites), USA (two sites)

and in the UK (two sites). Only data from the Australian participants were available for this analysis (n=188).

Eligible participants were: aged ≥ 18 years; with a life-limiting illness who did not qualify for long term oxygen therapy; partial pressure of oxygen in arterial blood (PaO_2) > 7.3 kPa breathing ambient air; mMRC ≥3 at screening despite optimal disease management; life-expectancy longer than one month; and stable medication for at least the previous week. Exclusion criteria included current smoking; a respiratory or cardiac event in the previous seven days; anemia (hemoglobin < 100g/L); partial pressure of carbon dioxide in arterial blood <8 $PaCO_2$> 6.7kPa; or cognitive impairment (Mini Mental State examination score < 24 points).[22]

Assessments

Baseline was defined as Day 1 (two days before randomisation) and assessments continued to Day 9 thus including seven treatment days.

Evening values of DES (using the question 'What is your breathlessness like right now?') mMRC ('What is your best exertional performance today?'), and a 11-point numerical rating scale (NRS) ('How is your breathlessness right now') between 0 (not breathless at all) and 10 (breathlessness as bad as you can imagine) was recorded by the study participant for each of the 9 days.[10]

Functional status was assessed by research personnel on Days 1, 3 and 9 using Eastern Cooperative Oncology Group (ECOG).[23, 24] ECOG was categorized as: asymptomatic (0), symptomatic but ambulatory (1), symptomatic, <50% in bed during the day (2), symptomatic,> 50% in bed but not bedbound (3), and bedbound (4).[23]

Ethical considerations

The study was approved by the Southern Adelaide Health Service Human Research Ethics Committee as well as local research and ethics committees or institutional review boards of all participating sites. All participants provided written informed consent.

Statistical analyses

Baseline patient characteristics were summarized using mean with standard deviation (SD) and median with range or interquartile range (IQR) for continuous variables with normal and skewed distribution, respectively. Categorical variables were expressed as frequencies and percentages.

The measurement properties of DES and mMRC were evaluated in concordance with international guidelines for the evaluation of patient reported outcomes measure.[25] Test-retest reliability of DES and mMRC were assessed using ratings on day 1 and 2 (before randomisation). Ratings were cross-tabulated and test-retest reliability was assessed using the weighted kappa statistics with linear weights. A kappa value of 0.7 or above is considered good.[18, 25]Construct validity (meaning the correlation with other relevant measures) was assessed using Kendall's tau B rank correlation coefficient, looking at associations between DES and mMRC values and NRS and ECOG scores, all from day 1. Responsiveness was assessed by the regression slope of NRS and DES over time from Day 1 through 9 for each individual participant, accounting for correlations. Patients with recorded ratings for fewer than half the days were excluded (n=11) from the responsiveness analyses. Statistical significance was defined as two-sided p-value < 0.05. Statistical analyses were conducted using the software packages Stata, version 14.1 (StataCorp LP; College Station, TX).

RESULTS

Patient characteristics

Table 2 shows baseline characteristics of the 188 included participants; 66% were males and the most common cause of breathlessness was COPD (70 %). Nearly 40% of the participants had previously been prescribed long term oxygen therapy. The mean DES and mMRC scores at baseline were 2.3 and 2.9, respectively (Table 2).

Score distribution and reliability

The distribution of mMRC on DES scores and their inter-relation is shown in Figure 1. Out of all respondents, 44% scored the highest category (4) on mMRC, indicating a ceiling effect in this setting, while only 6% scored the highest category (5) on DES. Most of the responses categorized as mMRC 2-4 scored DES at category 2. Nine individuals (6.6%) scored the highest category on mMRC and the lowest (1) on DES at the same time (Figure 1).

The relationship between both scales and the NRS rating is shown in figures 2a and b. Test-retest agreement was moderate to good for both scales (89% DES; 84% mMRC; $p < 0.0001$) with kappa values of approximately 0.6 for both scales (Table 3, e-figure 2a-b).

Construct validity

Both DES and mMRC were correlated with NRS breathlessness intensity scores and ECOG scores (Table 4). All correlations were highly statistically significant (DES and mMRC; $p<0.0001$, DES and NRS; $p<0.0001$, mMRC and NRS; $p=0.0468$, DES and ECOG; $p=0.003$ mMRC and ECOG; $p<0.0001$). The NRS was correlated more strongly with the DES (Kendall Tau-B=0.32) than the mMRC (Kendall's Tau-B=0.12).

Responsiveness

The change in DES and NRS scores over the nine-day period is shown in Figure 3. The change scores for both scales were approximately normally distributed. The mean change is less than zero in each case, indicating an overall tendency for both breathlessness scores to decrease over the study period. A change in DES was associated with change in NRS, r=0.3 ($p<0.0001$) (Figure 3a). The mMRC also showed a statistically significant association with change in NRS, r=0.16 (p=0.03) (Figure 3).

DISCUSSION

In this first validation study in people with life-limiting illnesses, DES, compared to mMRC, had similar or slightly better test-retest reliability and construct validity. Both scales were relatively unresponsive to change. The need of an instrument more adapted to this setting than mMRC is highlighted by that 44% of the participants were in the highest mMRC category. DES differentiated the group of patients with mMRC 4 and increasing DES scores were more closely correlated to increasing breathlessness intensity (NRS scores) compared with increasing mMRC scores. This study also identified a potential problem with DES as the comparative distribution shows that most mMRC values 3-5 equate to DES 2 values, indicating that the category may be too broad. Furthermore, the response options 2-5 are not mutually exclusive. The participant may both be breathless when walking around the house as well as when getting out of bed and the category 2 might be the first one that applies to most respondents. A further problem with measurements of this kind is the fact that the scales are used differently in different settings. There is a need for standardization on this issue, further highlighted by the fact that nine individuals rated the highest score on mMRC and the lowest on DES at the same time. Compared to mMRC, DES was more responsive to changes in breathlessness intensity (NRS scores) but correlations were weak. Although DES may be

more useful for description and discriminating patient populations and was more responsive than mMRC, it may have insufficient responsiveness to be used as endpoints in clinical trials.

MMRC has been shown to have a prognostic value for mortality in COPD, exceeding that of the level of airflow limitation.[15] Given the similarities in phrasing between mMRC and DES and the comparable or better reliability and construct validity, the associations with prognosis and clinical outcomes might be similar or better for DES than for mMRC in people with life-limiting illness. Whenever possible the multidimensional aspects of breathlessness should be assessed [26] [27]

Strengths of this study include the use of a quality data set with a large cohort of patients with life-limiting illness and chronic breathlessness in a randomized controlled trial, with standardized, longitudinal collection of clinically relevant data over nine days.

Potential limitations were that the eligibility criteria of the randomized controlled trial may limit the generalizability to all patients with life-limiting disease, which should be evaluated in further studies in this setting. The questions were not asked precisely the same which might affect the results, probably in the direction of underestimating mMRC. Full understanding of the impacts of chronic breathlessness may need a multi-dimensional measurement in research, but in clinical practice among people with life-limiting illness it might be more useful to focus on simple and unidimensional measurements.

This study has important implications for practice and research. For clinicians, DES is a discriminative tool that could be used for assessing symptom prevalence and functional impact of breathlessness to describe and select patient populations in clinical care and research. Both scales are insufficiently responsive to be used as an outcome measure of therapy but DES had better score distribution in severe illness with less ceiling effect. Further research should focus on the optimal questioning and standardizing the use to ensure a better

distribution. In the light of the problems showed with both scales, perhaps a combination of the two scales could prove useful to give a better distribution and differentiation of patients.

Conclusion

Compared to mMRC, DES had comparable or better measurement properties in terms of test-retest reliability and construct validity and could be used as a discriminative tool in this population, but both scales are to insensitive to change to be used as an outcome in clinical trials.

ACKNOWLEDGEMENTS

We would like to thank Professor Emeritus Martin Bland at Department of Health Sciences, University of York for his work with the statistical analysis, as well as all participating staff and patients.

REFERENCES

1. Johnson, M.J., et al., *Towards an expert consensus to delineate a clinical syndrome of chronic breathlessness.* Eur Respir J, 2017. **49**(5).
2. Parshall, M.B., 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): p. 435-52.
3. Gronseth, R., et al., *Predictors of dyspnoea prevalence: results from the BOLD study.* Eur Respir J, 2014. **43**(6): p. 1610-20.
4. Bowden, J.A., et al., *Predictors of chronic breathlessness: a large population study.* BMC Public Health, 2011. **11**: p. 33.
5. Currow, D.C., et al., *Do the trajectories of dyspnea differ in prevalence and intensity by diagnosis at the end of life? A consecutive cohort study.* J Pain Symptom Manage, 2010. **39**(4): p. 680-90.
6. Bausewein, C., et al., *Measurement of breathlessness in advanced disease: a systematic review.* Respir Med, 2007. **101**(3): p. 399-410.
7. Bausewein, C., S. Booth, and I.J. Higginson, *Measurement of dyspnoea in the clinical rather than the research setting.* Curr Opin Support Palliat Care, 2008. **2**(2): p. 95-9.
8. Banzett, R.B. and C.R.O.D. ScD, *SHOULD WE MEASURE DYSPNEA IN EVERYONE?* The European respiratory journal, 2014. **43**(6): p. 1547-1550.
9. Schilling, R.S., J.P. Hughes, and I. Dingwall-Fordyce, *Disagreement between observers in an epidemiological study of respiratory disease.* Br Med J, 1955. **1**(4905): p. 65-8.
10. Bestall, J.C., et al., *Usefulness of the Medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease.* Thorax, 1999. **54**(7): p. 581-6.
11. Fletcher, C., *DISCUSSION ON THE DIAGNOSIS OF PULMONARY EMPHYSEMA.* Proceedings of the Royal Society of Medicine, 1952. **45**(9): p. 576-586.
12. Mahler, D.A. and C.K. Wells, *Evaluation of clinical methods for rating dyspnea.* Chest, 1988. **93**(3): p. 580-6.
13. Hess, M.W., *The 2017 Global Initiative for Chronic Obstructive Lung Disease Report and Practice Implications for the Respiratory Therapist.* Respir Care, 2017.
14. Currow, D.C., et al., *Chronic breathlessness associated with poorer physical and mental health-related quality of life (SF-12) across all adult age groups.* Thorax, 2017.
15. Nishimura, K., et al., *Dyspnea is a better predictor of 5-year survival than airway obstruction in patients with COPD.* Chest, 2002. **121**(5): p. 1434-40.
16. Chhabra, S.K., A.K. Gupta, and M.Z. Khuma, *Evaluation of three scales of dyspnea in chronic obstructive pulmonary disease.* Ann Thorac Med, 2009. **4**(3): p. 128-32.
17. Perez, T., et al., *Modified Medical Research Council scale vs Baseline Dyspnea Index to evaluate dyspnea in chronic obstructive pulmonary disease.* Int J Chron Obstruct Pulmon Dis, 2015. **10**: p. 1663-72.
18. Terwee, C.B., et al., *Quality criteria were proposed for measurement properties of health status questionnaires.* J Clin Epidemiol, 2007. **60**(1): p. 34-42.
19. Dorman, S., et al., *Researching breathlessness in palliative care: consensus statement of the National Cancer Research Institute Palliative Care Breathlessness Subgroup.* Palliat Med, 2009. **23**(3): p. 213-27.
20. Heyse-Moore, L., *On dyspnoea in advanced cancer*. 1993, Southampton University.
21. Abernethy, A.P., et al., *Effect of palliative oxygen versus room air in relief of breathlessness in patients with refractory dyspnoea: a double-blind, randomised controlled trial.* Lancet, 2010. **376**(9743): p. 784-93.
22. Folstein, M.F., S.E. Folstein, and P.R. McHugh, *"Mini-mental state". A practical method for grading the cognitive state of patients for the clinician.* J Psychiatr Res, 1975. **12**(3): p. 189-98.

23. Oken, M.M., et al., *Toxicity and response criteria of the Eastern Cooperative Oncology Group.* Am J Clin Oncol, 1982. **5**(6): p. 649-55.
24. Karnofsky, D.A., WH; Craver, LF, Burhenal JH. , *The use of the nitrogen mustards in the palliative treatment of carcinoma - with particular reference to bronhogenic carcinoma.* Cancer, 1948. **1 (4)**: p. 634-56.
25. (PCORI), P.C.O.R.I., *The design and selection of Patient-Reported Outcomes Measures (PROM's) for Use in Patient Centered Outcomes Research.* Oxford Outcomes, 2012.
26. Banzett, R.B., et al., *Multidimensional Dyspnea Profile: an instrument for clinical and laboratory research.* Eur Respir J, 2015. **45**(6): p. 1681-91.
27. Yorke, J., et al., *Quantification of dyspnoea using descriptors: development and initial testing of the Dyspnoea-12.* Thorax, 2010. **65**(1): p. 21-26.

TABLES

Table 1. Dyspnea Exertion Scale (DES) and Medical Research Council (mMRC) breathlessness scale

Dyspnea Exertion Scale (DES)	Medical Research Council (mMRC) scale
1 = I am able to walk at my own pace on the level without getting out of breath	0 = Not troubled by breathlessness, except with strenuous exercise.
2 = I become breathless if I walk around the house or on the hospital ward on the level at my own pace	1 = Troubled by shortness of breath when hurrying on the level or walking up a slight hill
	2 =Breathless or has to stop for breath when walking at own pace on the level
	3 = Stops for breath after walking about 100 yards (90m) or after a few minutes on the level
3 = I become breathless if I move around in bed or get out of bed	4 = Breathless when dressing or undressing
4 = I become breathless on talking	
5 = I am breathless at rest	

Table 2. Baseline characteristics

Variable	**All (n=188)**
Age, mean (SD)	73.4 (10.1)
Gender (%)	
Male	124 (66)
Missing	1 (0.5)
Causes of breathlessness (%)	
COPD	131(70)
Primary lung cancer	24(13)
Other causes	36(19)
PaO_2, kPa Mean (SD)	10.1 (1.6)
$PaCO_2$,kPa Mean (SD)	5.2 (0.5)
Oxygen treatment (%)	38.8 (4.4)
DES (n=177) (%)	
1	33 (19)
2	86 (49)
3	27 (15)
4	25 (14)
5	6 (3)
Missing	9
mMRC(n=182) (%)	
1	25 (14)
2	47 (26)
3	29 (16)
4	81 (44)
Missing	6
ECOG(n=181) (%)	
1	52 (28)
2	80 (42)
3	49 (26)
Missing	7

SD= standard deviation, COPD = Chronic Obstructive Pulmonary Disease, PaO2= partial pressure of oxygen in arterial blood, PaC02= partial pressure of carbon dioxide in arterial blood, DES = Dyspnea Exertion Scale, mMRC = modified Medical Research Council, ECOG = Eastern Cooperative Oncology Group (ECOG)

Table 3: Test-retest reliability of DES and mMRC.

	Agreement	**Expected agreement**	**Kappa**	**P-value**
DES	89.12%	72.94%	0.598	<0.0001
mMRC	83.70%	59.01%	0.602	<0.0001

DES= Dyspnea Exertion Scale, mMRC= modified Medical Research Council, Agreement and expected agreement for ratings of breathlessness between two days. Test-retest reliability assessed using the weighted kappa statistics with linear weights**.**

Table 4: Associations between DES and mMRC values and NRS and ECOG scores

Comparison	**Correlation between the scores (Kendall's tau B)**	**P - value**
DES vs mMRC	0.32	<0.0001
DES vs NRS	0.32	<0.0001
mMRC vs NRS	0.12	0.0468
DES vs ECOG	0.23	0.0003
mMRC vs ECOG	0.30	<0.0001

Associations were measured using Kendall's tau B ranging from 1 (all rankings are the same) to -1 (all rankings are the reverse of the other). DES = Dyspnea Exertion Scale, mMRC =modified Medical Research Council, NRS = numerical rating scale, ECOG = Eastern Cooperative Oncology Group (ECOG)

FIGURES

Figure 1: Distribution of modified Medical Research Council (mMRC) scores vs. Dyspnea Exertion Scale (DES) scores

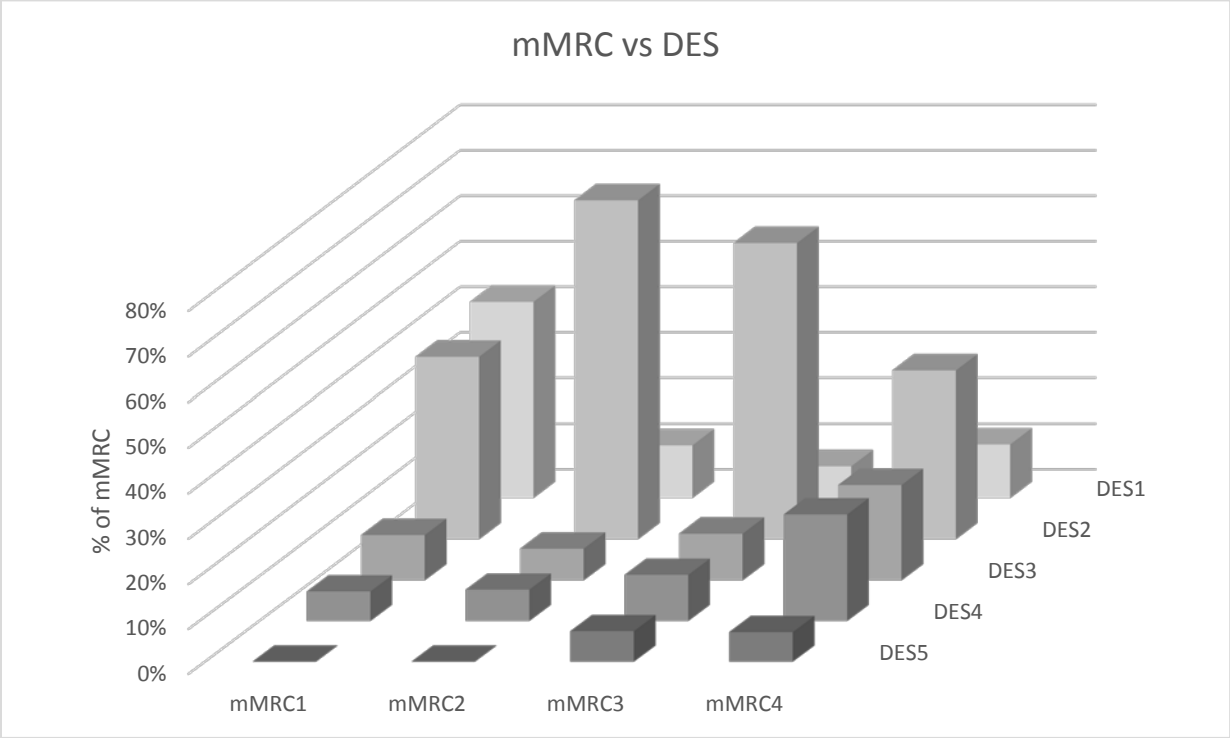

Figure 2a: Boxplot showing the distribution of Dyspnea Exertion Scale (DES) per Numerical Rating Scale score (NRS).

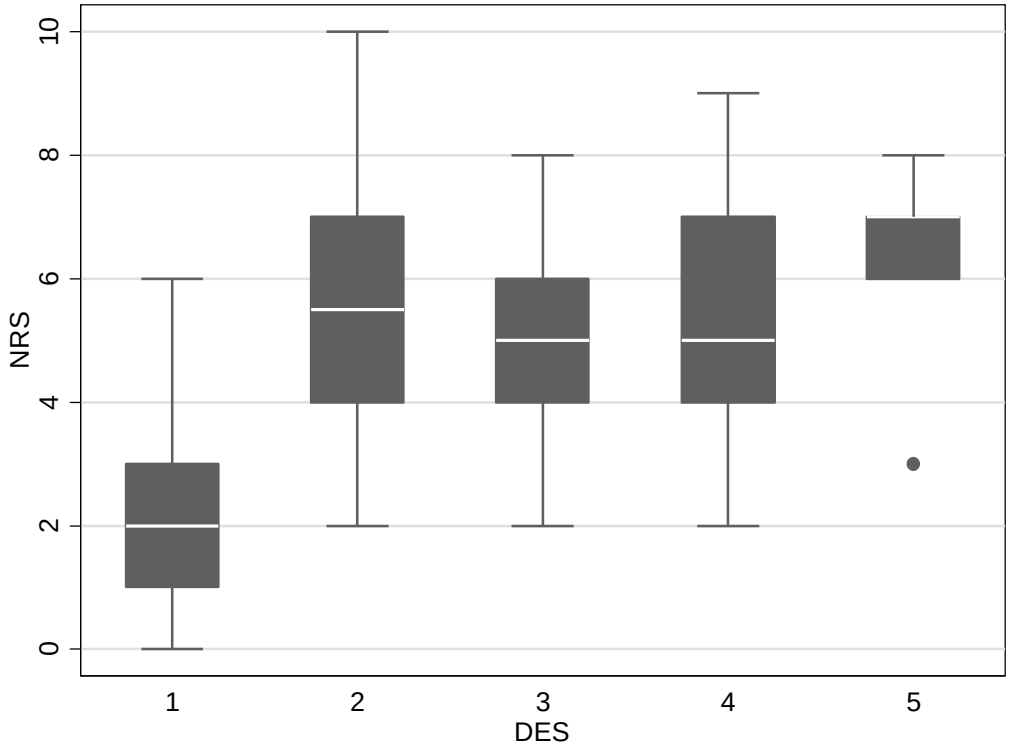

Figure 2b: Boxplot showing the distribution of modified Medical Research Council (mMRC) per Numerical Rating Scale score (NRS).

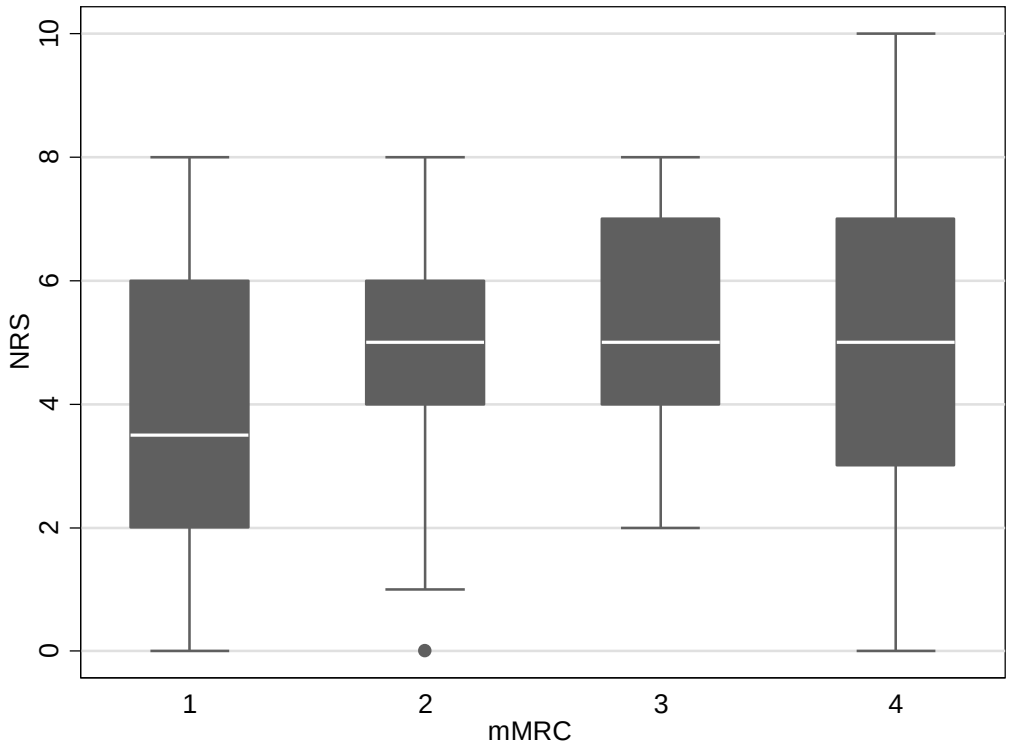

Figure 3a: Change in Dyspnea Exertion Scale (DES) plotted against change in Numerical Rating Scale (NRS) of breathlessness.

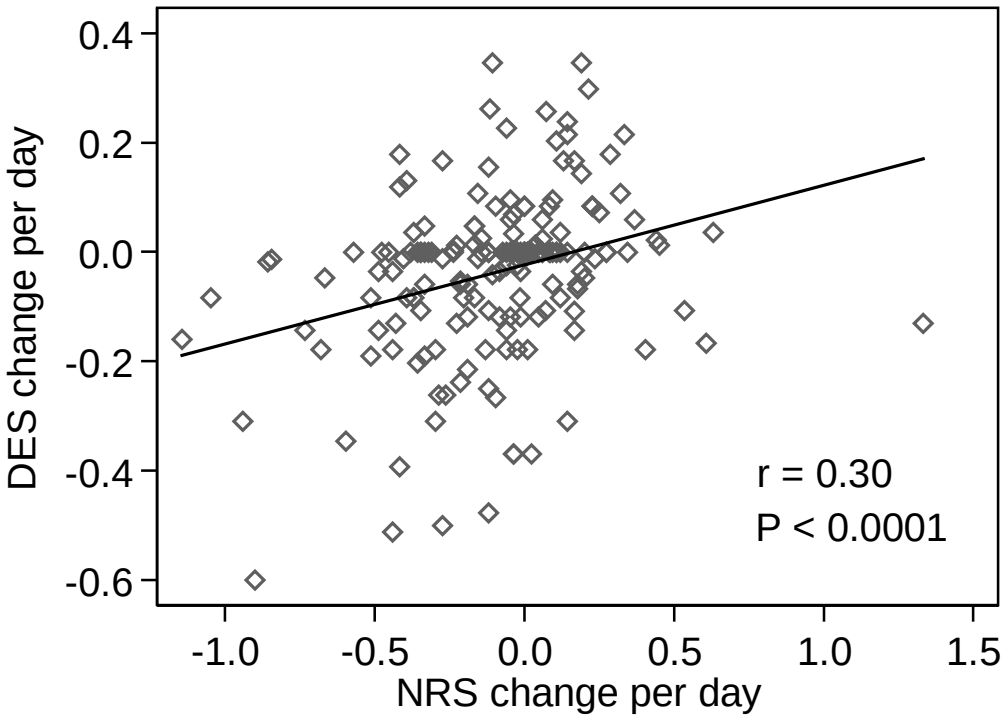

Figure 3b: Change in modified Medical Research Council (mMRC) plotted against change in Numerical Rating Scale (NRS) of breathlessness.