



Low agreement between mMRC rated by patients and clinicians: implications for practice

To the Editor:

Chronic breathlessness [1] causes immense suffering in cardiorespiratory diseases. The functional impact of activity-related breathlessness, measured on the modified Medical Research Council (mMRC) scale [2], is highly prognostic, informs disease evaluation and management, including in chronic obstructive pulmonary disease (COPD) [3], and is widely used for determining eligibility in clinical trials.

In clinical practice, mMRC is often rated by physicians based on the patient's medical history. It is not known to what extent mMRC ratings differ when administered by clinicians compared with patient self-report. The ratings may be influenced by other clinical characteristics, such as the patient's functional status. The New York Heart Association scale, which is similar to mMRC and is key for management of heart failure, is associated with functional status, measured using the Australia-modified Karnofsky Performance Status (AKPS) [4], but discriminates poorly between clinically important performance states in people with advanced disease [4].

The primary aim of this study was to evaluate the agreement between clinician- and patient-reported mMRC scores. Secondary aims were to evaluate whether the agreement differed by severity of activity-related breathlessness and how clinician and patient ratings correlated to the patient's functional status.

This was a pooled analysis of two randomised, placebo-controlled trials of morphine [5] and sertraline [6] for chronic breathlessness. Only data at screening and baseline were used (before any study treatment was initiated). Patients had severe life-limiting illnesses and chronic breathlessness defined as a clinician-rated mMRC ≥ 2 at screening despite optimal treatment for the underlying cause(s), as detailed elsewhere [5, 6]. Participants with missing data on clinician- or patient-reported mMRC ($n=68$) were excluded. No data were imputed.

mMRC was rated by clinicians at screening and was then self-reported by patients in their study diary at baseline (before randomisation). Patients' functional status was rated by clinicians at baseline using AKPS [7]. The primary analysis compared clinician and patient mMRC ratings conducted within 3 days. A sensitivity analysis was performed using ratings performed ≥ 4 days apart. Agreement was analysed using quadratic-weighted Cohen's κ , categorised according to LANDIS and KOCH [8]: 0=no (chance) agreement; 0.01–0.2=slight; 0.21–0.40=fair; 0.41–0.60=moderate; 0.61–0.80=substantial; ≥ 0.81 =high agreement. Associations between the mMRC ratings and patients' functional status (AKPS) were analysed using Kendall's τ . The study was approved by relevant human research ethics committees and all participants provided written, informed consent. Statistical analyses were conducted using the Statistical Package for the Social Sciences (SPSS) software (version 24.0; IBM Corporation, Armonk, NY, USA).

In total, 464 patients (294 from the morphine trial and 170 from the sertraline trial) had paired clinician and patient mMRC ratings. The time between clinician and patient mMRC ratings was a median 0 (interquartile range $-1-0$) days; 312 patients had ratings within 3 days (by 42 physicians) and were included in the primary analysis: mean \pm SD age was 73.8 \pm 8.8; 63.5% were male; the most common



@ERSpublications

Chronic activity-related breathlessness often assessed on the mMRC scale has low agreement between clinicians and patients. Self-ratings are important when available and improved objective assessment methods are emerging and should be used. <http://bit.ly/314wkVS>

Cite this article as: Ekström M, Chang S, Johnson MJ, *et al.* Low agreement between mMRC rated by patients and clinicians: implications for practice. *Eur Respir J* 2019; 54: 1901517 [<https://doi.org/10.1183/13993003.01517-2019>].

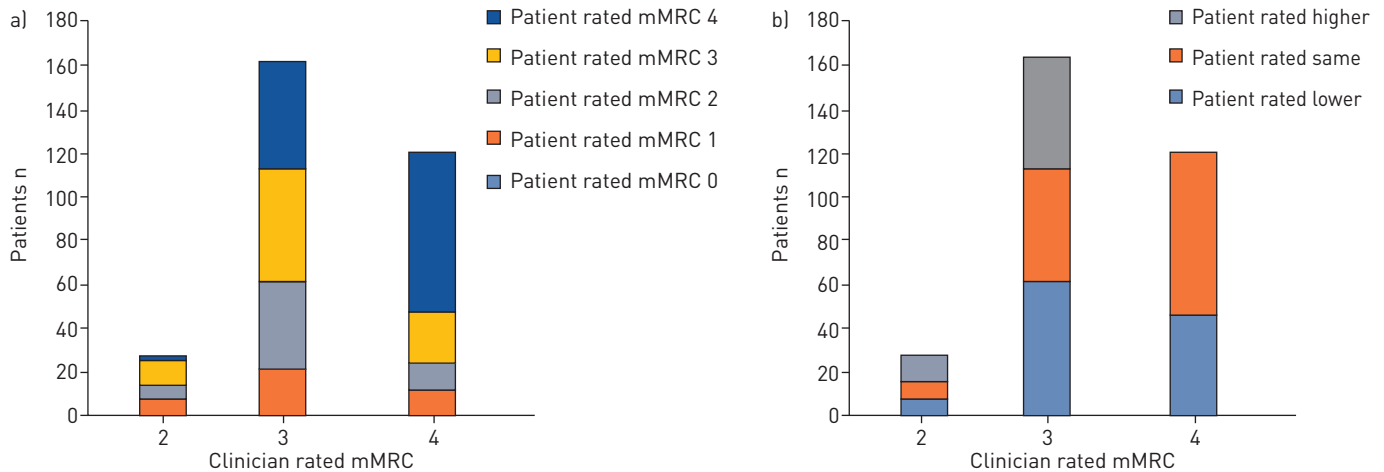


FIGURE 1 Agreement between a) clinician- and patient-rated modified Medical Research Council (mMRC) breathlessness scores; b) distribution of lower, similar and higher patient *versus* clinician ratings. Agreement was relatively low between patient- and clinician-rated mMRC, with an even distribution of under- and over-ratings for mMRC 2–3.

diagnoses were COPD (70.5%), interstitial lung disease (17.3%), lung cancer (13.8%) and heart failure (4.8%); and patients were ambulatory with a mean \pm SD AKPS of 61.5 \pm 10.1. Characteristics were similar between patients who were included and excluded from the primary analysis.

Agreement between clinician- and patient-reported mMRC (scored within 3 days; n=312) is shown in figure 1. The ratings differed considerably and the agreement for all categories was slight to fair, Cohen's κ 0.238 (95% CI 0.143–0.326). The rate of under- and overestimation by clinicians compared to patient self-reports was similar (figure 1). However, agreement was better for higher mMRC scores (25% for clinician mMRC 2, 31% for mMRC 3 and 61% for mMRC 4; $p<0.001$ using Mantel–Haenszel Chi-squared test). Functional status was more closely related to clinician-rated mMRC ($\tau=-0.42$; $p<0.001$) than patient-rated mMRC ($\tau=-0.22$; $p<0.001$). For scores >3 days apart (n=152), agreement was slightly lower (Cohen's κ 0.154, 95% CI 0.047–0.260), but findings were otherwise similar.

This study evaluated for the first time the agreement between clinician- and patient-rated mMRC. The main finding was that only a minority of ratings agreed, with similar rates of clinician under- and overestimation. These findings are consistent with reported disagreement between clinician and patient ratings of subjective measures including symptom intensity [9] and quality of life [10]. Our study is the first indication of substantial disagreement between clinicians and patients when assessing even a relatively objective measure, such as when breathlessness limits exertion.

Secondly, a novel finding was that as activity-related breathlessness worsened, agreement between patients and their clinicians improved. The subjective symptom of breathlessness might be under-detected by the clinician until becomes visible as a “clinical sign” of reduced function. Functional status was more closely related to clinician-rated than patient-rated mMRC. This could reflect that patients reduce or avoid physical activities to limit their breathing discomfort, which could lead to patients underestimating their activity-related breathlessness (as they become more inactive), contributing to symptom under-report. In addition, clinicians may incorporate other clinical information when rating breathlessness such as the patient's disease severity and functional status. In fact, this could make the clinician ratings even more predictive than the self-report of future clinical outcomes, which should be evaluated in studies with long-term outcome data.

A strength of the analysis was the large sample of patients with chronic breathlessness, with ratings using standardised scales in the setting of randomised controlled trials. A potential limitation was the time between the ratings, hence the primary analysis included ratings done within 3 days. Given that mMRC only has five levels that are quite broad and the chronicity of breathlessness in the study population, mMRC scores should be stable within time periods >3 days. As a clinician rated mMRC of 2–4 was an eligibility criterion, findings pertain mostly to moderate-to-severe chronic breathlessness. The improved agreement for higher mMRC scores might be partially related to getting closer to the upper limit of the scale. Higher agreement might also be found near the lower limit (mMRC 0–1), giving a U-shaped agreement for mMRC, which should be further explored. There were no data on how each clinician established a patient's mMRC. Involvement of patients in the clinician rating is possible, but would in fact make their scores more similar and thus overestimate the agreement.

The low agreement between clinician- and patient-rated mMRC has direct clinical implications, as mMRC is widely used to assess disease severity and prognosis, guide patient management, and select participants for interventional symptom trials [3, 11]. The findings highlight that activity-related breathlessness is a subjective experience that is only weakly related to other commonly measured clinical parameters (including functional status), and that symptom assessment should include self-report whenever possible to accurately capture patients' experiences [1, 12]. At the same time, given the complexity of chronic progressive diseases, comorbidities and symptoms, assessment necessitates a clinician's involvement, which may also mitigate symptom under-reporting by patients. Training of clinicians to adequately assess breathlessness and gain a better proxy mMRC where self-report is not possible would give more accurate representation of patient status, which is important in cardiorespiratory disease.

Improved methods to assess exertional breathlessness are needed for use in clinical care, for selecting participants to clinical trials and to measure treatment effects. The mMRC might under-report symptoms in patients with milder disease and who have become less active due to breathlessness [13], and is too unresponsive to detect change. Standardised tests for measuring changes in activity-related breathlessness have been validated in COPD [14, 15].

Magnus Ekström^{1,2}, Sungwon Chang², Miriam J. Johnson^{2,3}, Belinda Fazekas², Slavica Kochovska², Chao Huang³ and David C. Currow^{2,3}

¹Dept of Clinical Sciences, Division of Respiratory Medicine and Allergy, Lund University, Lund, Sweden.

²IMPACCT, Faculty of Health, University of Technology Sydney, Ultimo, Australia. ³Wolfson Palliative Care Research Centre, Hull York Medical School, University of Hull, Hull, UK.

Correspondence: Magnus Ekström, Dept of Medicine, Blekinge Hospital, SE-371 95 Karlskrona, Sweden. Email: pmekstrom@gmail.com

Received: 31 July 2019 | Accepted after revision: 06 Oct 2019

Conflict of interest: M. Ekström has nothing to disclose. S. Chang has nothing to disclose. M.J. Johnson has nothing to disclose. B. Fazekas has nothing to disclose. S. Kochovska has nothing to disclose. C. Huang has nothing to disclose. D. C. Currow received an unrestricted research grant from Mundipharma, is an unpaid member of an advisory board for Helsinn Pharmaceuticals, and has consulted Specialist Therapeutics and Mayne Pharma and received intellectual property payments from Mayne Pharma.

Support statement: This work was funded by the Dept of Health and Ageing, Australian Government; Svenska Sällskapet för Medicinsk Forskning; and National Health and Medical Research Council, grant 1010065. Funding information for this article has been deposited with the Crossref Funder Registry.

References

- 1 Johnson MJ, Yorke J, Hansen-Flaschen J, *et al.* Towards an expert consensus to delineate a clinical syndrome of chronic breathlessness. *Eur Respir J* 2017; 49: 1602277.
- 2 Bestall JC, Paul EA, Garrod R, *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: 581–586.
- 3 Global Initiative for Chronic Obstructive Lung Disease. Global Strategy for the Diagnosis, Management and Prevention of COPD. 2019. <http://goldcopd.org/> Date last accessed: March 15, 2019.
- 4 Johnson MJ, Bland JM, Davidson PM, *et al.* The relationship between two performance scales: New York Heart Association Classification and Karnofsky Performance Status Scale. *J Pain Symptom Manage* 2014; 47: 652–658.
- 5 Ekström M, Williams M, Johnson MJ, *et al.* Agreement between breathlessness severity and unpleasantness in people with chronic breathlessness: a longitudinal clinical study. *J Pain Symptom Manage* 2019; 57: 715–723.
- 6 Currow DC, Ekström M, Louw S, *et al.* Sertraline in symptomatic chronic breathlessness: a double blind, randomised trial. *Eur Respir J* 2019; 53: 1801270.
- 7 Abernethy A, Shelby-James T, Fazekas B, *et al.* The Australia-modified Karnofsky Performance Status (AKPS) scale: a revised scale for contemporary palliative care clinical practice [ISRCTN81117481]. *BMC Palliat Care* 2005; 4: 7.
- 8 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159–174.
- 9 Simon ST, Altfelder N, Alt-Epping B, *et al.* Is breathlessness what the professional says it is? Analysis of patient and professionals' assessments from a German nationwide register. *Support Care Cancer* 2014; 22: 1825–1832.
- 10 Bruera E. Poor correlation between physician and patient assessment of quality of life in palliative care. *Nat Clin Pract Oncol* 2006; 3: 592–593.
- 11 Ponikowski P, Voors AA, Anker SD, *et al.* 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). *Eur J Heart Fail* 2016; 18: 891–975.
- 12 Parshall MB, Schwartzstein RM, Adams L, *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: 435–452.
- 13 Soumagne T, Laveneziana P, Veil-Picard M, *et al.* Asymptomatic subjects with airway obstruction have significant impairment at exercise. *Thorax* 2016; 71: 804–811.
- 14 Maltais F, Aumann J-L, Kirsten A-M, *et al.* Dual bronchodilation with tiotropium/olodaterol further reduces activity-related breathlessness versus tiotropium alone in COPD. *Eur Respir J* 2019; 53: 1802049.
- 15 Borel B, Wilkinson-Maitland CA, Hamilton A, *et al.* Three-minute constant rate step test for detecting exertional dyspnea relief after bronchodilation in COPD. *Int J Chron Obstruct Pulmon Dis* 2016; 11: 2991–3000.