Comparisons of health status scores with MRC grades in a primary care COPD population: Implications for the new GOLD 2011 classification

P.W. Jones, L. Adamek, G Nadeau, N. Banik

1Division of Clinical Science, St George’s, University of London, UK
2Respiratory Centre of Excellence, GlaxoSmithKline, Uxbridge, United Kingdom
3Biostatistics and Epidemiology, GlaxoSmithKline, Munich, Germany

Correspondence: P.W. Jones. Division of Clinical Science, St George’s University of London, London, UK. E-mail: pjones@sghms.ac.uk
ABSTRACT

The 2011 GOLD strategy document recommends COPD assessment using symptoms and future exacerbation risk, employing two score cut-points: COPD Assessment Test (CAT) score $\geq 10$ or modified Medical Research Council Dyspnoea (mMRC) score $\geq 2$. To explore the equivalence of these two symptom cut-points, the relationship between CAT and mMRC scores, and St George’s Respiratory Questionnaire [SGRQ], the short form health survey and the Functional Assessment of Chronic Illness Therapy Fatigue scores were retrospectively analysed using a primary care dataset.

Data from 1817 patients (mean±SD FEV$_1$: 1.6±0.6L) showed a significant association between mMRC and all health status scores (ANOVA, p<0.0001). mMRC Grade 1 was associated with significant levels of health status impairment (SGRQ 39.4±15.5; CAT 15.7±7.0); even patients with mMRC Grade 0 had modestly elevated scores (SGRQ 28.5±15.1; CAT 11.7±6.8). mMRC $\geq 2$ categorised 57.2% patients with low symptom (Groups A and C) versus 17.2% with the CAT. Using mMRC cut-point ($\geq 1$) resulted in similar GOLD group categorisations as the CAT (18.9%).

The mMRC showed a clear relationship with health status scores; even low mMRC grades were associated with health status impairment. Cut-points of mMRC $\geq 1$ and CAT $\geq 10$ were approximately equivalent in determining low-symptom patients. The GOLD assessment framework may require refinement.

Keywords

Chronic obstructive pulmonary disease; New GOLD classification; COPD Assessment Test; modified Medical Council Research Scale; health-related quality of life; primary care
INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is a common chronic inflammatory disease, characterised by persistent airflow limitation, which is both preventable and treatable [1]. In individual patients, the overall severity of the disease is influenced by exacerbations and comorbidities [1]. Until recently, spirometry remained the standard method for grading COPD severity [2]. However, it is now recognised that, at the individual patient level, forced expiratory volume in one second (FEV₁) is an unreliable marker of the severity of breathlessness, exercise limitation and health status impairment [3]. The 2011 Global Initiative for Chronic Obstructive Lung Disease (GOLD) strategy document now recommends that COPD management and treatment should consider both disease impact, determined by assessment of symptoms and activity limitation, and future risk of exacerbations, determined from airflow limitation or exacerbation history [1]. This combined assessment of COPD results in the grouping of patients into one of four categories: A: low risk, less symptoms; B: low risk, more symptoms; C: high risk, less symptoms; D: high risk, more symptoms.

GOLD recommends either the COPD Assessment Test (CAT) or the modified Medical Research Council (mMRC) dyspnoea scale for assessing symptoms. The CAT is a short, 8-item, health status questionnaire developed to provide a simple tool for quantifying the symptoms and impacts of COPD [4]. It has been shown to distinguish between different severities of COPD and is highly correlated with the St George’s Respiratory Questionnaire (SGRQ) [5]. The original MRC dyspnoea scale was described by Fletcher in 1952 and was developed to help physicians establish clinical grades of breathlessness (five grades) for their patients with emphysema, based on their ability to perform physical activities [6]. A modified version of this scale is used today (mMRC) which has more simplified statements and refers to ‘people’ instead of ‘men’, but is based on a similar five stages of breathlessness due to physical exertion [7]. [Note: the original MRC Grades ranged 1-5, confusingly the mMRC grades patients 0-4]. Bestall et al showed the MRC grade to be
significantly associated with shuttle distance, SGRQ scores, Chronic Respiratory Questionnaire scores and The Nottingham Extended Activities of Daily Living scores [8]. However, this study only evaluated patients with MRC grades 3 to 5 and not patients with milder disease (MRC grades 1 and 2). In the absence of other evidence, GOLD recommended a CAT ≥ 10 and mMRC ≥ 2 as equivalent symptom cut-points for categorising patients into low or high symptom groups.

The primary aim of this analysis was to test the equivalence, assumed by GOLD, between these two symptom cut-points, by examining the relationship between health status scores and mMRC grades in patients with COPD in patients with a wide range of severity, using the Health-Related Quality of Life in COPD in Europe Study (HEED) database [10]. A secondary aim was to compare the demographic and clinical characteristics, and health status scores of patients grouped using the new GOLD combined assessment framework.

METHODS

The HEED study was a large, cross-sectional, observational study conducted in primary care, full details of which have been previously published. [9] Patients visiting their primary care physician (PCP) for any reason and with a diagnosis of COPD (aged 40-80 years; post-bronchodilator forced expiratory volume in 1 s [FEV₁]/forced vital capacity [FVC] ratio of <70%; smoking pack history of at least 10 pack years) were invited to participate.

At a single study visit, patients completed the following health-related quality of life questionnaires: COPD specific St George’s Respiratory Questionnaire (SGRQ-C) [10]; the Short Form Health Survey (SF-12) [11]; the CAT [4]; the Functional Assessment of Chronic Illness Therapy (FACIT) fatigue scale [12]. Dyspnoea was assessed using the mMRC [7]. The number of patient-reported exacerbations in the last 6 months was recorded. An exacerbation was defined as a worsening of symptoms that required oral corticosteroids and/or antibiotics and/or hospitalisation. Post-bronchodilator FEV₁, FVC and FEV₁/FVC ratio
were also recorded; spirometry was either performed at the study visit or from documented evidence within the previous 6 months.

**Statistical analysis**

Sample size calculations have been described in an earlier publication [9]. Descriptive statistics using Statistical Analysis Systems version 9.1.3 software (SAS Inc, Cary, USA), were used to report mean health status scores in each of the mMRC grades. A one-way analysis of variance (ANOVA) was applied to test the association between mMRC grade and each health status score.

Descriptive statistics were used to describe the demographic and baseline characteristics of the patients within each GOLD assessment category. Data on exacerbations over the preceding 6 months were collected under three headings: 1) exacerbations treated with antibiotics; 2) exacerbations treated with oral corticosteroids (OCS); 3) exacerbations requiring hospitalisation). Since the total exacerbation number was not collected in an individual, the sum of all three exacerbation categories was used and annualised. Using this approach, patients with no exacerbations in 6 months were classified into GOLD categories A and B. Patients with \( \geq 1 \) exacerbation in the preceding 6 months had an extrapolated exacerbation rate of \( \geq 2 \), so were categorised into groups C or D.

A concordance analysis was applied to quantify the association between the frequencies of patients categorised into GOLD groups A to D using symptom cut-points CAT \( \geq 10 \) versus mMRC \( \geq 2 \) and CAT \( \geq 10 \) versus mMRC \( \geq 1 \), respectively. A weighted Kappa coefficient was calculated (a value of 1 indicates a perfect agreement).

**RESULTS**

**Study Population**

The demographic and clinical characteristics of the reported population have been described previously [9]. The mean age was 65 years and mean duration of COPD was 9 years. The mean FEV\(_1\) was 1.6L (57% of predicted) and over 40% were current smokers. In the
previous 6 months, just over half of all patients had had an exacerbation requiring antibiotics, 35% had an exacerbation requiring OCS and 10% had been hospitalised due to an exacerbation.

**Clinical characteristics of patients by new GOLD classification**

The demographic and clinical characteristics of patients split by the new GOLD Groups, classified by CAT ≥10 or MRC ≥2 cut points are shown in Table 1. The mean age of patients was similar across GOLD groups, whether classified by CAT or mMRC and there was no apparent difference in mean number of comorbidities across GOLD groups by either classification. The highest level of COPD symptoms (cough, sputum and dyspnoea) and lowest level of FEV₁ were observed in Group D, regardless of classification method. A higher proportion of Group A patients, as classified by the mMRC, reported respiratory symptoms compared to those placed in this category by the CAT; this was also true for patients in Group B (symptoms of sputum and cough) and Group C (all symptoms).

**Health status scores by new GOLD classification**

Patients in all four groups, as categorised by the mMRC ≥2 cut-point, had worse health status scores and more fatigue (encompassing tiredness, weakness, and difficulty in conducting usual activities) compared with the equivalent group categorised by CAT ≥10 (Table 2). The differences in scores between these two categorisation methods groups exceeded the minimum clinically important difference (MCID) for each questionnaire (13, 14; Quality Metric Inc, private communication), with the exception of SF-12 PCS scores in Group D. For patients categorised by mMRC as having low symptoms (Groups A and C), SGRQ scores were much higher (by approximately 3 times the MCID) than those categorised by CAT.

**Relationship between health status scores and mMRC grade**

The results of the one-way ANOVA to test the association of health status scores with mMRC grade showed a significant relationship between mMRC grade and SGRQ, SF-12 PCS, CAT and FACIT scores (p<0.0001 for all) (Figure 1). The differences in SGRQ and FACIT scores between mMRC grades 0-1, 1-2, 2-3 and 3-4 all exceeded the MCIDs for
these questionnaires (13, 14). The differences in SF-12 PCS score also exceeded its MCID between mMRC grades 0-1, 1-2 and 2-3, but not between grades 3-4. Notably, even patients with mMRC grade 0 (‘only breathless with strenuous exercise’) had modestly elevated CAT and SGRQ scores, and a mean SF-12 score of 44.5, which is below the normalised score of 50 expected of the general population (11, 15). For patients with mMRC grade 1 (‘breathless when hurrying on level or walking up a slight hill’) or above, significant levels of health status impairment were observed.

**Proportion of patients in each GOLD group using CAT ≥ 10 and mMRC ≥ 2 symptom cut-points (GOLD recommendation)**

The proportion of patients categorised into groups A to D differed according to the use of a GOLD symptom cut-point of mMRC ≥ 2 (Figure 2a) or CAT ≥10 (Figure 2b). The mMRC classified 57.2% patients as having low symptoms (Groups A and C) compared with 17.2% with the CAT. The concordance analysis comparing the frequencies of patients categorised by these two variables, the mMRC ≥ 2 and CAT ≥10 cut-points, resulted in a weighted Kappa coefficient of 0.626.

**Proportion of patients in each GOLD group using CAT ≥ 10 and mMRC ≥ 1 symptom cut-points (amended criteria)**

As an exploratory exercise, health status scores were investigated when patients were categorised using an mMRC cut-point of ≥1. Using this cut-point, patients categorised by mMRC ≥1 had similar mean health status and fatigue scores to those found with a CAT cut-point of ≥10 (Table 2). Using this cut-point, the proportion of patients categorised in to groups A to D was similar to that with the CAT (Figure 2c). The mMRC ≥ 1 classified 18.9% of patients as having low symptoms (groups A and C) compared with 17.2% with the CAT. The concordance analysis comparing the frequencies of patients categorised by these two variables, the mMRC ≥ 1 and CAT ≥10 cut-points, resulted in a weighted Kappa coefficient of 0.792; indicating that there was a higher degree of agreement between these variables and their respective cut-points than between the classification applying CAT ≥10 and mMRC ≥2.
DISCUSSION

The new GOLD assessment system incorporates symptoms and risk of exacerbations, in addition to an assessment of airflow limitation (1). Two alternative symptomatic cut-points are recommended for use. In the absence of published evidence, mMRC ≥2 and CAT ≥10 were thought to be equivalent. This analysis has shown that there is a clear relationship between mMRC grade and health status scores across the whole severity range of the mMRC scale, which supports and extends the findings of Bestall et al [8]. However, we also show that an mMRC score ≥2 appears not to be equivalent to a CAT score ≥10. Patients categorised as having low symptoms using this mMRC cut-point have significantly worse health status scores, and experienced more fatigue, than patients categorised using the CAT with a cut-point of ≥10. In this patient population (mean FEV₁ of 57% predicted) mMRC ≥2 categorised the majority of patients as having low symptoms (i.e. they were placed in GOLD Groups A and C). The SGRQ scores in Group A and C patients were much higher than those observed in COPD-diagnosed patients identified in a population study by Ferrer et al [16] and were very similar to the average SGRQ scores in COPD patients identified in another prevalence study [17]. By contrast, use of the CAT≥10 cut-point placed a smaller proportion of patients in the low symptom groups, although it should be noted that even these patients still had significant health status impairment - as judged by SGRQ and SF-12.

The MRC dyspnoea scale was developed as a standard set of questions by experts in the field of chronic bronchitis and emphysema for measuring dyspnoea. Fletcher reported that the use of the questions enabled different observers to get reasonably repeatable results and that there was general agreement between the answers to these questions and an objective measure of dyspnoea [6]. It should be noted that most MRC Grades contain two different activities, but neither the validity of such combinations nor the equivalence of the two combined activities has ever been tested. mMRC Grade 1 describes a patient who reports being breathless when hurrying on the level or when walking up a slight hill. From a
COPD perspective that might reflect a relatively less symptomatic patient than others with the condition; but from a broader perspective, outside the setting of COPD, being breathless when hurrying on the level indicates a significant level of symptoms that has an impact on normal daily activities. In this study, the mean SGRQ score in mMRC Grade 1 patients was 39.4±15.5, which lies within 0.5 standard deviation of the mean score in patients recruited to recent clinical trials [18, 19]. Patients with mMRC Grade 1 are clearly symptomatic as judged by their health status scores.

The CAT was developed in a very different way from the MRC; it focuses on all aspects of COPD and was developed following a rigorous selection of the items based on interviews and focus groups with COPD patients, supported by interviews with community physicians and pulmonologists [4]. The primary objective was to create a questionnaire made up of the smallest number of items that formed a unidimensional instrument with reliable measurement properties. Its testing was robust at each stage of development. The CAT correlates very well with the SGRQ [5] and a Bland and Altman plot shows that the two instruments perform in a very similar way across the scaling range of both instruments [20]. A CAT score $\geq 10$ has been shown to have a significant impact on the daily lives of patients with COPD [20]. Patients with a CAT score of 10 or above are likely to be breathless on most days, get exhausted easily and take a long time to do housework. Higher CAT scores are associated with greater disease impacts.

An unexpected finding was that 15% of the patients in Group B (i.e. with mMRC$\geq 2$) recorded no breathlessness using a simple self-reported measure of respiratory symptoms (Table 1). This observation lends support to the use of standardised instruments such as the mMRC and CAT for assessing symptomatic effects of COPD, rather relying on the patients’ response to global questions about symptoms and their severity.

The intention of the new GOLD assessment framework is for it to be used in clinical settings, so physicians need be assured that the suggested cut-points are correct and meaningful.
Misclassification of patients could potentially affect their future management and treatment. This analysis provides the first evidence-based approach to choosing the correct cut-points. Following our observation of a lack of concordance between CAT≥10 and mMRC≥2, we carried out an exploratory analysis using different mMRC cut-points for the GOLD symptomatic classification. This was based on our observation that the mean CAT score in mMRC Grade 0 patients was 11.7±6.8. The characteristics of patients categorised using an mMRC ≥1 cut-point matched those categorised by CAT≥10 much more closely than the GOLD mMRC ≥2 cut-point. The distribution of the patients into the four GOLD categories by CAT≥10 and mMRC ≥1 was very similar. The association between the classifications using a CAT≥10 and the mMRC is stronger (weighted kappa coefficient closer to 1) when applying a cut-off of 1 instead of 2 for the mMRC-based classification. This does not mean that exactly the same patients were placed into the different categories using CAT and mMRC≥1, but it shows that the characteristics of the patients so categorised was similar. The differences in classification of patients using CAT or MRC ≥ 1 would influence treatment in only a small proportion of patients. By contrast, differences in the classification of patients using CAT or MRC ≥ 2 show that, using this mMRC cut-point, a sizeable proportion of patients with symptoms that have a significant impact upon their daily activity would not be prescribed maintenance treatment from which they would likely derive some benefit. These data suggest that the new GOLD assessment framework symptom cut-points may require some future modification with respect to MRC.

One of the aims of this analysis was to use a large COPD dataset to examine the characteristics of patients grouped using the new GOLD combined assessment framework in a typical clinical practice setting. Overall, there were no marked differences in demographic characteristics when patients were categorised using the mMRC or the CAT. The overall level of comorbidities and cardiovascular comorbidities were similar across GOLD groups using either classification, which agrees with the findings for numbers of comorbidities split by the GOLD spirometric staging system [9].
A limitation of this analysis is that the individual total numbers of exacerbations over 12 months were not directly available and were derived from the sum of the numbers of exacerbations treated with antibiotics and/or treated with oral corticosteroids and/or leading to hospitalization and therefore may have over-estimated the true number of total exacerbations. Also, spirometry was not performed in a standardised way across centres but fulfilled local requirements at each site. This study was conducted in primary care in patients with confirmed COPD and therefore may not be applicable to the total COPD population and does not represent patients with undiagnosed disease. A further limitation was that this was a cross-sectional study and we cannot speculate about movement across category boundaries over time.

The new GOLD combined COPD assessment provides a much-needed framework for measuring the impact of COPD in terms of current symptoms and future exacerbation risk. This analysis showed a clear relationship between mMRC and other health status scores, and that even patients with low mMRC grades have perceptible health status impairment. However, the current mMRC cut-point of ≥ 2 appears to over-estimate the proportion of patients with low symptoms; these data may provide evidence for the future modification of the group cut-off points with respect to mMRC scores.

**Acknowledgments**

Funding for this study was provided by GlaxoSmithKline (GSK) (Study number: 111749). The authors would also like to acknowledge editorial support in the form of development of the manuscript first draft, assembling tables and figures, collating author comments, copyediting, and referencing which was provided by Kate Hollingworth of Continuous Improvement Ltd. This support was funded by GSK.
REFERENCES


2. The Global Initiative for Chronic Obstructive Lung disease (GOLD) revised 2010.


Table 1: Demography and clinical data by GOLD group using CAT ≥10 and mMRC ≥ 2 symptom cut-points, for the reported COPD population

<table>
<thead>
<tr>
<th>CAT ≥ 10</th>
<th>mMRC ≥ 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
</tr>
<tr>
<td>N (% of total population)</td>
<td>147 (8.1)</td>
</tr>
<tr>
<td>Males, n (%)</td>
<td>114 (77.6)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>66.3±8.6</td>
</tr>
<tr>
<td>COPD duration (years)</td>
<td>6.5±5.3</td>
</tr>
<tr>
<td>Smoking pack years</td>
<td>42.4±25.0</td>
</tr>
<tr>
<td>Current smoker, n (%)</td>
<td>52 (35.4)</td>
</tr>
<tr>
<td>Number of comorbidities</td>
<td>1.6±1.3</td>
</tr>
<tr>
<td>Number of cardiovascular comorbidities</td>
<td>0.8±0.9</td>
</tr>
<tr>
<td>Exacerbations per patient in last 6 months requiring</td>
<td></td>
</tr>
<tr>
<td>Antibiotics</td>
<td>0</td>
</tr>
<tr>
<td>Oral corticosteroids</td>
<td>0</td>
</tr>
<tr>
<td>Hospitalisation</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
</tr>
<tr>
<td>COPD symptoms on study day, n (%)</td>
<td></td>
</tr>
<tr>
<td>Cough</td>
<td>73 (49.7)</td>
</tr>
<tr>
<td>Sputum</td>
<td>61 (41.5)</td>
</tr>
<tr>
<td>Dyspnoea</td>
<td>54 (36.7)</td>
</tr>
<tr>
<td>FEV₁ (L)</td>
<td>1.9±0.5</td>
</tr>
<tr>
<td>FEV₁ % of predicted</td>
<td>70.7±15.5</td>
</tr>
</tbody>
</table>

Data presented as mean±SD unless otherwise stated. GOLD patient group definitions: A: low risk, less symptoms; B: low risk, more symptoms; C: high risk, less symptoms; D: high risk, more symptoms

In last 6 months
Table 2: Health status scores by GOLD group using CAT ≥ 10 and mMRC ≥ 2 symptom cut-points and an exploratory cut-point using mMRC ≥ 1, for the reported COPD population

<table>
<thead>
<tr>
<th>Patient group</th>
<th>mMRC ≥ 2</th>
<th>CAT ≥ 10</th>
<th>mMRC ≥ 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SGRQ total score(^1)</td>
<td>SF-12 PCS(^2)</td>
<td>FACIT-Fatigue(^3)</td>
</tr>
<tr>
<td>A</td>
<td>30.5 (28.8, 32.2)</td>
<td>43.2 (42.3, 44.0)</td>
<td>40.5 (39.6, 41.5)</td>
</tr>
<tr>
<td>B</td>
<td>48.3 (45.2, 51.4)</td>
<td>34.8 (33.3, 36.2)</td>
<td>31.8 (29.9, 33.7)</td>
</tr>
<tr>
<td>C</td>
<td>38.6 (37.3, 39.9)</td>
<td>40.3 (39.7, 41.0)</td>
<td>38.0 (37.3, 38.7)</td>
</tr>
<tr>
<td>D</td>
<td>58.1 (56.7, 59.4)</td>
<td>32.3 (31.7, 32.9)</td>
<td>27.9 (27.0, 28.8)</td>
</tr>
</tbody>
</table>

Data presented as mean (95% confidence interval). mMRC: modified Medical Research Council; CAT: COPD Assessment Test; SGRQ: St George’s Respiratory Questionnaire; SF-12 PCS: Short Form health survey physical component score; FACIT-Fatigue: Functional Assessment of Chronic Illness Therapy (FACIT) Fatigue scale

\(^1\)Reported as SGRQ scores following transformation from SGRQ-C scores; SGRQ total score: a lower score represents a better QoL.

\(^2\)SF-12 scores: a higher score represents a better QoL.

\(^3\)FACIT-fatigue scores: higher score indicates less fatigue
Figures Legend

Figure 1: Health status scores by mMRC grade for the reported COPD population:

a) SGRQ,
b) SF12-PCS

![Graph showing SF12-PCS score vs mMRC grade]

mMRC grade

SF-12 PCS score

0 1 2 3 4

0 10 20 30 40 50 60 70 80 90 100

c) CAT

![Graph showing CAT score vs mMRC grade]

mMRC grade

CAT score

0 5 10 15 20 25 30 35 40

d) FACIT

Data shown as mean±SD. One-way ANOVA of the association between mMRC grade and health status scores (p<0.0001 for all)
Figure 2: Proportion of patients in each GOLD group using symptom cut-points: mMRC ≥ 2; CAT ≥ 10; mMRC ≥ 1 for the reported COPD population