International Journal of COPD

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

# Application of the new GOLD COPD staging system to a US primary care cohort, with comparison to physician and patient impressions of severity

This article was published in the following Dove Press journal:
International Journal of COPD
30 July 2015
Number of times this article has been viewed

Douglas W Mapel[1]
Anand A Dalal[2]
Phaedra T Johnson[3]
Laura K Becker[3]
Alyssa Goolsby Hunter[3]

[1]Epidemiology and Health Outcomes Research, Lovelace Clinic Foundation, Albuquerque, NM, [2]US Health Outcomes and Medical Policy, GlaxoSmithKline, Research Triangle Park, NC, [3]Life Sciences, Optum, Eden Prairie, MN, USA

**Background:** In 2011, the traditional Global Initiative for Chronic Obstructive Lung Disease (GOLD) COPD spirometry-based severity classification system was revised to also include exacerbation history and COPD Assessment Test (CAT) and modified Medical Research Council Dyspnea Scale (mMRC) scores. This study examined how COPD patients treated in primary care are reclassified by the new GOLD system compared to the traditional system, and each system's level of agreement with patient's or physician's severity assessments.

**Methods:** In this US multicenter cross-sectional study, COPD patients were recruited by 83 primary care practitioners (PCPs) to complete spirometry testing and a survey. Patients were classified by the traditional spirometry-based system (stages 1–4) and under the new system (grades A, B, C, D) using spirometry, exacerbation history, mMRC, and/or CAT results. Concordance between physician and patient-reported severity, spirometry stage, and ABCD grade based on either mMRC or CAT scores was examined.

**Results:** Data from 445 patients with spirometry-confirmed COPD were used. As compared to the traditional system, the GOLD mMRC system reclassifies 47% of patients, and GOLD CAT system reclassifies 41%, but the distributions are very different. The GOLD mMRC system resulted in relatively equal distributions by ABCD grade (33%, 22%, 19%, 26%, respectively), but the GOLD CAT system put most into either B or D groups (9%, 45%, 4%, and 42%). The addition of exacerbation history reclassified only 19 additional patients. Agreement between PCPs' severity rating or their patients' self-assessment and the new ABCD grade was very poor ($\kappa$=0.17 or less).

**Conclusion:** As compared to the traditional system, the GOLD 2011 multidimensional system reclassified nearly half of patients, but how they were reclassified varied greatly by whether the mMRC or CAT questionnaire was chosen. Either way, the new system had little correlation with the PCPs or their patients' impressions about the COPD severity.

**Keywords:** assessment, comorbidity, exacerbation, Global Initiative for Chronic Obstructive Lung Disease, spirometry, stratification

## Introduction

The new Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2011 system for COPD severity assessment added chronic symptoms and exacerbation history to the traditional system of rating the degree of airflow obstruction by spirometry. It has been studied in a variety of research cohorts, but its impact in primary care is uncertain.

In this analysis of 445 patients with spirometry proven COPD managed in primary care practices from across the US, we find that the new GOLD system does reclassify

Correspondence: Douglas W Mapel
Lovelace Clinic Foundation, 2309 Renard Place SE Suite 103, Albuquerque, NM 87106-4264, USA
Tel +1 505 235 1434
Email dmapel@comcast.net

substantial proportions of COPD patients as compared to just spirometry alone, but how they are reclassified varies greatly by which symptoms questionnaire is chosen. Furthermore, the new reclassifications do not have any better agreement with physician's or patient's own impressions about COPD severity than the traditional system.

Clinicians and clinical scientists are interested in assessing COPD severity to have objective measures of lung impairment, prognostic information, and parameters on which to guide treatment. Traditionally, COPD severity has been solely based on the degree of airflow obstruction, in terms of percent of predicted forced expiratory volume in 1 second ($FEV_1$), as measured by spirometry. The use of $FEV_1$ to classify the severity of COPD was used in Drs Charles Fletcher and Richard Peto's studies of the natural history of COPD in the West London cohort of a half-century ago,[1] and spirometry-based severity systems have proven to be valid predictors of survival in many COPD cohorts worldwide since then.[1–3] Traditional systems that use spirometry to describe COPD progression, such as that recommended by the GOLD Committee in its original 2001 guidelines, have also been shown in prospective studies to be predictive of a variety of other clinical outcomes including health-related quality of life,[4,5] and the risk for episodes of acute deterioration in lung function known as COPD exacerbations.[6,7]

However, traditional spirometry-based COPD severity scales capture only one dimension of respiratory impairment, airflow, and neglect the multiple health dimensions negatively impacted by COPD, such as chronic symptoms and comorbidities. The correlations between $FEV_1$ and most patient-reported clinical outcomes are not very strong, even in prospective studies comparing changes in lung function to symptoms scores or multi-dimensional measures.[8–10] In the average pulmonary clinic, it is not difficult to find patients who have "severe" COPD by spirometry but are minimally symptomatic, while others with only "moderate" airflow obstruction are functionally disabled by their lung disease. Traditional severity classification systems also do not address how to classify the large number of current and ex-smokers with restrictive spirometry characteristics, whose survival prognosis is at least as poor as those with moderate airflow obstruction.[11,12] The limits of using spirometry measures to define COPD are also highlighted by the persistent debate about the problem of potential over-diagnosis of COPD when the traditional definition of obstruction as an $FEV_1$/forced vital capacity (FVC) ratio of less than 0.70 is applied to older populations.[13]

In 2011 the GOLD Committee recommended a new COPD assessment system that combines spirometry testing with measures of chronic respiratory symptoms, namely, the COPD Assessment Test (CAT) or modified British Medical Research Council Dyspnea Scale (mMRC),[14,15] along with an estimation of the future risk for adverse outcomes, as determined by either the recent history of acute COPD exacerbations or the percent of predicted $FEV_1$.[16] The tiered treatment recommendations that were based on the old spirometry paradigm were also revised to correspond to the new paradigm. The goal of these changes was to improve the clinical assessment and management of COPD.[17] Since the introduction of the new GOLD assessment system there has been interest in understanding how it compares to the traditional spirometry-based staging system, but most studies to date have been conducted with COPD patients recruited from university specialty clinics or research cohorts enrolled in longitudinal studies.[18–29] Very few studies have been based on primary care COPD populations.[30] Understanding how the new GOLD COPD assessment system relates to the older spirometry-based severity system is a practical problem for primary care practitioners (PCPs) who need to be able to rate the severity of their patient's lung disease and communicate that to the patient and to other health care providers.[31]

The primary objective of this analysis is to examine in a primary-care-based cohort how COPD patients staged by the traditional GOLD spirometry-based severity system are reclassified by the new GOLD 2011 assessment systems. Because the history of exacerbations is an important component of the new GOLD system, the severity stages and assessment groups are further stratified by exacerbation history. We also examine how old and new classification systems align with patients' and their PCPs' perceptions of COPD severity.

# Methods

## Participants and recruitment strategy

This was a cross-sectional observational study of 899 COPD patients treated in individual primary care practices from across the US. A total of 95 PCPs (General Internal Medicine or Family Practice) were recruited to participate in the study, and 83 PCPs enrolled at least one patient. Their practice characteristics are described in an earlier report.[32] Investigators identified potential subjects in electronic records using a stratified random sampling approach (ie, selection of each nth patient) to ensure unbiased selection. Patients were recruited by sites using a scripted telephone call and/or mailed letter. Patients aged 40 or older with English language

ability and documented care for at least 1 year at the PCP's clinic were included in the study. Patients were excluded if they had conditions that contraindicated the forced expiratory maneuver needed for spirometry, or were unable to complete study procedures, or had participated in a clinical trial within the prior 12 months. For this analysis, we only included patients who met the American Thoracic Society (ATS) definition of spirometry proven COPD (ie, $FEV_1$/FVC ratio $<0.70$ on tests meeting ATS quality standards), and who provided all information needed for GOLD staging and appropriate self-assessment. Of the 899 enrolled in the study, eight withdrew before completing spirometry testing, leaving 891 who completed the spirometry phase. Of these, only 666 performed spirometry meeting ATS quality standards, and provided complete clinical information needed to calculate the new GOLD stage. Four hundred and fifty-three of these were confirmed to have spirometry confirmed COPD, and of these, only 445 properly completed the self-assessment questionnaire, and thus are the cohort included in these analyses.

## Data collection

Data collection was performed by investigators during a scheduled office visit. During the visit, physicians recorded the patient's clinical history, spirometry results obtained during the visit, and health care resource utilization in a web-based case report form. Prior to spirometry testing, investigators recorded their global assessment of the patient's COPD severity at the time of the study visit on a 5-point scale, ranging from 1 (no clinical symptoms or disease impact) to 5 (very severe). The 5-point scale was intended to correspond to the original GOLD COPD staging system, which ranged from stage 0 for persons with risk factors or symptoms but no airflow obstruction, and stages 1–4 (mild, moderate, severe, and very severe) for those proven to have airflow obstruction. Patients completed a paper questionnaire to collect standardized assessments including the CAT, mMRC, and a general assessment of severity at the time of the study visit on a 5-point scale, ranging from 1 (very mild) to 5 (very severe). Data were collected from February 2012 to November 2012. This study was approved and overseen by Sterling Institutional Review Board (Atlanta, Georgia), study number 3,872. Informed consent was obtained from all subjects.

## Spirometry procedures

Sites were provided an electronic, hand-held, MicroLoop™ portable spirometer and associated Spirometry PC Software™ for the study. Following ATS guidelines, relaxed spirometry testing was first used to capture three slow vital capacity results, and then forced spirometry testing was used to capture technically acceptable results for FVC and $FEV_1$. Up to eight efforts were required from each patient to obtain up to three acceptable tests per ATS guidelines. All spirometry measurements are reported pre-bronchodilator because it was not feasible to do pre- and post-bronchodilator testing in all clinics. Patients were asked not to use their COPD medications on the morning of the test. Predicted values and the percentage of predicted $FEV_1$ (%$pFEV_1$) were calculated using National Health and Nutrition Examination Survey III reference values.[33]

Prior to patient enrollment, investigators and study site staff completed real-time, study-specific training via an online meeting platform. Training addressed study procedures, including standard ATS spirometry procedures and use of the MicroLoop™ spirometer. Following enrollment of the first three patients at each study site, spirometry results were sent to an independent respiratory therapist experienced and certified in pulmonary function testing for quality control review. Additional guidance and training was provided to study site personnel if needed.

## Severity classification

Patients were classified into their traditional obstruction severity stage (stages 1–4, described as mild, moderate, severe, and very severe, respectively) based on their %$pFEV_1$ using GOLD guidelines.[16] Patients were classified into their new GOLD mMRC grade (ABCD), and their GOLD CAT grade (ABCD), by stratifying them by their %$pFEV_1$ and their mMRC or CAT scores, as per the new GOLD recommendations. Finally, we also classified patients by their PCPs' recorded history of exacerbations within the last 12 months. PCP and patient's self-assessed overall severity ratings were also used for classification. Very few patients were self-described as "very mild" or physician-described as "no clinical symptoms or disease impact", so these were combined with the mild or stage 1 category for all comparisons.

## Statistics

Statistical comparisons of continuous variables were made with student's *t*-tests and analysis of variance, as appropriate. Counts and percentages were compared using chi-square analyses. To compare agreement between perceived severity measures and the spirometry-based severity results, a Cohen's kappa coefficient was used.

This approach evaluates disagreement between levels of severity and provides a summary result ranging from 0 (no agreement) to 1 (perfect agreement). All analyses utilized a two-sided *P* of 0.05 for significance and were performed using SAS 9.2.

## Results

The demographic characteristics of the 445 COPD patients included in the analysis are presented in Table 1. Most were older (mean age 68 years) and well-established patients of the participating PCP, with a mean attendance in their clinic of 11 years. Most were current (38%) or former smokers (56%), with mean lifetime smoking histories of 39 and 52 pack-years, respectively. Approximately two thirds were overweight or obese, and 95% were being treated for at least one other chronic comorbidity.

**Table 1** Demographic characteristics

| | Total (N=445) | |
|---|---|---|
| | n | % |
| Age group (years) | | |
| 40–49 | 15 | 3 |
| 50–59 | 87 | 20 |
| 60–64 | 74 | 17 |
| 65–69 | 80 | 18 |
| 70–74 | 75 | 17 |
| 75–79 | 57 | 13 |
| 80+ | 57 | 13 |
| Sex | | |
| Male | 227 | 51 |
| Female | 218 | 49 |
| Education level | | |
| Did not complete high school | 61 | 14 |
| High school or equivalent (eg, GED) | 161 | 36 |
| Some college | 166 | 37 |
| College graduate | 33 | 7 |
| Graduate school | 24 | 5 |
| Race/ethnicity | | |
| Non-Hispanic White | 383 | 86 |
| Black or African American | 37 | 8 |
| Hispanic | 10 | 2 |
| Other or unspecified | 15 | 3 |
| Body Mass Index (BMI) | | |
| Underweight (<18.5) | 24 | 5 |
| Normal weight (18.5–24.9) | 121 | 27 |
| Overweight (25–29.9) | 145 | 33 |
| Obese (BMI of 30 or greater) | 155 | 35 |
| Comorbid conditions | | |
| Other respiratory | 113 | 25 |
| Hypertension | 291 | 65 |
| Cardiovascular | 200 | 45 |
| Neurological | 47 | 11 |
| Endocrine | 127 | 29 |
| Gastrointestinal | 144 | 32 |
| Renal/urologic | 60 | 13 |
| Ophthalmologic | 63 | 14 |
| Mental health | 152 | 34 |
| Muscle/bone | 199 | 45 |
| Blood | 36 | 8 |
| None | 23 | 5 |

**Abbreviation:** GED, General Educational Development.

### Stratification by traditional spirometry

The majority of patients in this cohort had moderate or severe airflow obstruction according to the traditional spirometry stage system (Table 2). Patients' self-assessments of their COPD severity were poorly congruent with their spirometry-based stage (κ=0.13), and more were wrong about their severity stage than correct (46% underestimated and 13% overestimated) (Figure 1). The PCP's severity ratings were also inconsistent and tended to underestimate their patient's severity; 34% were accurate as compared to the traditional spirometry stage, with 57% underestimated and 9% overestimated, for an overall kappa of 0.11 (Figure 2). Agreement between patient and their physician's assessments was also poor, with doctor's impressions tending to be less severe than the patient's (Figure 2).

### Stratification by the mMRC or CAT

Patients were then reclassified by the new GOLD system using their mMRC or CAT scores (Table 3). Substantial proportions of patients from the old severity system are reclassified, but how they are reclassified varies greatly by whether the mMRC or CAT system is selected.

After application of the new GOLD mMRC system, 48% (n=206) of the patients are re-stratified higher or lower than their spirometry level when distributed into the GOLD A, B, C, or D groups (Table 3). Among persons with mild airflow obstruction (stage 1), 81% are allocated to group A, and the remainder to group B. At the other end of the spectrum, patients with the most severe airflow obstruction (stage 4) tend to be in group D (70% vs 30% in group C). Patients with moderate airflow obstruction (stage 2) are relatively evenly distributed between groups A and B, and patients with severe obstruction (stage 3) are relatively evenly distributed between C and D. Therefore, the mMRC system is re-stratifying patients in the middle ranges of airflow obstruction according to their chronic symptoms, while those with the highest and lowest degrees of obstruction tend to stay in the highest (A) and lowest (D) groups. However, the agreement between the GOLD mMRC level and either the physician's global impression of severity or the patients' self-perception of severity is still poor (κ=0.17 and 0.13, respectively) (Figures 1 and 2).

**Table 2** Traditional spirometry severity, with patient and physician assessments of COPD severity prior to spirometry

| Traditional spirometry GOLD stage (N=445) N (column %) | Patient self-assessment of COPD severity (N=445) N (column %) | Physician assessment of patient severity (N=445) N (column %) |
|---|---|---|
| Stage 1 (mild) ($FEV_1$>80% predicted) N=43 (10%) | $FEV_1$ stage 1 (n=24)<br>$FEV_1$ stage 2 (n=63)<br>$FEV_1$ stage 3 (n=22)<br>$FEV_1$ stage 4 (n=2)<br>N=111 (25%) | $FEV_1$ stage 1 (n=32)<br>$FEV_1$ stage 2 (n=113)<br>$FEV_1$ stage 3 (n=43)<br>$FEV_1$ stage 4 (n=12)<br>N=200 (45%) |
| Stage 2 (moderate) ($FEV_1$ 50%–79% predicted) N=202 (45%) | $FEV_1$ stage 1 (n=15)<br>$FEV_1$ stage 2 (n=101)<br>$FEV_1$ stage 3 (n=73)<br>$FEV_1$ stage 4 (n=23)<br>N=212 (48%) | $FEV_1$ stage 1 (n=7)<br>$FEV_1$ stage 2 (n=63)<br>$FEV_1$ stage 3 (n=48)<br>$FEV_1$ stage 4 (n=14)<br>N=132 (30%) |
| Stage 3 (severe) ($FEV_1$ 30%–49% predicted) N=144 (32%) | $FEV_1$ stage 1 (n=4)<br>$FEV_1$ stage 2 (n=29)<br>$FEV_1$ stage 3 (n=48)<br>$FEV_1$ stage 4 (n=23)<br>N=104 (23%) | $FEV_1$ stage 1 (n=4)<br>$FEV_1$ stage 2 (n=24)<br>$FEV_1$ stage 3 (n=48)<br>$FEV_1$ stage 4 (n=23)<br>N=99 (22%) |
| Stage 4 (very Severe) ($FEV_1$<30% predicted) N=56 (13%) | $FEV_1$ stage 1 (n=0)<br>$FEV_1$ stage 2 (n=9)<br>$FEV_1$ stage 3 (n=1)<br>$FEV_1$ stage 4 (n=8)<br>N=18 (4%) | $FEV_1$ stage 1 (n=0)<br>$FEV_1$ stage 2 (n=2)<br>$FEV_1$ stage 3 (n=5)<br>$FEV_1$ stage 4 (n=7)<br>N=14 (3%) |

**Abbreviations:** GOLD, Global Initiative for Chronic Obstructive Lung Disease; $FEV_1$, forced expiratory volume in 1 second.

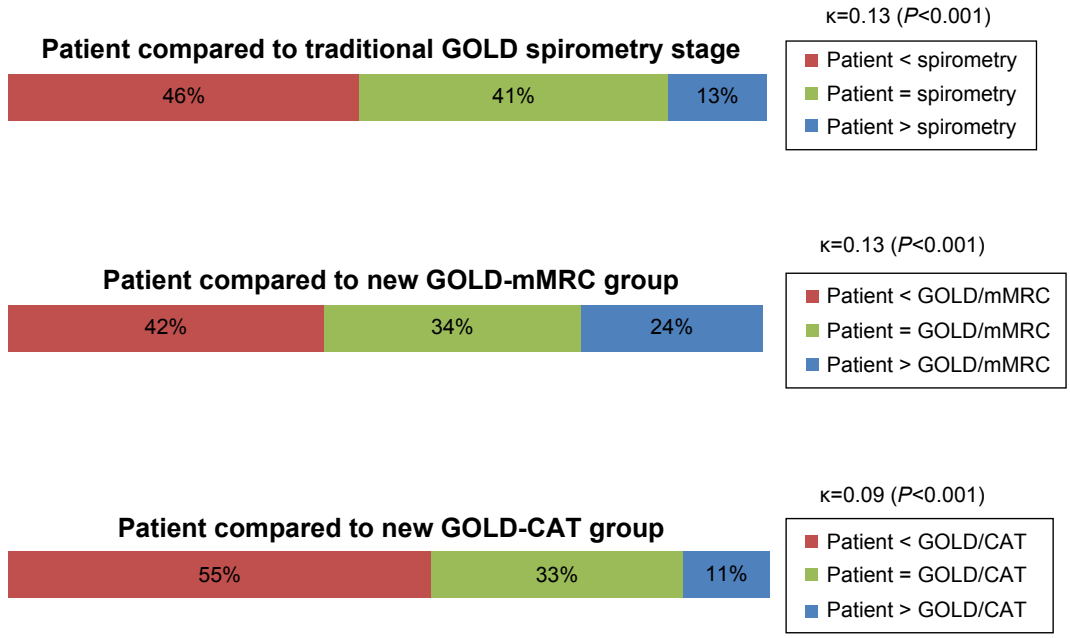


**Figure 1** Comparison of patient assessment of severity versus traditional GOLD spirometry stages, and new GOLD mMRC and CAT groups.
**Notes:** The colored bars represent the proportion of patients within each level of comparison; for example, 46% of patients rated their COPD as being less severe than the severity rating measured by spirometry. The kappa coefficient (κ) describes the agreement about COPD severity within each level of comparison and provides a summary result ranging from 0 (no agreement) to 1 (perfect agreement). While κ values less than 0.20 indicate very poor agreement, the *P*-values less than 0.001 suggest that these are still better than purely random associations.
**Abbreviations:** GOLD, Global Initiative for Chronic Obstructive Lung Disease; mMRC, modified Medical Research Council Dyspnea Scale; CAT, COPD Assessment Test.

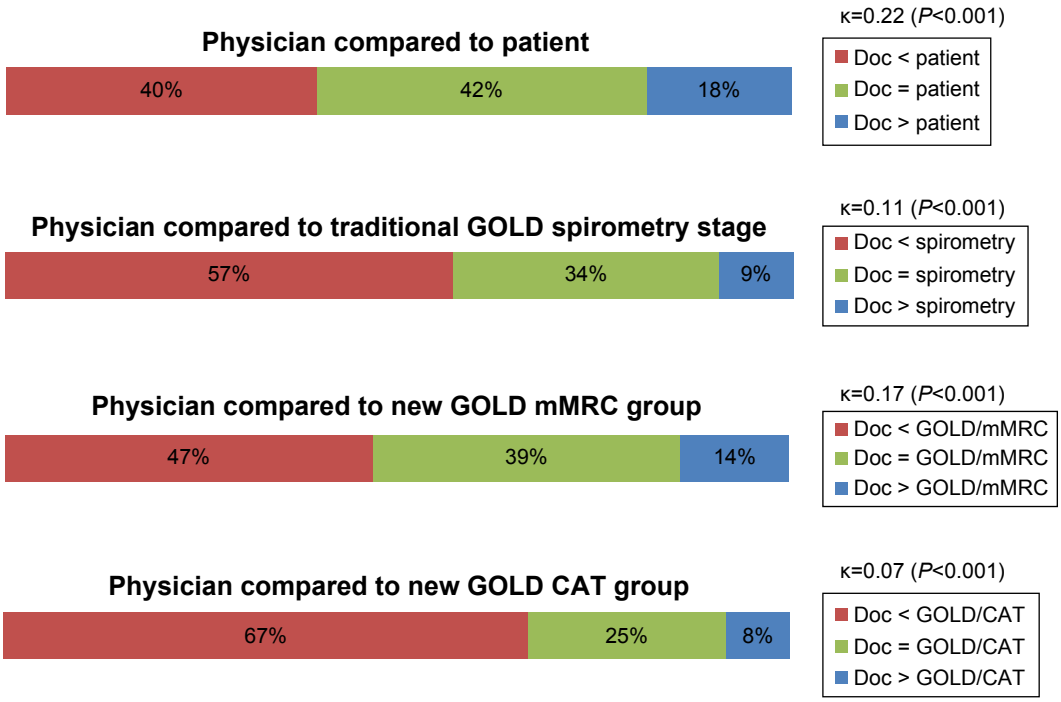


**Figure 2** Comparison of primary care physician assessment of severity versus the patient, traditional GOLD spirometry level, and the GOLD mMRC and CAT groups.
**Notes:** The colored bars represent the proportion of patients within each group of comparison; for example, 40% of physicians rated their patient's COPD as being less severe than the severity rating estimated by the patients themselves.
**Abbreviations:** GOLD, Global Initiative for Chronic Obstructive Lung Disease; mMRC, modified Medical Research Council Dyspnea Scale; CAT, COPD Assessment Test; Doc, doctor.

After reclassification by the new GOLD system using the CAT scores, 41% (n=179) of patients were re-stratified into a level higher or lower than their spirometry-based severity, but the distributions were much different than the mMRC results (Table 3). Among patients with the mildest obstruction (stage 1), only 38% are in group A; among those with moderate obstruction (stage 2), 88% are in group B; and among those with severe obstruction (stage 3), 90% are found in group D; thus the ability of the CAT to discriminate patients in the middle ranges of airflow obstruction by their chronic symptoms is very limited. Furthermore, the agreements between GOLD CAT severity level and either physician impression or patient self-assessment are even worse than by spirometry grade alone (κ=0.07 and 0.09, respectively) (Figures 1 and 2).

**Table 3** New GOLD grade by mMRC and CAT results, stratified by traditional GOLD spirometry stage

| GOLD grade | GOLD – mMRC (N=433)* (n) [%] of $FEV_1$ stage N (column %) | GOLD – CAT (N=438)* (n) [%] of $FEV_1$ stage N (column %) |
|---|---|---|
| A | Stage 1 (n=34) [81%]<br>Stage 2 (n=109) [55%]<br>N=143 (33%) | Stage 1 (n=16) [38%]<br>Stage 2 (n=24) [12%]<br>N=40 (9%) |
| B | Stage 1 (n=8) [19%]<br>Stage 2 (n=89) [45%]<br>N=96 (22%) | Stage 1 (n=26) [62%]<br>Stage 2 (n=173) [88%]<br>N=199 (45%) |
| C | Stage 3 (n=67) [48%]<br>Stage 4 (n=16) [30%]<br>N=83 (19%) | Stage 3 (n=15) [10%]<br>Stage 4 (n=1) [2%]<br>N=16 (4%) |
| D | Stage 3 (n=73) [52%]<br>Stage 4 (n=38) [70%]<br>N=111 (26%) | Stage 3 (n=128) [90%]<br>Stage 4 (n=55) [98%]<br>N=183 (42%) |

**Note:** *Patients who did not complete the mMRC or CAT questionnaires were excluded.
**Abbreviations:** GOLD, Global Initiative for Chronic Obstructive Lung Disease; $FEV_1$, forced expiratory volume in 1 second; mMRC, modified Medical Research Council Dyspnea Scale; CAT, COPD Assessment Test.

## Stratification by exacerbation history

We then stratified the history of exacerbations within the last 12 months by the GOLD spirometry, GOLD mMRC, and GOLD CAT systems (Table 4). We noted that physicians identified 14.8% of patients as "frequent exacerbators" (two or more exacerbations requiring steroids in the previous 12 months) while only 13.3% of patients self-reported two or more exacerbations requiring steroids, creating a possible misclassification error due to recall bias if patient history alone is used (data for patient-reported exacerbations not shown).

As expected, the incidence of exacerbations within the last year increased with the severity of airflow obstruction (GOLD $FEV_1$ stages 1–4) (Table 4). The percentage of

**Table 4** Twelve-month exacerbation history stratified by traditional GOLD spirometry stage, and new GOLD mMRC or CAT groups

| | | GOLD $FEV_1$ Stage 1 | GOLD $FEV_1$ Stage 2 | GOLD $FEV_1$ Stage 3 | GOLD $FEV_1$ Stage 4 | GOLD mMRC Group A | GOLD mMRC Group B | GOLD mMRC Group C | GOLD mMRC Group D | GOLD CAT Group A | GOLD CAT Group B | GOLD CAT Group C | GOLD CAT Group D |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Exacerbations | Column N | 43 | 202 | 144 | 56 | 143 | 96 | 83 | 111 | 40 | 199 | 16 | 183 |
| 0 | n | 37 | 153 | 81 | 25 | 118 | 67 | 49 | 54 | 35 | 150 | 13 | 92 |
| | % | 86 | 76 | 56 | 45 | 83 | 70 | 59 | 49 | 88 | 75 | 81 | 50 |
| 1 | n | 4 | 31 | 34 | 14 | 18 | 17 | 19 | 29 | 5 | 29 | 2 | 46 |
| | % | 9 | 15 | 24 | 25 | 13 | 18 | 23 | 26 | 13 | 15 | 13 | 25 |
| ≥2 | n | 2 | 18 | 29 | 17 | 7 | 12 | 15 | 28 | 0 | 20 | 1 | 45 |
| | % | 5 | 9 | 20 | 30 | 5 | 13 | 18 | 25 | 0 | 10 | 6 | 25 |

**Abbreviations:** GOLD, Global Initiative for Chronic Obstructive Lung Disease; $FEV_1$, forced expiratory volume in 1 second; mMRC, modified Medical Research Council Dyspnea Scale; CAT, COPD Assessment Test.

patients who had one exacerbation or frequent exacerbations also increased linearly by GOLD mMRC group (GOLD mMRC groups A to D, Table 4). Of the 433 patients who were reclassified under GOLD mMRC, seven of these group A patients and 12 group B patients would be promoted to groups C and D, respectively, because of their high risk for exacerbations. Therefore, after adjusting the COPD mMRC system by exacerbation history according to the physician, 33% were in group A, 22% in group B, 19% in group C, and 26% in group D.

The exacerbation history by GOLD CAT group did not increase steadily with severity (GOLD CAT groups A to D, Table 4). None of the 40 GOLD CAT group A patients were frequent exacerbators, so all would stay in group A. Of the 199 patients in GOLD CAT group B, 20 were frequent exacerbators and would therefore be upgraded to group D. Therefore, after adjusting the GOLD CAT system by exacerbation history according to the physician, 9% were in group A, 45% in group B, 4% in group C, and 42% in group D.

## Discussion

The new GOLD COPD assessment system adds chronic respiratory symptoms information and recent exacerbation history to reclassify persons into a two-dimensional matrix that should better characterize the disease impact on chronic symptoms and risk for exacerbations and be a more accurate guide for therapy.[16,17] In this study of primary care COPD patients, we found as expected that the traditional COPD severity system based solely on spirometry did not correlate well with either patient or physician perception of severity, although there was a weak correlation between exacerbation history and degree of airflow obstruction. The new GOLD system using the mMRC questionnaire does reclassify relatively equal proportions of patients with stage 2 (moderate) and stage 3 (severe) airflow obstruction based on their chronic dyspnea, so it appears to at least have the potential to characterize these patients in a clinically useful way. The history of exacerbations also increases steadily by mMRC level, which helps validate that the system is working as expected. However, the distributions by the CAT scores are so heavily skewed toward the B and D groups that only 13% of the cohort are left in either A or C, making it unlikely that the CAT questionnaire will add much to the spirometry assessment of any given patient. Adding exacerbation history to either the mMRC or CAT stratification reclassifies only an additional 4% of all patients, and the exacerbation histories are variable depending on whether they come from the physician or the patient. Finally, the agreement between

physician and patient assessments of COPD severity and the GOLD mMRC and CAT levels are not substantially better than those of the traditional spirometry-based system, and may be worse. In summary, the results of reclassifying primary care COPD patients by the new GOLD assessment system varies greatly by whether the CAT or mMRC system is chosen, and it is not clear that either adds much practical benefit to the traditional spirometry-based system.

A key motivation to move away from the traditional spirometry-only COPD assessment was the poor correlation between the degree of airflow obstruction and other clinical outcomes, which made the basis for treatment recommendations unstable.[16,17] It was presumed that adding the additional dimensions of chronic respiratory symptoms and exacerbation history would result in a more practical tool that would stratify patients into groups with similar characteristics that merit specific treatments. For example, among patients with increased chronic respiratory symptoms, those with stage 1 or stage 2 airflow obstruction would have indication for use of long-acting bronchodilators, and those with stage 3 or stage 4 airflow obstruction might merit "triple therapy" (inhaled corticosteroid + long-acting β2-agonist and long-acting anticholinergic). However, if a score of 10 is used as the cut-point for the CAT questionnaire, then the effect in a primary care population is to put approximately nine out of ten patients into a more aggressive treatment group, and one might reasonably ask whether it is worth the effort of giving the questionnaire to ten patients to find the one who does not merit additional medication. The mMRC does has the advantage of differentiating the patients with moderate or severe obstruction into more balanced groups, and thus it has the potential of differentiating more patients into clinically relevant treatment groups, but prospective analyses will be needed to confirm this.

An interesting finding in this project is that the agreement between patient or physician assessment of COPD severity and the objective measures of severity were not improved by the new GOLD assessment systems. An important feature of qualitative clinical assessment tools is the "face validity", which is the transparency or relevance of a test as it appears to participants.[34,35] The COPD assessment systems that were only based on spirometry were lacking in face validity because too often the measure (mild, moderate, or severe stage) did not correspond well with the patient's experience of symptoms and other disease consequences. If it works as expected, then the face validity of the new GOLD system should be improved as compared to the spirometry-only system because of the addition of the symptoms questionnaires. Unfortunately, our data suggest that face validity is not substantially improved by the GOLD mMRC system, and possibly worsened by the GOLD CAT system. Therefore, it is likely that the new GOLD system will also be limited by one of the main criticisms of the traditional spirometry-based system, that being a poor correlation with clinical perceptions about disease severity and health status. Furthermore, to become practical tools for primary care, severity assessment systems will need to be validated as accurate prognostic measures, such as in their ability to predict morbidity (eg, risk for COPD exacerbations) and mortality.

There are some important limitations of this study that should be considered. Patients were well-established COPD patients who had been seeing their PCPs an average of 11 years. It is not known whether results would be similar in newly diagnosed COPD patients or among patients in other countries. There are also likely to be selection biases based on patients' willingness to participate in research that could affect survey responses. We used standardized questionnaires that were administered by staff specifically trained for this study, and it is possible that patients participating in clinical research will behave and answer questions differently than they would in normal clinical conditions.

Another limitation is the problem of language – how does one define mild, moderate, or severe COPD? There are several problems introduced by defining severity by a simple linear 1-to-4 system based on spirometry that are further complicated by converting it to a 2-by-2 matrix that combines spirometry, either the mMRC or the CAT score, and exacerbation history. For example, the matrix scheme does not follow the usual progression of disease in that persons in group B (increased chronic symptoms but low risk of future adverse events) are more likely to progress to group D (increased chronic symptoms and future adverse events) than to group C (minimal chronic symptoms but high risk of adverse events). In a study of 6,628 COPD patients from the Copenhagen Heart Study, GOLD mMRC-derived group B patients had significantly worse 3-year all-cause hospitalization rates and survival than group C patients.[18] To compare how well the new GOLD system compares to the traditional GOLD severity scale in terms of how well it matches patients' self-assessments and their physicians' global assessment, we have assumed that group B equates to "moderate" and group C equates to "severe", but acknowledge that this is not a very stable assumption.

Our results were similar to those found in a study of COPD patients derived from general practices from the United Kingdom.[30] Haughney et al performed a retrospective

survey of 6,283 COPD patients who had $FEV_1$ and mMRC data. By the new GOLD system 36% were A, 19% B, 20% C, and 25% D, as compared to the spirometry grades of 17% stage 1, 52% stage 2, 26% stage 3, and 5% stage 4. They also found some degree of gradation in exacerbation risk by category, but their system for identifying exacerbations was based on chart review, and direct comparison between their exacerbation data and ours is not feasible from the published data. The studies that have compared GOLD mMRC and CAT classification in non-primary care populations also found wide discrepancies between them.[22–25,28,29]

## Conclusion

We found that the new GOLD COPD system reclassifies a substantial number of primary care patients as compared to the traditional spirometry-based severity system, but the reclassification is highly variable depending on whether the mMRC or CAT system is chosen. Furthermore, the poor agreement between the patients' and physicians' global assessments of severity scales even by the GOLD mMRC system makes it doubtful that this new system is capturing the major determinates that affect their perceptions about their disease progression or current status. It remains to be seen whether the new system improves PCPs' decisions about treatment or helps patients understand their lung disease.

## Disclosure

Dr Mapel is a paid consultant for GlaxoSmithKline, Ikaria, and Boehringer Ingelheim, and has received research funding from GlaxoSmithKline, AstraZeneca, Pfizer Pharmaceuticals, and Boehringer Ingelheim. Dr Dalal is an employee of GlaxoSmithKline in the US Health Outcomes division. The other authors are employees of Optum, which is a clinical research company that was contracted to conduct this study. The other authors have no conflicts of interest to disclose.

## References


1. Fletcher CM. *The natural history of chronic bronchitis and emphysema*. Oxford: Oxford University Press; 1976.
2. Peto R, Speizer FE, Cochrane AL, et al. The relevance in adults of airflow obstruction, but not of mucus hypersecretion, to mortality from chronic lung disease. Results from 20 years of prospective observation. *Am Rev Respir Dis*. 1983;128(3):491–500.
3. Anto JM, Vermeire P, Vestbo J, Sunyer J. Epidemiology of chronic obstructive pulmonary disease. *Eur Respir J*. 2001;17(5):982–994.
4. Jones PW, Anderson JA, Calverley PM, et al. Health status in the TORCH study of COPD: treatment efficacy and other determinants of change. *Respir Res*. 2011;12:71.
5. Weatherall M, Marsh S, Shirtcliffe P, Williams M, Travers J, Beasley R. Quality of life measured by the St George's Respiratory Questionnaire and spirometry. *Eur Respir J*. 2009;33(5):1025–1030.
6. Hoogendoorn M, Feenstra TL, Hoogenveen RT, Al M, Molken MR. Association between lung function and exacerbation frequency in patients with COPD. *Int J Chron Obstruct Pulmon Dis*. 2010;5:435–444.
7. Hurst JR, Vestbo J, Anzueto A, et al. Susceptibility to exacerbation in chronic obstructive pulmonary disease. *N Engl J Med*. 2010;363(12):1128–1138.
8. Jones P, Miravitlles M, van der Molen T, Kulich K. Beyond FEV(1) in COPD: a review of patient-reported outcomes and their measurement. *Int J Chron Obstruct Pulmon Dis*. 2012;7:697–709.
9. Westwood M, Bourbeau J, Jones PW, Cerulli A, Capkun-Niggli G, Worthy G. Relationship between $FEV_1$ change and patient-reported outcomes in randomised trials of inhaled bronchodilators for stable COPD: a systematic review. *Respir Res*. 2011;12:40.
10. Tsiligianni I, Kocks J, Tzanakis N, Siafakas N, van der Molen T. Factors that influence disease-specific quality of life or health status in patients with COPD: a review and meta-analysis of Pearson correlations. *Prim Care Respir J*. 2011;20(3):257–268.
11. Shavelle RM, Paculdo DR, Kush SJ, Mannino DM, Strauss DJ. Life expectancy and years of life lost in chronic obstructive pulmonary disease: findings from the NHANES III Follow-up Study. *Int J Chron Obstruct Pulmon Dis*. 2009;4:137–148.
12. Ford ES, Mannino DM, Wheaton AG, Giles WH, Presley-Cantrell L, Croft JB. Trends in the prevalence of obstructive and restrictive lung function among adults in the United States: findings from the National Health and Nutrition Examination surveys from 1988–1994 to 2007–2010. *Chest*. 2013;143(5):1395–1406.
13. Rennard SI, Vestbo J, Agusti A. What is chronic obstructive pulmonary disease anyway?: Continua, categories, cut points, and moving beyond spirometry. *Am J Respir Crit Care Med*. 2013;187(10):1036–1037.
14. Jones PW, Harding G, Berry P, Wiklund I, Chen WH, Kline Leidy N. Development and first validation of the COPD Assessment Test. *Eur Respir J*. 2009;34(3):648–654.
15. Bestall JC, Paul EA, Garrod R, Garnham R, Jones PW, Wedzicha JA. 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):581–586.
16. goldcopd.org [homepage on the Internet]. Global Strategy for the Diagnosis, Management and Prevention of COPD. 2014 GOLD COPD Committee. Available from: http://www.goldcopd.org/. Accessed December 6, 2014.
17. Vestbo J, Hurd SS, Rodriguez-Roisin R. The 2011 revision of the global strategy for the diagnosis, management and prevention of COPD (GOLD) – why and what? *Clin Respir J*. 2012;6(4):208–214.
18. Lange P, Marott JL, Vestbo J, et al. Prediction of the clinical course of chronic obstructive pulmonary disease, using the new GOLD classification: a study of the general population. *Am J Respir Crit Care Med*. 2012;186(10):975–981.
19. Johannessen A, Nilsen RM, Storebo M, Gulsvik A, Eagan T, Bakke P. Comparison of 2011 and 2007 Global Initiative for Chronic Obstructive Lung Disease guidelines for predicting mortality and hospitalization. *Am J Respir Crit Care Med*. 2013;188(1):51–59.
20. Han MK, Muellerova H, Curran-Everett D, et al. GOLD 2011 disease severity classification in COPDGene: a prospective cohort study. *Lancet Respir Med*. 2013;1(1):43–50.
21. Soriano JB, Alfageme I, Almagro P, et al. Distribution and prognostic validity of the new Global Initiative for Chronic Obstructive Lung Disease grading classification. *Chest*. 2013;143(3):694–702.
22. Kim S, Oh J, Kim YI, et al. Differences in classification of COPD group using COPD assessment test (CAT) or modified Medical Research Council (mMRC) dyspnea scores: a cross-sectional analyses. *BMC Pulm Med*. 2013;13:35.
23. Jones PW, Adamek L, Nadeau G, Banik N. Comparisons of health status scores with MRC grades in COPD: implications for the GOLD 2011 classification. *Eur Respir J*. 2013;42(3):647–654.
24. Jones PW, Nadeau G, Small M, Adamek L. Characteristics of a COPD population categorised using the GOLD framework by health status and exacerbations. *Respir Med*. 2014;108(1):129–135.

25. Casanova C, Marin JM, Martinez-Gonzalez C, et al. New GOLD classification: longitudinal data on group assignment. *Respir Res*. 2014;15:3.
26. de Torres JP, Marin JM, Martinez-Gonzalez C, et al. Clinical application of the COPD assessment test: longitudinal data from the COPD History Assessment in Spain (CHAIN) cohort. *Chest*. 2014;146(1):111–122.
27. de Torres JP, Casanova C, Marin JM, et al. Prognostic evaluation of COPD patients: GOLD 2011 versus BODE and the COPD comorbidity index COTE. *Thorax*. 2014;69(9):799–804.
28. Rieger-Reyes C, Garcia-Tirado FJ, Rubio-Galan FJ, Marin-Trigo JM. Classification of chronic obstructive pulmonary disease severity according to the new Global Initiative for Chronic Obstructive Lung Disease 2011 guidelines: COPD assessment test versus modified Medical Research Council scale. *Arch Bronconeumol*. 2014;50(4):129–134.
29. Price DB, Baker CL, Zou KH, Higgins VS, Bailey JT, Pike JS. Real-world characterization and differentiation of the Global Initiative for Chronic Obstructive Lung Disease strategy classification. *Int J Chron Obstruct Pulmon Dis*. 2014;9:551–561.
30. Haughney J, Gruffydd-Jones K, Roberts J, Lee AJ, Hardwell A, McGarvey L. The distribution of COPD in UK general practice using the new GOLD classification. *Eur Respir J*. 2014;43(4):993–1002.
31. Gruffydd-Jones K. GOLD guidelines 2011: what are the implications for primary care? *Prim Care Respir J*. 2012;21(4):437–441.
32. Mapel DW, Dalal AA, Johnson P, Becker L, Goolsby Hunter A. A Clinical Study of COPD Severity Assessment by Primary Care Physicians and their Patients as Compared to Spirometry. *Am J Med*. Epub 2015 Jan 13.
33. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med*. 1999;159(1):179–187.
34. Gravetter FJF, Lori-Ann B. *Research Methods for the Behavioral Sciences*, 4th Edition. Belmont, CA: Wadsworth; 2012.
35. Lloyd H, Jenkinson C, Hadi M, Gibbons E, Fitzpatrick R. Patient reports of the outcomes of treatment: a structured review of approaches. *Health Qual Life Outcomes*. 2014;12:5.