

Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation



Original Research

Mortality and Exacerbations by Global Initiative for Chronic Obstructive Lung Disease Groups ABCD: 2011 Versus 2017 in the COPDGene® Cohort

Rachel N. Criner, MD¹ Wassim W. Labaki, MD² Elizabeth A. Regan, MD, PhD³ Jessica M. Bon, MD, MS⁴
Xavier Soler, MD, PhD⁵ Surya P. Bhatt, MD⁶ Susan Murray, ScD⁷ John E. Hokanson, PhD⁸
Edwin K. Silverman, MD, PhD⁹ James D. Crapo, MD³ Jeffrey L. Curtis, MD^{2,10} Fernando J. Martinez, MD, MS²
Barry J. Make, MD³ MeiLan K. Han MD, MS² Carlos H. Martinez, MD, MPH;² for the COPDGene® Investigators

Abstract

Background: The Global initiative for chronic Obstructive Lung Disease (GOLD) ABCD groupings were recently modified. The GOLD 2011 guidelines defined increased risk as forced expiratory volume in 1 second (FEV₁) < 50% predicted or ≥ 2 outpatient or ≥ 1 hospitalized exacerbation in the prior year, whereas the GOLD 2017 guidelines use only exacerbation history. We compared mortality and exacerbation rates in the Genetic Epidemiology of COPD Study cohort (COPDGene®) by 2011 (exacerbation history/FEV₁ and dyspnea) versus 2017 (exacerbations and dyspnea) classifications.

Methods: Using data from COPDGene®, we tested associations of ABCD groups with all-cause mortality (Cox models, adjusted for age, sex, race and comorbidities) and longitudinal exacerbations (zero-inflated Poisson models).

Results: In 4469 individuals (mean age 63.1 years, 44% female), individual distributions in 2011 versus 2017 systems were: A, 32.0% versus 37.0%; B, 17.6% versus 36.3%; C, 9.4% versus 4.4%; D, 41.0% versus 22.3%; (observed agreement 76% [expected 27.8%], Kappa 0.67, $p < 0.001$). Individuals in group D-2011 had 1.1 ± 1.6 exacerbations/year (mean \pm standard deviation [SD]) versus 1.4 ± 1.8 for D-2017 (median follow-up 3.7 years). Using group A as reference, for both systems, mortality (median follow-up 6.8 years) was highest in group D (D-2011, [hazard ratio] HR 5.2 [95% confidence interval (CI) 4.2, 6.4]; D-2017, HR 5.5 [4.5, 6.8]), lowest for group C (HR 1.9 [1.4, 2.6] versus HR 1.9 [1.3, 2.8]) and intermediate for group B (HR 2.6 [2.0, 3.4] versus HR 3.4 [2.8, 4.1]). GOLD 2011 had better mortality discrimination (area under the curve [AUC] 0.68) than GOLD 2017 (AUC 0.66, $p < 0.001$ for comparison) but similar exacerbation rate prediction.

Conclusions: Relative to the GOLD 2011 consensus statement, discriminate predictive power of the 2017 ABCD classification is similar for exacerbations but lower for survival.

Abbreviations: Global Initiative for Chronic Obstructive Lung Disease, **GOLD**; forced expiratory volume in 1 second, **FEV₁**; Genetic Epidemiology of COPD Study, **COPDGene®**; standard deviation, **SD**; hazard ratio, **HR**; confidence interval, **CI**; area under the curve, **AUC**; chronic obstructive pulmonary disease, **COPD**; modified Medical Research Council dyspnea scale, **mMRC**; COPD Assessment Test, **CAT**; St George's Respiratory Questionnaire, **SGRQ**; social security death index, **SSDI**; COPDGene® Longitudinal Follow-up Program, **LFU**; receiver operating curve, **ROC**; 6-minute walking distance, **6MWD**

Funding Support: COPDGene® is supported by funding from the National Heart, Lung, and Blood Institute (R01 HL089897 and R01 HL089856, HL138188). COPDGene® is also supported by the COPD Foundation through contributions made to an Industry Advisory Board comprised of AstraZeneca, Boehringer-Ingelheim, Novartis, Pfizer, GlaxoSmithKline, Siemens and Sunovion.

Date of Acceptance: August 24, 2018

Citation: Criner RN, Labaki WW, Regan EA, et al; for the COPDGene Investigators. Mortality and exacerbations by Global Initiative for Chronic Obstructive Lung Disease Groups ABCD: 2011 versus 2017 in the COPDGene® Cohort. *Chronic Obstr Pulm Dis.* 2019;6(1):64-73. doi: <https://doi.org/10.15326/jcopdf.6.1.2018.0130>

For personal use only. Permission required for all other uses.

- 1 Department of Internal Medicine, University of Michigan, Ann Arbor
- 2 Division of Pulmonary and Critical Care Medicine, University of Michigan, Ann Arbor
- 3 Division of Pulmonary Medicine, National Jewish Health, Denver, Colorado
- 4 Division of Pulmonary and Critical Care Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania
- 5 Division of Pulmonary, Critical Care and Sleep Medicine, University of California, San Diego
- 6 Division of Pulmonary, Allergy and Critical Care Medicine, University of Alabama, Birmingham
- 7 School of Public Health, University of Michigan, Ann Arbor
- 8 School of Public Health, University of Colorado, Aurora
- 9 Channing Division of Network Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts
- 10 Veterans' Administration Ann Arbor Healthcare System, Ann Arbor, Michigan

Keywords:

chronic obstructive pulmonary disease; COPD; Global Initiative for Chronic Obstructive Lung Disease ABCD groupings 2011 and 2017; exacerbation mortality

Address correspondence to:

MeiLan Han, MD, MS
 Division of Pulmonary and Critical Care Medicine
 3916 Taubman Center
 1500 E. Medical Center Drive
 University of Michigan Medical Center
 Ann Arbor, MI 48109-0360
 Phone: 734-763-2540
 Email: Mrking@med.umich.edu

This article contains an online supplement.

Introduction

The Global initiative for chronic Obstructive Lung Disease (GOLD) grading system for chronic obstructive pulmonary disease (COPD) has evolved greatly over the past 20 years. Initially, only forced expiratory volume in 1 second (FEV₁) was used to classify COPD severity.¹ This decision was later criticized because FEV₁ is an imperfect measure of the individual disease experience, which is better captured using additional measures such as dyspnea and walking distance.²⁻⁴ In 2011, the GOLD committee proposed a classification system of groups ABCD, based in a 2-step process that uses the worst value of 1 of 2 predictors of risk (FEV₁ or exacerbation history), combined with a measure of the burden of symptoms (based in the modified Medical

Research Council dyspnea [mMRC] or the COPD Assessment Test [CAT] scores).^{5,6}

The ABCD system was considered as an improvement, which acknowledged that COPD is a multidimensional disease with different factors determining disease impact and brought an opportunity to provide more individualized care to COPD patients.⁷ However, the 2011 ABCD system lacked superiority to the earlier simple spirometry criteria in predicting mortality and there was confusion regarding group D as it was defined by both lung function and exacerbation history.^{6,8-11} Therefore, in the most recent iteration of the GOLD consensus statement released in 2017,^{2,3,6,8-10} spirometry-determined airflow limitation was separated from clinical parameters. Currently, the ABCD classification is based on measures of symptoms (on the horizontal axis) and risk (on the vertical axis), which is determined only by exacerbation history.⁶

The 2017 ABCD system is an assessment tool, designed to aid clinicians in creating optimal treatment regimens for their patients based on disease characteristics, and not intended for prognostication.⁶ Nonetheless, it is important to analyze the impact of the 2017 ABCD system in comparison to the 2011 ABCD system, in regards to association with exacerbation and mortality, as they guide management for millions of COPD patients worldwide.

Using cross-sectional and longitudinal data from the Genetic Epidemiology of COPD Study (COPDGene®) cohort, we examined whether there are differences in mortality and exacerbation prediction between the 2011 ABCD group (defined using the worst descriptor of risk, either spirometry or exacerbation history, with mMRC as descriptor of symptoms) and the new 2017 ABCD group (exacerbation history and symptoms). Specifically, we investigated if, using GOLD 2017 criteria relative to GOLD 2011 criteria: (1) the distribution of individuals between groups ABCD are altered; (2) the mortality and exacerbation risk differ between groups ABCD; and (3) the discrimination ability to predict mortality between groups ABCD are changed.

Methods

Study Design and Participants

This is a longitudinal analysis of individuals with COPD enrolled in the COPDGene® cohort (Clinicaltrials.gov identifier NCT00608764), which comprises 21

academic clinical centers within the United States. In its initial phase, COPDGene[®] enrolled over 10,000 ever-smokers at risk for and with COPD. Individuals were required to be between 45-80 years of age with a ≥ 10 pack-years smoking history and free of exacerbations in the 4 weeks before enrollment. Additional inclusion and exclusion criteria have been published.¹²

For the current analyses, we included the 4469 COPDGene[®] participants fulfilling the fixed-ratio spirometry definition of COPD as a post-bronchodilator $FEV_1/FVC < 0.7$. All participants provided informed consent, and institutional review board (IRB) approval was obtained in all participating centers. Using self-administered, validated questionnaires, COPDGene[®] investigators collected data on demographics, biometrics, medical history, and medications. Participants also answered the modified American Thoracic Society Respiratory Epidemiology Questionnaire¹³ to identify exacerbation frequency, the St George's Respiratory Questionnaire (SGRQ)¹⁴ for obstructive disease-specific quality of life measurement, and the mMRC scale for dyspnea assessment.

Exposure

We applied GOLD 2011 and GOLD 2017 criteria to all enrolled individuals on whom vital status was available as of October 2016 (median follow-up time 6.8, interquartile range 6.0, 7.7 years); based on data collected at Visit 1, we classified their GOLD stage and severity based on ABCD groups.⁷ Symptoms were first assessed using the mMRC dyspnea scale. Individuals scoring mMRC 0-1 were low symptom burden. To determine the *risk* descriptor in the 2011 grouping system, we used spirometry (with low risk defined as GOLD spirometry stages 1-2 [$FEV_1 > 50\%$ predicted]) and exacerbation history (low risk 0-1 outpatient exacerbations and no inpatient treated exacerbations; high risk 2 or more outpatient-treated exacerbations or 1 or more exacerbations leading to a hospital admission). A combination of a descriptor of *symptom burden* (mMRC) and *risk* (the worst risk descriptor, either spirometry or exacerbation history) were used to create groups ABCD, resulting in group A as low risk/low symptoms, group B low risk/high symptoms, Group C high risk/low symptoms, and group D high risk/high symptoms.

The GOLD 2017 criteria did not use spirometry to guide ABCD grouping; instead *burden of symptoms*

(similar mMRC criteria) and exacerbation history alone were used as descriptors of *risk* (low risk 0-1 outpatient exacerbations and no inpatient treated exacerbations; high risk 2 or more outpatient-treated exacerbations or 1 or more exacerbations leading to a hospital admission).⁶ The combination of mMRC and exacerbations was used to create the 2017 ABCD groups.

Outcomes

Our outcomes were all-cause mortality and longitudinal, patient-reported exacerbations. Information on exacerbations was obtained using a previously described automated telecommunication system, with longitudinal follow-up data on exacerbations available up to October of 2013 (median follow-up 3.7, interquartile range 2.8, 4.8 years).¹⁵ Briefly, all participants agreed to be contacted by email or telephone on a regular basis and to provide information about development of new exacerbations in the interval since the initial visit. Participants were followed longitudinally via either an automated telephone or web-based system on an every 6-month basis. Individuals not reached through the automated system were contacted by a research coordinator for a phone-based interview. Exacerbations were defined as any episode of increased cough, sputum production, or shortness of breath lasting >48 hours and requiring treatment with antibiotics, systemic steroids, or both; exacerbations requiring hospital admission or emergency department visits were considered severe.

We confirmed participant vital status using multiple sources; information from the social security death index (SSDI), next-of-kin interviews, clinical records, and the COPDGene[®] Longitudinal Follow-up (LFU) Program were utilized. The mortality dataset was created in December 2016. Care was taken to avoid ascertainment bias, which can occur when deaths are reported more consistently than alive status. Due to individual center IRB restrictions, only 8675 individuals had vital status searched by SSDI. Deaths and vital status for those who were searched via SSDI were back censored 3 months from the search date to account for expected lag time between a death and its appearance in the SSDI dataset. The centralized SSDI search was performed on October 14, 2016. For a subset of 333 individuals participating in the LFU program and for whom SSDI search could not be performed, but had a survey completed in the 7 months prior, ensuring

that they were being actively followed, they were back censored 6 months prior to dataset generation date (June 18, 2016). For those in the LFU program who expired prior to the 6-month censor date, follow-up time terminating in death was included in the dataset if an LFU contact in the prior 7 months indicated that the individual was being “actively followed” at the time of death.

Statistical Analysis

Descriptive statistics used means and standard deviation, or proportions, according to the variable of interest. Bivariate associations between each ABCD grouping system and mortality and annualized exacerbation frequency were performed. Agreement of assignment of individuals to ABCD classification groups between the GOLD 2011 and GOLD 2017 systems was examined using the kappa statistic. We assessed differences in all-cause mortality across groups A to D in each classification using Kaplan-Meier analysis. Using Cox proportional hazard models, we also analyzed the association between each group in the GOLD 2011 and 2017 systems with mortality. Cox models were adjusted for demographics (age, sex, race), cardiovascular comorbidities (a combination of myocardial infarction, angina, congestive heart failure), cardiovascular risk factors (diabetes, obesity, hypertension, hyperlipidemia), musculoskeletal disease (combination of osteoarthritis and osteoporosis) and asthma. We tested associations of ABCD groups with exacerbations using both zero-inflated Poisson regression models, equally adjusted. The discrimination ability of the 2011 (mMRC and spirometry) and 2017 (mMRC and exacerbation history) systems were tested by a comparison of the area under the receiver operating curves (ROC) for mortality during follow-up and any exacerbation during follow-up.

We also performed 2 additional sensitivity analyses. First, because the GOLD consensus statement allows for symptom burden with CAT, in addition to mMRC, we looked for an alternative measure of symptoms. Because COPDGene[®] did not collect information on CAT during the first visit, we tested a modified ABCD grouping alternative, combining SGRQ (cut-point of 25 based on previous data showing that this value is comparable to a CAT score of 10)^{16,17} as measure of risk, with spirometry (2011) or exacerbations (2017) and replicated all analyses with these 2 grouping systems.

Second, we examined whether the associations were maintained when the outcome was severe (hospital or emergency department treated) exacerbations, using the 2011 and 2017 ABCD groups. All analyses were performed using the statistical package STATA v.12 (Stata Corp. College Station, Texas), and a *p* value <0.05 was considered statistically significant.

Results

Description of Participants

We included 4469 COPDGene[®] enrollees who fulfilled the fixed ratio definition of COPD, GOLD stages 1-4. Individuals were middle-aged (mean 63.1 years) with sizeable fractions of women (44%) and African-Americans (22.6%) (Table 1). The group had a significant burden of common comorbidities, included hypertension (48.2%), musculoskeletal disease (37.1%), gastroesophageal reflux disease (29.1%), and asthma (22.4%). The distribution of individuals based on the 2011 versus 2017 systems showed fair agreement (kappa score 0.67, *p*<0.001). Relative to the 2011 system, the 2017 classification scheme resulted in increased percentages of individuals categorized as groups A and B and a decrease in groups C and D (Table 2).

Mortality by GOLD ABCD Groups

As anticipated, Kaplan-Meier analysis showed greatest all-cause mortality over the study period in group D individuals, whether classified by 2011 GOLD criteria (Fig. 1A) or 2017 GOLD criteria (Fig. 1B); 8-year survival in this group was <60%. Eight-year survival (roughly 80%) did not differ between groups B and C in the 2011 classification; in contrast, unadjusted survival using the 2017 system was greater for group C than group B (Fig. 1). Nevertheless, adjusted Cox proportional models (using group A as reference) demonstrated that the risk of death using either system was lowest for group C individuals, intermediate for group B, and greatest for group D (Table 3). Although a small difference, GOLD 2011 exhibited significantly better discrimination in predicting mortality, as defined by area under the curve (AUC) for receiver-operator analysis, than the 2017 classification (AUC 0.68 [95% CI 0.67, 0.70] versus 0.66 [0.65, 0.68], *p*<0.001 for the comparison) (Table 3).

Table 1. Characteristics of Participants with COPD from the COPDGene® Cohort^{a,b}

Demographics	
Age in Years	63.1 (8.6)
Female Sex (%)	44.1
African Americans (%)	22.6
Body Mass Index in kg/m ²	27.9 (6.1)
Smoking History	
Current Smokers (%)	43.4
Smoking History	51.6 (27.2)
Respiratory Health	
Chronic Bronchitis (%)	25.9
mMRC Dyspnea Score ≥ 2 (%)	58.6
FEV ₁ in Liters	1.66 (0.79)
FEV ₁ % Predicted	57.4 (22.8)
SGRQ Total	36.8 (22.9)
6MWD in Feet	1232 (408)
GOLD Spirometry Stage (%)	
1	17.8
2	42.8
3	25.9
4	13.5
Comorbidities (%)	
Hypertension	48.2
Coronary Artery Disease	14.3
Angina	5.4
Diabetes Mellitus	12.3
Gastroesophageal Reflux	29.1
Asthma	22.4
Musculoskeletal Disease	37.1

^aData presented as mean (standard deviation), unless otherwise specified

^bn=4469

COPD=chronic obstructive pulmonary disease; COPDGene®=Genetic Epidemiology of COPD Study; mMRC=modified Medical Research Council dyspnea; FEV₁=forced expiratory volume in 1 second; SGRQ=St George's Respiratory Questionnaire; 6MWD=6-minute walk distance; GOLD=Global initiative for Chronic Obstructive Lung Disease

Table 2. Distribution of Participants with COPD by ABCD groups, GOLD 2011 and 2017

	GOLD 2011 (%)	GOLD 2017 (%)
Group A	32.0	37.0
Group B	17.6	36.3
Group C	9.4	4.4
Group D	41.0	22.3
	Agreement	0.76
	Expected	0.28
	Kappa	0.67

COPD=chronic obstructive pulmonary disease; GOLD=Global initiative for chronic Obstructive Lung Disease

All Exacerbations by GOLD ABCD Groups

Similarly, the unadjusted exacerbation frequency was greatest in group D by either GOLD classification scheme (GOLD 2011, 1.1 ± 1.6 exacerbations/year [mean ± SD]; GOLD 2017, 1.4 ± 1.8). In adjusted models with group A as a reference, group D had a markedly elevated fold-increase in exacerbations (GOLD 2011, hazard ratio (HR) 5.6 [95% confidence interval (CI) 4.9, 6.4]; GOLD 2017, 6.3 [5.5, 7.1]). Despite the changes in ABCD groupings seen in Table 2, there were no significant differences in exacerbation rates between groups B and C as assigned by the 2011 versus 2017 classification schemes (Table 3), although both were significantly elevated relative to group A in their respective schemes (B-2011, fold-increase HR 2.4 [95% CI 2.1, 2.9] versus B-2017, 2.7 [2.4, 3.0]; C-2011, 2.9 [2.4, 3.5] versus C-2017, 2.8 [2.3, 3.5]). Only a small difference between GOLD classifications was seen for prediction of all exacerbations (GOLD 2011, AUC 0.67 versus GOLD 2017, AUC 0.65, $p = 0.006$ for the comparison) (Table 3).

Analyses of Severe Exacerbations

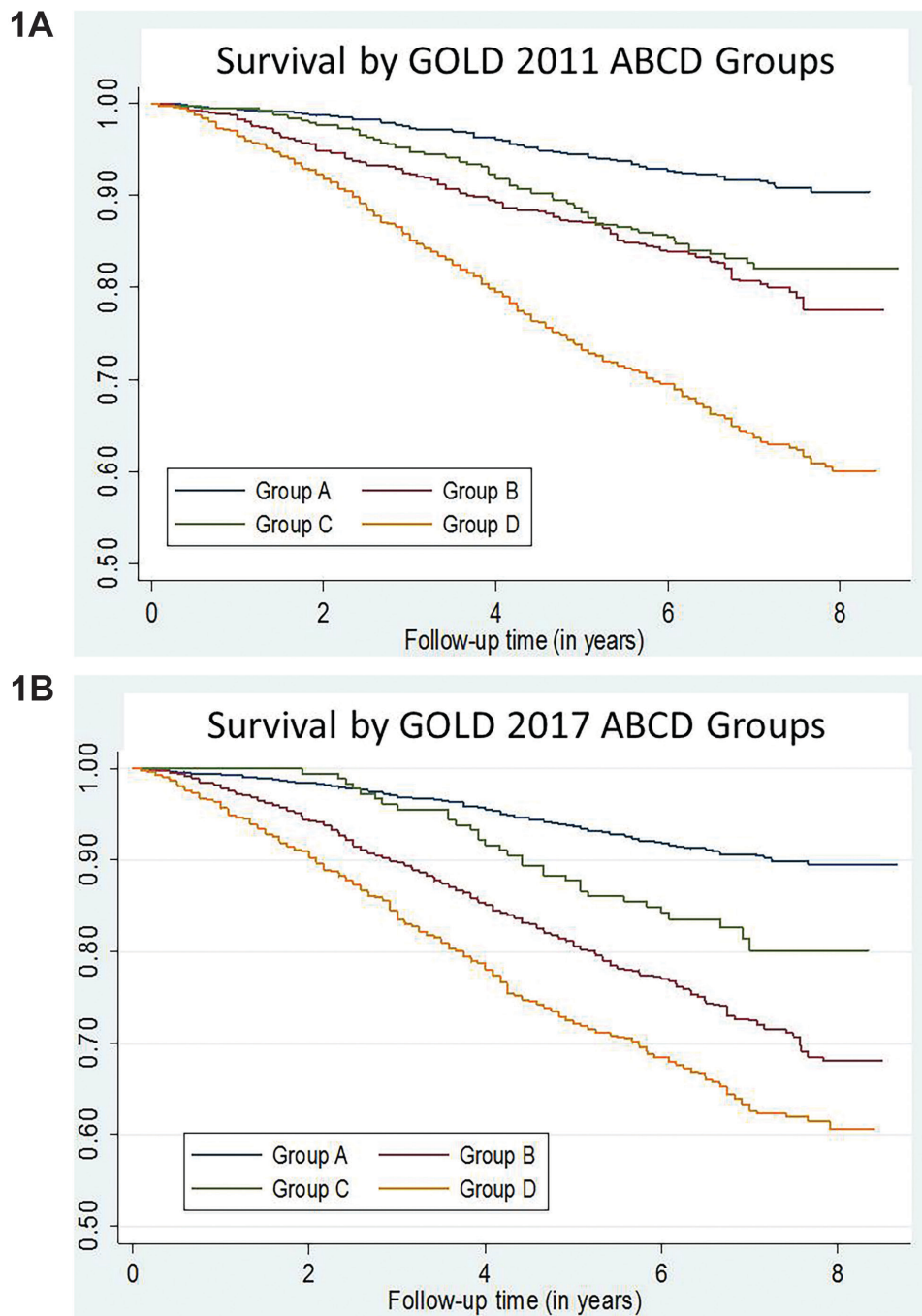
When defined by GOLD 2017, the annual rate of severe exacerbations also increased progressively from groups A through D (Online data supplement Table 1). By contrast, for GOLD 2011 (using group A as reference), although group D had the highest annual exacerbation rate (1.1 ± 2.2), group B had a similar rate to group C (0.4 ± 1.1 versus 0.4 ± 1.4). There was no significant difference between the 2011 and 2017 systems when predicting severe exacerbation frequency (AUC 0.69 versus 0.69, $p = 0.88$ for the comparison) (Online data supplement Table 1).

Analyses Using SGRQ as Measure of Symptoms

Finally, we substituted SGRQ for mMRC as a measure of respiratory symptoms and assessed the effect on mortality rates. Relative to the 2011 classification, this modification also increased the frequency of individuals categorized as being in groups A and B, had a negligible effect on group C, and decreased those in group D (GOLD 2011, 44.7% versus GOLD 2017, 24.5%) (Online data supplement Table 2). The agreement between schemes was again fair (kappa 0.66).

Results of this modification were similar to those using mMRC as measure of symptoms (Online data supplement Table 3). Unadjusted mortality according

Figure 1. Survival Curves for ABCD Groups by GOLD-2011^a and GOLD-2017^b



^apanel A

^bpanel B

GOLD=Global initiative for chronic Obstructive Lung Disease.

Table 3. Mortality^a and Exacerbation^b Frequency of Participants with COPD by ABCD Groups, GOLD 2011 and GOLD 2017^c

	Mortality		Exacerbations		
	GOLD 2011	GOLD 2017		GOLD 2011	GOLD 2017
Group A					
Frequency	7.4 %	8.2 %	Annual Rate	0.2 (0.6)	0.2 (0.6)
HR (95% CI)	Reference	Reference	Fold Increase	Reference	Reference
Group B					
Frequency	16.5 %	23.8 %	Annual Rate	0.5 (1.0)	0.6 (1.1)
HR (95% CI)	2.6 (2.0, 3.4)	3.4 (2.8, 4.1)	Fold Increase	2.4 (2.1, 2.9)	2.7 (2.4, 3.0)
Group C					
Frequency	15.0 %	16.2 %	Annual Rate	0.6 (1.1)	0.7 (1.2)
HR (95% CI)	1.9 (1.4, 2.6)	1.9 (1.3, 2.8)	Fold Increase	2.9 (2.4, 3.5)	2.8 (2.3, 3.5)
Group D					
Frequency	30.9 %	31.1 %	Annual Rate	1.1 (1.6)	1.4 (1.8)
HR (95% CI)	5.2 (4.2, 6.4)	5.5 (4.5, 6.8)	Fold Increase	5.6 (4.9, 6.4)	6.3 (5.5, 7.1)
Overall Discrimination					
AUC (95% CI)	0.682 (0.665, 0.700)	0.664 (0.646, 0.683)	AUC (95% CI)	0.666 (0.651, 0.681)	0.653 (0.638, 0.669)
	<i>p</i> <0.001			<i>p</i> =0.006	

^aMortality models based on COX proportional hazard models, adjusted for demographics (age, sex, race), cardiovascular comorbidities (a combination of myocardial infarction, angina, congestive heart failure), cardiovascular risk factors (diabetes, obesity, hypertension, hyperlipidemia), musculoskeletal disease (combination of osteoarthritis and osteoporosis) and asthma.

^bExacerbation models based on zero-inflated Poisson models, equally adjusted.

^cData presented as mean (standard deviation), unless otherwise specified.

COPD=chronic obstructive pulmonary disease; GOLD=Global initiative for chronic Obstructive Lung Disease; HR=hazard ratio; CI=confidence interval; AUC=area under the curve

to the 2011 scheme was greatest in group D (29.8%) and indistinguishable in groups B and C (15.6% versus 13.3%), whereas in the 2017 scheme, it was again greater in group B than group C (22.5% versus 14.1%). However, in this case, adjusted Cox proportional models (using group A as reference) demonstrated that for both GOLD classification schemes, mortality risk was greater in group B (GOLD 2011, HR 2.6 [95% CI 2.0, 3.3]; GOLD 2017, 3.4 [2.7, 4.1]) than in group C (GOLD 2011, HR 1.8 [95% CI 1.2, 2.6]; GOLD 2017, 1.8 [1.0, 3.1]) (Online data supplement Table 3)

Importantly, even for this alternative definition of respiratory symptoms, GOLD 2011 again showed stronger mortality discrimination by ROC analysis than GOLD 2017 (GOLD 2011, AUC 0.68 versus GOLD 2017, AUC 0.66, *p*<0.001 for the comparison) (Online data supplement Table 3). Once again, prediction of all exacerbation differences between classification schemes was minimal (GOLD 2011, AUC 0.66 versus GOLD 2017, AUC 0.66, *p* = 0.05).

Discussion

This analysis of longitudinal data on 4469 individuals

in the COPDGene[®] cohort, comparing the newly published GOLD 2017 COPD classification to the 2011 criteria, has 4 principal findings. First, though group D continued to have the highest rates of deaths and exacerbations in both classifications, GOLD 2011, using either spirometry or exacerbation history as risk descriptor, provided more accurate mortality discrimination in all ABCD groups. Second, we found a higher mortality rate in group B than in group C in both GOLD 2011 and GOLD 2017. Third, the 2 classifications showed similar overall performance in predicting the frequency of exacerbations, including severe exacerbations. Finally, GOLD 2017 distributed more individuals in our cohort to group B and fewer to group D, compared to GOLD 2011, and in both classifications, group C was the smallest.

Although the GOLD classification system is intended to guide management, several studies over the years have compared the association between GOLD groups and mortality and exacerbation rates. The thrust of this analysis and those of others address whether successive modification of the GOLD criteria is actually producing better prognostic instruments and whether a single scheme identifies those both at risk

for exacerbations and death. Our finding that, in both classifications, mortality was highest in group D and lowest in group A agrees with and extends studies that examined the 2011 criteria,^{2,3,18} or which compared 2007 and 2011 GOLD criteria.^{2,8,19} Similar or better performance for mortality prediction of the 2007 GOLD criteria, relative to the 2011 criteria, was seen in the GenKOLS and UPLIFT cohorts.^{8,11} COCOMICS also found similar mortality rates for 2007 and 2011 criteria at 1, 3 and 10 years.⁹ In agreement, a patient-pooled analysis by Soriano and colleagues also found similar predicted mortality rates for GOLD 2007 and 2011 over a 10-year follow-up period. Collectively, these analyses of earlier GOLD classifications suggested that symptoms and exacerbation history add little to prediction of death.¹⁰

Hence, our finding that GOLD 2011 is a better predictor of mortality than GOLD 2017 is a finding that warrants testing in other cohorts with longitudinal mortality data. Importantly, this disparity persisted when alternate definitions were used for classification, such as using SGRQ for mMRC. Our results agree with an analysis showing that survival rates based on spirometry-only criteria were lower than those based on exacerbation history-only criteria.⁹ Interestingly, in that study, combining spirometry and exacerbation history did yield additively decreased survival rates.⁹ Aligned with these findings of the additive value of information on spirometry to exacerbation history, we found that a model including both GOLD 2017 ABCD groups and spirometry severity could increase the discrimination ability for mortality, obtaining an AUC 0.73 (data not shown), significantly better than ABCD groups alone, 2011 or 2017. Unfortunately, combining these data will create at least 8 groups, which may limit practical utility.

There was only a minimal difference between GOLD 2011 and 2017 when predicting exacerbations, including severe exacerbations, suggesting that simplification of the 2017 ABCD system to exclude spirometric data performs just as well as the 2011 ABCD system in predicting these key health care utilization events. These results align with a large population-based study in Copenhagen by Lange and colleagues, in which they found that the 2011 ABCD classification predicted future exacerbations better than the 2007 criteria, implying that symptoms and exacerbation history are key contributors to the natural history of COPD.³ Lange et al and our findings contrast

to Han and colleagues, who performed an analysis using the same COPDGene[®] cohort as in this study and focused only on the GOLD 2011 classification^{3,20}; they subcategorized individuals in group D into D1 (met FEV₁ criteria only), D2 (met exacerbation criteria only) and D3 (met FEV₁ and exacerbation criteria) and found that those meeting both exacerbation and spirometry criteria had the highest number of and most severe exacerbations.²⁰ It is unlikely that the lack of difference between GOLD 2011 and GOLD 2017 in the current analysis in exacerbation prediction performance results from massive changes in allocation of individuals, as in both, group D had the greatest number of exacerbations, groups B and C were intermediate and group A had the least, as seen in other cohorts. Instead, our current analysis suggests that exacerbation history performs just as well as lung function combined with exacerbation history in predicting future exacerbations.⁵

Our finding that groups B and C had similar annual exacerbation rates, in both GOLD 2011 and GOLD 2017 grouping systems and even when using alternate definitions such as SGRQ, contrasts with prior studies. GOLD group C was the smallest category in our analysis, in agreement with multiple studies comparing GOLD 2007 and 2011.^{8,18,21,22} Group C has been suggested to be superfluous, as its treatment options are very similar to those of group B,¹⁰ and due to crossover of individuals between groups B and C during the 3-year follow-up.² However, our data argue against combining them, as we found a significantly higher mortality rate in group B. The agreement of this finding with 2 studies comparing GOLD 2007 and GOLD 2011 criteria^{8,19} suggests that the mortality difference was unlikely to have resulted solely from the smaller sample size of group C. This mortality gap has been proposed to reflect high rates of cardiovascular disease and malignancy in group B.³ An initial analysis of the COCOMICS cohort found similar mortality rates between these groups at up to 3 years follow-up,⁹ but a separate study using patient-based pooled analysis found a higher mortality risk in group C thereafter.¹⁰

We acknowledge several limitations. While the ABCD grouping aims to identify those at risk for poor outcomes in order to guide therapy, it was not designed specifically as a mortality or exacerbation risk prediction tool. While the GOLD 2011 exhibited significantly better discrimination in predicting mortality, as defined by AUC for receiver-operator

analysis, than the 2017 classification, the confidence intervals for the AUCs do overlap. A third limitation is that we used all-cause, rather than COPD-specific, mortality, an approach also used in other investigations with similar goals. Nevertheless, COPDGene® has used multiple approaches to ascertain vital status of its enrollees, providing strength to our findings. Finally, the GOLD consensus statement is intended for global application to COPD patients, whereas we analyzed a non-population-based sample from 21 U.S. centers, although large enough to have a wide scope and representation of individuals with different disease severity.

In summary, we present an analysis contrasting the new GOLD 2017 and GOLD 2011 classification schema for its association with exacerbation risk and mortality. In our sizeable cohort, the elimination of FEV₁ as a classifier in GOLD 2017 reduced ability to discriminate survival, highlighting the need to consider severity of airflow obstruction in assessing mortality risk. Our findings support the GOLD Scientific Committee decision to remove spirometry for exacerbation risk assessment.

Acknowledgements

Author contributions: MeiLan K. Han and Carlos H. Martinez performed the data analysis. MeiLan K. Han, Carlos H. Martinez, and Susan Murray interpreted the data. Rachel N. Criner, MeiLan K. Han, and Carlos H. Martinez drafted the manuscript. Rachel N. Criner, Meilan K. Han, Carlos H. Martinez, Wassim M. Labaki, Elizabeth A. Regan, Jessica M. Bon, Xavier Soler, Surya P. Bhatt, Susan Murray, John E. Hokanson, Edwin K. Silverman, James D. Crapo, Jeffrey L. Curtis, Fernando J. Martinez, and Barry J. Make revised the final draft. All authors approved the final version to be published.

Declaration of Interest

Fernando J. Martinez reports COPD-related personal fees and non-financial support from the American College of Chest Physicians; grants, personal fees, and non-financial support from AstraZeneca and GlaxoSmithKline; personal fees from Columbia University, Haymarket Communications, Integritas,

Inthought Research, MD Magazine, Methodist Hospital Brooklyn, New York University, Pearl Pharmaceuticals, Unity, UpToDate, WebMD/MedScape, and Western Connecticut Health Network; personal fees and non-financial support from ConCert, Genentech, Inova Fairfax Health System, Continuing Education, Miller Communications, National Association for Continuing Education, Novartis, PeerView, Prime, Puerto Rican Respiratory Society, Chiesi, Roche, Sunovion, Theravance, and Boehringer Ingelheim; and non-financial support from ProterrixBio.

Barry J. Make reports grants from National Heart, Lung, and Blood Institute; personal fees from Novartis, CSL Bering, Verona, Boehringer Ingelheim, Theravance, and Circassia; grants and personal fees from AstraZeneca; receives research funds from Pearl, which is provided to and controlled by National Jewish Health; reviews clinical trial data and serves on the Data and Safety Monitoring Board for Spiration; receives grants from and serves on the advisory board of GlaxoSmithKline; and participates in continuing medical education activities from Consensus Medical Education, Integrity Medical Education, Mt. Sinai Medical Center, WebMD, Up-To-Date, National Jewish Health, SPIRE Learning, American College of Chest Physicians, Projects in Knowledge, Hybrid Communications, Peer Review Institute, Cleveland Clinic, Medscape, and Ult Med Academy.

Surya P. Bhatt and Susan Murray declare grants from the National Institutes of Health. Edwin K. Silverman reports grants from National Institutes of Health, personal fees from Novartis, and grant and travel support from GlaxoSmithKline. Jeffrey L. Curtis reports grants from the National Institutes of Health/ National Heart Lung and Blood Institute, Department of Veterans Affairs, Department of Defense, National Institutes of Health's National Institute of Allergy and Infectious Diseases and MedImmune Corp., Ltd. MeiLan K. Han reports personal fees from GlaxoSmithKline, AstraZeneca, and Boehringer Ingelheim, research support from Sunovion, and non-financial and research support from Novartis. All other authors have nothing to declare.

References

1. Rabe KF, Hurd S, Anzueto A, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med.* 2007;176(6):532-555. doi: <https://doi.org/10.1164/rccm.200703-456SO>
2. Agusti A, Edwards LD, Celli B, et al. Characteristics, stability and outcomes of the 2011 GOLD COPD groups in the ECLIPSE cohort. *Eur Respir J.* 2013;42(3):636-646. doi: <https://doi.org/10.1183/09031936.00195212>
3. 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. doi: <https://doi.org/10.1164/rccm.201207-1299OC>
4. Celli BR, Cote CG, Lareau SC, Meek PM. Predictors of Survival in COPD: more than just the FEV₁. *Respir Med.* 2008;102 Suppl 1:S27-35. doi: [https://doi.org/10.1016/S0954-6111\(08\)70005-2](https://doi.org/10.1016/S0954-6111(08)70005-2)
5. Hurst JR, Vestbo J, Anzueto A, et al. Susceptibility to exacerbation in chronic obstructive pulmonary disease. *New Eng J Med.* 2010;363:1128-1138. doi: <https://doi.org/10.1056/NEJMoa0909883>
6. Vogelmeier CF, Criner GJ, Martinez FJ, et al. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease 2017 Report. GOLD executive summary. *Am J Respir Crit Care Med.* 2017;195(5):557-582. doi: <https://doi.org/10.1164/rccm.201701-0218PP>
7. Vestbo J, Hurd SS, Agusti AG, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med.* 2013;187(4):347-365. doi: <https://doi.org/10.1164/rccm.201204-0596PP>
8. 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.* 2013;188(1):51-59. doi: <https://doi.org/10.1164/rccm.201212-2276OC>
9. 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. doi: <https://doi.org/10.1378/chest.12-1053>
10. Lange P, Celli B, Agusti A, et al. Lung-function trajectories leading to chronic obstructive pulmonary disease. *New Eng J Med.* 2015;373:111-122. doi: <https://doi.org/10.1056/NEJMoa1411532>
11. Goossens LM, Leimer I, Metzdorf N, Becker K, Rutten-van Molken MP. Does the 2013 GOLD classification improve the ability to predict lung function decline, exacerbations and mortality: a post-hoc analysis of the 4-year UPLIFT trial. *BMC Pulm Med.* 2014;14:163. doi: <https://doi.org/10.1186/1471-2466-14-163>
12. Regan EA, Hokanson JE, Murphy JR, et al. Genetic epidemiology of COPD (COPDGeneCOPDGene) study design. *COPD.* 2010;7(1):32-43. doi: <https://doi.org/10.3109/15412550903499522>
13. Ferris BG. Epidemiology standardization project (American Thoracic Society). *Am Rev Respir Dis.* 1978;118:1-120
14. Jones PW, Quirk FH, Baveystock CM, Littlejohns P. A self-complete measure of health status for chronic airflow limitation. The St George's Respiratory Questionnaire. *Am Rev Respir Dis.* 1992;145(6):1321-1327. doi: <https://doi.org/10.1164/ajrccm/145.6.1321>
15. Stewart JI, Moyle S, Criner GJ, et al. Automated telecommunication to obtain longitudinal follow-up in a multicenter cross-sectional COPD study. *COPD.* 2012;9(5):466-472. doi: <https://doi.org/10.3109/15412555.2012.690010>
16. Jones PW, Tabberer M, Chen WH. Creating scenarios of the impact of COPD and their relationship to COPD Assessment Test (CAT) scores. *BMC Pulm Med.* 2011;11:42. doi: <https://doi.org/10.1186/1471-2466-11-42>
17. Jones PW, Brusselle G, Dal Negro RW, et al. Properties of the COPD assessment test in a cross-sectional European study. *Eur Respir J.* 2011;38(1):29-35. doi: <https://doi.org/10.1183/09031936.00177210>
18. Puhan MA, Siebeling L, Zoller M, Muggensturm P, ter Riet G. Simple functional performance tests and mortality in COPD. *Eur Respir J.* 2013;42:956-963. doi: <https://doi.org/10.1183/09031936.00131612>
19. 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. doi: <https://doi.org/10.1136/thoraxjnl-2014-205770>
20. Couper D, Lavange LM, Han M, et al. Design of the subpopulations and intermediate outcomes in COPD Study (SPIROMICS). *Thorax.* 2014;65(5):491-494. doi: <https://doi.org/10.1136/thoraxjnl-2013-203897>
21. Casanova C, Marin JM, Martinez-Gonzalez C, et al. New GOLD classification: longitudinal data on group assignment. *Respir Res.* 2014;15:3. doi: <https://doi.org/10.1186/1465-9921-15-3>
22. Casanova C, Marin JM, Martinez-Gonzalez C, et al. Differential effect of modified Medical Research Council Dyspnea, COPD Assessment Test, and Clinical COPD Questionnaire for symptoms evaluation within the new GOLD staging and mortality in COPD. *Chest.* 2015;148:159-168. doi: <https://doi.org/10.1378/chest.14-2449>