Respiratory Medicine 111 (2016) 30–38

Contents lists available at ScienceDirect

Respiratory Medicine

journal homepage: www.elsevier.com/locate/rmed

Differences between absolute and predicted values of forced expiratory volumes to classify ventilatory impairment in chronic obstructive pulmonary disease

William Checkley [a, b, *], Marilyn G. Foreman [c], Surya P. Bhatt [d], Mark T. Dransfield [d], MeiLan Han [e], Nicola A. Hanania [f], Nadia N. Hansel [a], Elizabeth A. Regan [g], Robert A. Wise [a], COPDGene Study Investigators

[a] Division of Pulmonary and Critical Care, School of Medicine, Johns Hopkins University, Baltimore, USA
[b] Program in Global Disease Epidemiology and Control, Department of International Health, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, USA
[c] Division of Pulmonary and Critical Care Medicine, Morehouse School of Medicine, Atlanta, USA
[d] Division of Pulmonary, Allergy and Critical Care Medicine, School of Medicine, University of Alabama at Birmingham, Birmingham, USA
[e] Division of Pulmonary and Critical Care Medicine, School of Medicine, University of Michigan, Ann Arbor, USA
[f] Section of Pulmonary, Critical Care, and Sleep Medicine, Baylor College of Medicine, Houston, USA
[g] Division of Rheumatology, Department of Medicine, National Jewish Health, Denver, USA

A R T I C L E I N F O

Article history:
Received 17 July 2015
Received in revised form
26 October 2015
Accepted 5 November 2015
Available online 11 November 2015

Keywords:
COPD
Reference equations
Severity criteria

A B S T R A C T

The Global Initiative for Chronic Obstructive Lung Disease (GOLD) severity criterion for COPD is used widely in clinical and research settings; however, it requires the use of ethnic- or population-specific reference equations. We propose two alternative severity criteria based on absolute post-bronchodilator FEV_1 values (FEV_1 and FEV_1/height^2) that do not depend on reference equations. We compared the accuracy of these classification schemasto those based on % predicted values (GOLD criterion) and Z-scores of post-bronchodilator FEV_1 to predict COPD-related functional outcomes or percent emphysema by computerized tomography of the lung. We tested the predictive accuracy of all severity criteria for the 6-minute walk distance (6MWD), St. George's Respiratory Questionnaire (SGRQ), 36-item Short-Form Health Survey physical health component score (SF-36) and the MMRC Dyspnea Score. We used 10-fold cross-validation to estimate average prediction errors and Bonferroni-adjusted t-tests to compare average prediction errors across classification criteria. We analyzed data of 3772 participants with COPD (average age 63 years, 54% male). Severity criteria based on absolute post-bronchodilator FEV_1 or FEV_1/height^2 yielded similar prediction errors for 6MWD, SGRQ, SF-36 physical health component score, and the MMRC Dyspnea Score when compared to the GOLD criterion (all $p > 0.34$); and, had similar predictive accuracy when compared with the Z-scores criterion, with the exception for 6MWD where post-bronchodilator FEV_1 appeared to perform slightly better than Z-scores ($p = 0.01$). Subgroup analyses did not identify differences across severity criteria by race, sex, or age between absolute values and the GOLD criterion or one based on Z-scores. Severity criteria for COPD based on absolute values of post-bronchodilator FEV_1 performed equally as well as did criteria based on predicted values when benchmarked against COPD-related functional and structural outcomes, are simple to use, and may provide a more accessible and comparable approach to severity classification worldwide, especially in settings where prediction equations are not available.

1. Introduction

Several methods have been used to classify severity of ventilatory impairment (i.e., airflow obstruction) in chronic obstructive pulmonary disease (COPD). Currently, the most widely adopted standard to classify ventilatory impairment in COPD is based on the

* Corresponding author. Division of Pulmonary and Critical Care, School of Medicine, Johns Hopkins University, 1800 Orleans Ave Suite 9121, Baltimore, USA.
E-mail address: wcheckl1@jhmi.edu (W. Checkley).

http://dx.doi.org/10.1016/j.rmed.2015.11.004
0954-6111/

Global Initiative for Chronic Obstructive Lung Disease (GOLD) classification which stratifies patients by percent predicted FEV_1 [1]. Similarly, the joint American Thoracic Society and European Respiratory Society guidelines also rank severity of obstructive ventilatory defects using percent predicted forced expiratory values, but use different stratification [2]. These classification schemes are based on the percent predicted FEV_1, which, in turn, depends upon the population being studied and the reference equations that are being used.

The use of reference equations to classify disease severity may lead to difficulties in applying and interpreting this information in both individuals and research populations. Because reference equations differ between laboratories and research studies, people with the same lung function may be categorized differently depending on which reference equations were used. The use of reference equations for stratifying ventilatory impairment may also lead to some confusing conclusions regarding functional impairment. For example, the GOLD staging system using reference equations derived for Caucasians in NHANES would classify a 68-inch tall 75 year-old man with COPD and an FEV_1 of 1.5 L to have moderate disease (Stage II, 53% predicted) whereas a 40 year-old man with COPD of the same height and the same FEV_1 would be classified as having severe disease (Stage III, 38% predicted). Since maximum ventilatory capacity with exercise is determined by multiplying the FEV_1 by 35 [3], we would expect that functional capacity would be similar for a ventilatory-limited individual with the same absolute FEV_1 regardless of age, sex or race. In the above example, the older individual would have similar or greater ventilatory impairment and more functional limitation compared to the younger individual even though the former would be classified as having more severe disease.

Furthermore, using percent predicted to stratify severity of ventilatory impairment may lead to inconsistent assessments of severity across races and sex. Because predicted values of lung function are lower for women and African Americans, compared to men and Caucasians, women and African–Americans might need to have lower lung function to qualify for workers compensation or disability benefits when their degree of impairment is equivalent. Similarly, reference equations derived from resource-poor settings in low- and middle-income countries might underestimate the magnitude of ventilatory impairment in epidemiologic studies [4].

Miller and Pedersen examined spirometric predictors of mortality in a large general population sample, and found that multiples of absolute FEV_1 representing the lowest 1 percentile of the population were a stronger predictor of survival than percent predicted FEV_1 [5]. They also found that absolute FEV_1 values divided by height-cubed or height-squared provided better prediction of survival in a general population and in a COPD study sample than did FEV_1 alone [6,7]. In our proposed approach, we modified the Miller–Pedersen model to determine whether it can be used to classify severity of functional or structural impairment in a COPD population. We sought to test the hypothesis that a simplified classification system for severity of obstructive ventilatory defects based on either absolute FEV_1, or height-adjusted FEV_1 irrespective of age, sex, or race would be predictive of functional limitations, disease impact, quality of life, and severity of emphysema. Moreover, we hypothesized that use of absolute values for the classification of ventilatory impairment would perform similarly in the statistical prediction of COPD-related functional outcomes as would the GOLD classification or one based on Z-scores. We also wanted to know not only about functional characteristics, but also the structural or anatomic measures of COPD. To test these hypotheses, we analyzed data from the COPDGene study which included subjects with a wide range of lung function abnormalities and measures of functional impairment, disease impact, quality of life, and percent emphysema by computerized tomography of the lung.

2. Methods

2.1. Study setting

The COPDGene study is a multicenter investigation of the genetic epidemiology of smoking-related lung disease which recruited 10,300 subjects at 21 clinical centers. Subjects were selected for participation if they were: aged 45–80 years; smoked cigarettes for ≥10 pack-years; and, were willing to undergo testing that included spirometry, chest CT scan, and blood collection for biomarker and genetic analysis [8]. Participants in the COPDGene Study were two-thirds non-Hispanic white and one-third black. All participants provided written informed consent. The study was approved by the Institutional Review Boards at participating institutions.

2.2. Study design

We selected participants from COPDGene who had post-bronchodilator spirometry data available for analysis and identified those who met criteria for COPD. We defined COPD as a post-bronchodilator FEV_1/FVC < lower limit of normal (LLN) using the Global Lung Function Initiative reference equations that adjust for age, sex, height, and race [9]. We defined severity of ventilatory impairment using four different criteria (Table 1). The first was based on % predicted values of FEV_1 [1] using the Global Lung Function Initiative reference equations that adjust for age, sex, height, and race [9]. The second and third were based on absolute values of FEV_1 and height-adjusted FEV_1, respectively. The last criterion was based on post-bronchodilator FEV_1 Z-score thresholds [10], using the Global Lung Function Initiative reference equations that adjust for age, sex, height, and race [9]. We then selected *a priori* four COPD-related functional outcomes for evaluation: 6-minute walk distance (6MWD), St. George's Respiratory Questionnaire (SGRQ), 36-item Short-Form Health Survey (SF-36) physical health component score and the Modified Medical Research Council (MMRC) Dyspnea Score. We also examined percent emphysema on chest CT scan as the COPD-related structural outcome. Chest CT scans were performed using multi-detector helical CT scanners with 16 or more detectors. Protocols for scanner types have been published [11]. Emphysema and airway disease severity and distribution were obtained from the inspiratory CT acquisition and air trapping from the expiratory acquisition. SLICER software (www.slicer.org) was used to calculate percent

Table 1
Approaches to classify ventilatory impairment.

	Proposed approaches to classify ventilatory impairment		Z-scores to classify ventilatory impairment	GOLD guidelines to classify ventilatory impairment
	FEV_1 (L)	FEV_1/height^2 (L/m^2)	FEV_1 Z-score	% Predicted FEV_1
Mild	≥2	≥0.8	>−2	≥80
Moderate	1 to 1.99	0.6 to 0.79	−2.99 to −2.0	50 to 79
Severe	0.5 to 0.99	0.4 to 0.59	−3.99 to −3.0	30 to 49
Very Severe	0 to 0.49	0 to 0.39	≤−4	0 to 29