**Severity of airflow obstruction in Chronic Obstructive Pulmonary Disease (COPD): Proposal for a new classification.**

Sonia Coton,1 William M. Vollmer,2 Eric Bateman,3 Guy B. Marks,4 Wan Tan,5 Filip Mejza,6 Sanjay Juvekar,7 Christer Janson,8 Kevin Mortimer,9 Mahesh PA,10 A. Sonia Buist,11 Peter G.J. Burney12 for the Burden of Obstructive Lung Disease Study investigators.13

1. Research Department of Primary Care and Population Health, University College, London, UK
2. Kaiser Permanente Center for Health Research, Portland, OR, USA
3. Division of Pulmonology, Department of Medicine, University of Cape Town, Cape Town, South Africa
4. Woolcock Institute of Medical Research and South Western Sydney Clinical School, UNSW, Sydney, Australia
5. iCapture Center for Cardiovascular and Pulmonary Research, University of British Columbia, Vancouver, BC, Canada
6. II Department of Internal Medicine, Jagiellonian University Medical College, Cracow, Poland
7. Vadu HDSS, KEM Hospital Research Centre, Pune, India
8. Department of Medical Sciences: Respiratory, Allergy and Sleep Research, Uppsala University, Uppsala, Sweden
9. Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool, UK
10. Department of Pulmonary Medicine, JSS Medical College, Mysore, India
11. Oregon Health and Sciences University, Portland, OR, USA
12. National Heart and Lung Institute, Imperial College, London, UK
13. A full list of the BOLD investigators can be found at <http://www.boldstudy.org/sites.html>

Key words: (4-6) Spirometry; Method Comparison; Epidemiological Studies; GOLD Classification

Address for correspondence: Prof. Peter Burney,

National Heart & Lung Institute

1 Manresa Road,

London SW3 6LR

London UK

Tel: +44 (0)207 594 7942

e-mail: [p.burney@imperial.ac.uk](mailto:p.burney@imperial.ac.uk)

Abstract: 198

Words: 2923

Table: 6

Figures: 2

Supplementary Tables: 2

References: 20

**Abstract**

Current classifications of Chronic Obstructive Pulmonary Disease (COPD) severity are complex and do not grade levels of obstruction. Obstruction is a simpler construct and independent of ethnicity.

Using data from the Burden of Obstructive Lung Disease (BOLD) study we constructed an index of severity based on the FEV1/FVC ratio with cut-points dividing the population intofour similar sized strata to those created by the GOLD criteria using FEV1. We measured agreement between classifications and validity of the FEV1-based classification in identifying the level of obstruction as defined by the new groupings. We compared the strengths of association of each classification with quality of life, MRC dyspnoea score and the self-reported exacerbation rate.

Agreement between classifications was only fair. FEV1-based criteria for moderate COPD identified only 79% of those with moderate obstruction and misclassified half of the participants with mild obstruction as having more severe COPD. Both scales were equally strongly associated with quality of life, exertional dyspnoea and respiratory exacerbations.

Severity assessed using the FEV1/FVC ratio is only in moderate agreement with the severity assessed using FEV1 but is equally strongly associated with other outcomes. Severity assessed using the FEV1/FVC ratio is likely to be independent of ethnicity.

[198 words]

2923 words

**Background**

Although Chronic Obstructive Pulmonary Disease (COPD) is associated with chronic airflow obstruction (CAO), the terms are not synonymous. Chronic airflow obstruction is a physiological state that can be measured, even if the best criteria for its definition are still disputed. ([1](#_ENREF_1)) COPD, on the other hand, remains a clinical condition that has been described rather than defined.(2) It has been shown empirically that survival is strongly associated with the one-second Forced Expiratory Volume (FEV1) and most indicators of COPD severity, including GOLD,([2](#_ENREF_2)) BODE,([3](#_ENREF_3)) ADO([4](#_ENREF_4)) and the method proposed by Miller([5](#_ENREF_5)) use this as at least part of their scoring.

The FEV1 is a good marker of severity because it is highly correlated with the FVC and the total lung capacity (TLC) and these are better predictors of outcome than the level of obstruction. ([6-8](#_ENREF_6)) Nevertheless there are also benefits in quantifying the severity of the obstruction itself. First this provides clearer differentiation between severity of COPD assessed using the FEV1 and severity of CAO assessed using the FEV1/FVC ratio. Second, because the FEV1/FVC ratio, unlike either the FEV1 or the FVC, is independent of ethnicity, such a measure would be valid across ethnic groups.([9](#_ENREF_9), [10](#_ENREF_10))

Using data from the Burden of Obstructive Lung Disease (BOLD) study we have constructed an index of severity based on the FEV1/FVC ratio with cut-points that divide the population intofour strata of similar size to those created by application of the GOLD spirometric criteria using the FEV1. We have compared these two indices of severity, we have tested the extent to which the FEV1-based criteria relate to the severity of obstruction as assessed by the FEV1/FVC ratio, and we have compared construct validity of the two scores by assessing their association with measures of mental and physical quality of life.

**Methods**

**BOLD Study Methods**

The methods of the BOLD study are described elsewhere.([11](#_ENREF_11)) Sites were asked to recruit a minimum of 600 non-institutionalised adults aged ≥40 years; using population-based sampling plans. All study sites used standardised questionnaires to collect information on age, sex, quality of life, medical history and respiratory symptoms. The questionnaires were translated following the International Quality of Life Assessment (IQOLA) project protocol and agreed versions were administered in the participant’s native language in face-to-face interviews, by trained certified staff. Responders were defined as those who had questionnaire data and post-bronchodilator spirometry regardless of quality.

Pre- and post-bronchodilator spirometry were measured with the ndd EasyOne spirometer, using 200 micrograms of inhaled salbutamol by spacer and a 15 minute interval. Spirometry was performed exclusively by trained, certified technicians who received regular feedback on quality to maintain pre-set quality standards during the study. Only tests that met the acceptability and repeatability criteria from the American Thoracic Society (ATS) and European Respiratory Society (ERS) were included for analyses.([12](#_ENREF_12)) The analyses in this paper use only the post-bronchodilator results.

Ethical approval was obtained by each site from the local ethical committee and written informed consent was obtained from every participant.

**Chronic airways obstruction severity definitions**

Chronic airways obstruction was defined as a post-bronchodilator FEV1/FVC ratio below the lower limit of normal (LLN) for age and sex, based on reference equations for Caucasians derived from the third US National Health and Nutrition Examination Survey (NHANES III).([13](#_ENREF_13)) Severity of COPD was graded using post-bronchodilator percent predicted FEV1­ with cut-offs defined as: mild; ≥80%, moderate; <80%≥50%, severe; <50%≥30%, and very severe; <30%. We refer to this as “FEV1-based staging” and the definitions of the stages are derived from the GOLD recommendations.

We generated the FEV1/FVC-based severity criteria based entirely on the post-bronchodilator FEV1/FVC ratio. For participants with acceptable ([12](#_ENREF_12)) post-bronchodilator spirometry we calculated the percentile value for post-bronchodilator FEV1/FVC ratio, using pooled data from the BOLD centres and the regression estimates from Hankinson et al 1999.([13](#_ENREF_13)) The FEV1/FVC Caucasian prediction equation coefficients for: age (b1), the prediction intercept (b0), and the LLN intercept were extracted directly from this paper,([13](#_ENREF_13)) and we computed the mean square error (MSE) of the FEV1/FVC Caucasian prediction equation for males and females separately. The regression coefficients and MSE are presented in Table 1.

The ith FEV1/FVC percentile was calculated using equation (1). Using an iterative process a participant was assigned the ith percentile if the observed post-bronchodilator FEV1/FVC ratio was greater than ith-1 percentile and less than the ith percentile. This was done for males and females separately.

Secondly, we chose FEV1/FVC percentile cut-points so that the proportion of participants in each newly-defined group matched the proportion of participants in the categories of the GOLD staging variable. For example, if there were 2% of participants in the very severe GOLD staging category, we would choose the FEV1/FVC percentile that defined 2% of participants with the lowest ratio. This classification is referred to as the ‘FEV1/FVC-based staging’ throughout the paper.

**Health outcomes definitions**

Dyspnoea was defined according to the modified MRC 0-4 scale; 0 for normal, 4 for being unable to leave the house due to breathlessness.

Quality of life (QoL) was assessed by participant response to the Medical Outcomes Study - Short Form 12 questionnaire (SF-12). For the purposes of this analysis, we have used the summary scores for: mental health (MCS), and physical health (PCS) derived according to standard SF-12 algorithms. ([14](#_ENREF_14)). The summary scores range from 0 to 100; with 100 representing the best quality of life.

Respiratory exacerbations were defined as a history of at least one reported episode of breathing problems that interfered with usual daily activities or caused the participant to miss work in the previous 12 months.

**Statistical Methods**

Agreement between FEV1-based and FEV1/FVC-based staging was assessed using the kappa statistic. Unweighted, linear, and quadratic weighting schemes were employed; the weighting schemes applied penalties as agreement became further apart. The weights assigned to classifications differing by (0, 1, 2 and 3) categories were: (1, 0.667, 0.33, and 0) for linear; and (1, 0.889, 0.556, and 0) for quadratic, respectively. The linear and quadratic weighting schemes were applied using pre-recorded weights available within the Stata Kappa command. ([15](#_ENREF_15))

The FEV1/FVC-based staging was used as the reference from which to assess the sensitivity and specificity of each FEV1-based stage when classifying the severity of obstruction.

The relationships between each classification and the health-related outcomes: dyspnoea, QoL (MSC and PCS), and presence or absence of a “respiratory exacerbation”, were explored with regression modelling. Models were fitted separately for each classification and each health-related outcome. Linear regression modelling was used to estimate the relationship between each classification and reported dyspnoea, mental health (MCS QoL) and physical health (PCS QoL). Negative binomial regression modelling was used to estimate the relationship between each classification and respiratory exacerbations, adjusting for over dispersion. Identical regression models were fitted separately for each site, allowing for sampling weights and survey design. The results were subsequently pooled across all sites and average effects for each outcome were estimated using random effects meta-analysis. Heterogeneity was summarised using the I2 statistic.

All analyses were performed using Stata 12 (Stata Corporation, College Station, TX, USA).

**Results**

Twenty-four BOLD sites were included and a total of 20,240 participants from all sites were classified as responders (had core questionnaire data and some spirometry regardless of quality). Of the responders 16,996 (84%) had acceptable post-bronchodilator spirometry and complete questionnaire data. A total of 1,993 participants had spirometrically confirmed chronic airflow obstruction defined as post-bronchodilator FEV1/FVC ratio below LLN, these participants were the study population for this analysis.

The severity of COPD in participants with spirometrically confirmed chronic airflow obstruction was graded according to the FEV1-based staging criteria: 611 (31%) mild, 999 (50%) moderate, 326 (16%) severe and 57 (3%) very severe (Table 2) Men and women had a similar prevalence of obstruction across categories of the FEV1-based staging. The highest prevalence occurred in the moderate category: 51% in men and 49% in women.

The FEV1/FVC percentile was calculated for all participants with confirmed chronic airflow obstruction, using equation (1). Initially we used iterative steps of 0.001 to calculate the FEV1/FVC percentiles. However, the small proportion of participants in the very severe FEV1-based staging category required very small iterative steps to distinguish the cut points. Percentiles between zero and 1x10-8 were assigned in steps of 1x10-12; percentiles between 1x10-8 and 3x10-5 were assigned in steps of 1x10-8; and percentiles between 3x10-5 and 0.05 were assigned in steps of 1x10-5.

The FEV1/FVC percentile cut points chosen to match the upper limit of the proportion of participants in each FEV1-based staging category were calculated as mild = 5th percentile; moderate = 1.78th percentile; severe = 0.0027th percentile; and very severe = 3.0thx10-10 percentile. These cut-points were used to define the FEV1/FVC-based staging in equation (1).

By design the marginal distributions within categories are the same for the FEV1-based and FEV1/FVC-based staging (Table 3), but the two criteria classify different individuals as: mild, moderate, severe and very severe. The FEV1-based staging criteria agreed 54%, 60%, 47%, and 49% percent of the time with the FEV1/FVC-based staging criteria of mild, moderate, severe, and very severe, respectively (Table 3). Overall agreement between the two criteria was 55% (unweighted), 85% (linear weighting), and 94% (quadratic weighting). The corresponding kappa statistics were 0.30, 0.42 and 0.57, respectively, showing fair to moderate agreement between the two classifications (Table 4).

Considering the FEV1/FVC-based staging as the diagnostic standard for the severity of obstruction we calculated the sensitivity and specificity of the FEV1-based staging. The likelihood that the FEV1-based criteria would identify at least a given level of obstruction fell from 79.5% (95% CI: 77.3%, 81.6%) for moderate or greater obstruction, to 61.4% (56.3%, 66.3%) with severe or greater obstruction, to 50.8% (37.3%, 64.4%) for very severe obstruction. Conversely the FEV1-based criteria were less likely to underestimate the level of obstruction as levels of obstruction increased (Table 5). The specificity (and 95% CI) of the FEV1-based staging was; 53.5% (49.5%, 57.5%), 90.8% (89.3%, 92.2%) and 98.6% (97.9%, 99.0%) for moderate or greater obstruction; severe or greater obstruction; and very severe obstruction, respectively.

The two criteria had the same ability to predict dyspnoea, MCS QoL, PCS QoL and respiratory exacerbations (Figure 1). Predicted MCS QoL and PCS QoL declined with increasing severity for both classifications. Predicted dyspnoea increased with increasing severity for both classifications. The predicted log count of respiratory exacerbations did not indicate any pattern with either classification. The I2 statistic for heterogeneity was large for all the models fitted, ranging from 22% to 99.9%. The p-value for the I2 statistic indicated statistical significance in all but 4 estimates. These estimates were: FEV1/FVC-based staging mild and moderate obstruction with QoL PCS; FEV1-based staging mild obstruction with QoL MCS; and FEV1-based staging moderate obstruction with respiratory exacerbations. (Figure 1)

Table 6 and Figure 2 give the criteria for the FEV1/FVC-based staging cut points. The FEV1/FVC-based staging cut points decreased with increasing age for both males and females, reflecting the natural decrease in lung function as age increases. The FEV1/FVC-based staging cut points were slightly lower for males when compared to females.

**Discussion**

We have established spirometric criteria for classifying the severity of airflow obstruction using the FEV1/FVC ratio which divides the population into same sized groups as the FEV1-based classification suggested by GOLD. We have shown that the classifications of individuals on the two scales show only moderate agreement, but that both scales have similar associations with quality of life, exertional dyspnoea and reported exacerbations. Finally we have provided both equations and nomograms for each sex separately indicating the limits to the mild, moderate, severe and very severe grades of obstruction by age.

The advantages of the new classification are that it avoids confusion between the severity of obstruction and the severity of spirometric restriction and it avoids the difficulties that arise when comparing the lung volumes of people from different ethnic backgrounds.

FEV1 is almost always included in severity scores for COPD. The FEV1, however, does not by itself measure the level of obstruction as it may be reduced in either obstructive or restrictive conditions. The assumption has been that restrictive disease is so rare that it can be ignored, and that because the FEV1 is both easier and more accurately recorded than the FEV1/FVC ratio, the FEV1 can be used as a surrogate measure of obstruction. Fletcher et al. used FEV1/Height3 as the main outcomein their monograph on the natural history of chronic bronchitis and emphysema, ([16](#_ENREF_16)) because they suspected some undetected technical error in their measurements of vital capacity (page 57 in reference 16), and FEV1 is a reasonable marker of increasing obstruction in a longitudinal study. However the FEV1 is a poor marker of obstruction in cross-sectional studies because the FEV1 is more influenced by variations in the FVC than by variations in the FEV1/FVC ratio. In BOLD study centres the correlation coefficients for FEV1 against FVC are typically above 0.9 (supplementary Figure 1), those for FEV1/FVC against FVC are typically below 0.1 (supplementary Figure 2).

The FEV1/FVC ratio is independent of ethnicity ([9](#_ENREF_9), [10](#_ENREF_10)) unlike the FEV1. There is disagreement over the interpretation of the lower lung volumes found in non-European populations. Some argue that these are so common as to be “normal”([10](#_ENREF_10)) while others argue that a low FVC still represents a disadvantage regardless of ethnicity.([17](#_ENREF_17)) When looking for a classification of CAO there is benefit in using a scale that avoids this area of uncertainty.

Our data come from a large study that has highly standardised and quality assured lung function data. All the curves have been individually checked by a central monitoring group and the criteria for adequate spirometry suggested by the ATS and ERS then applied.([12](#_ENREF_12)) The data are collected across very diverse populations.

Although the relations of the two different scales to the other markers of severity, physical and mental quality of life, dyspnoea and a history of respiratory exacerbation were very similar, they varied markedly between centres. This reflects differences between centres in the threshold values of lung function at which participants report symptoms. This is not surprising as we would expect these threshold values to be culturally dependent. Nevertheless this variability indicates that the mean values reported here for each stage should not be interpreted as the expected value at any particular site, let alone as a true global mean value.

An alternative to the FEV1/FVC-based staging presented in this paper would be to use the predictive values for the FEV1/FVC, which are readily available from most spirometers. This is similar to the approach made by GOLD in using the FEV1% predicted to stage COPD severity. The disadvantage of this method is, however, that cut-offs based on the percentage of a predicted value are affected by age. As age increases and predicted value falls the absolute value represented by a percentage of the predicted value also falls. A centile approach, as used here, provides a more consistent measure across age groups.([1](#_ENREF_1))

We have not assessed the association with mortality as the data are not yet available for the BOLD participants. Evidence from several studies, however, suggests that FVC is an important determinant of survival,([6](#_ENREF_6), [18](#_ENREF_18)) and the Atheroma Risk in Communities (ARIC) study data suggests that after adjusting for smoking history the FEV1/FVC ratio is a comparatively poor predictor of mortality.([7](#_ENREF_7)) Using FEV1 as a marker of severity in COPD, as in the GOLD,([2](#_ENREF_2)) BODE([3](#_ENREF_3)) and ADO([4](#_ENREF_4)) scores, works well precisely because it is well correlated with the FVC and not, as often supposed, because it is itself a marker of obstruction.

As obstruction increases it becomes more difficult to empty the lungs, the residual volume of air in the lungs at the end of expiration increases, the forced vital capacity declines and there is a paradoxical increase in the FEV1/FVC ratio. This is frequently used to argue against using the ratio as a marker of obstruction. However, in a stratified selection of subjects from a BOLD follow up study we measured FEV1, FVC and TLC for 111 participants.([19](#_ENREF_19)) The FEV1/TLC and the FEV1/FVC ratio, as expected, were highly correlated (r=0.83) and there was no association between the difference in ratios and their mean value (=-0.06; p=0.32), (unpublished observations). This suggests that although there is a difference between the FEV1/FVC ratio and the FEV1/TLC ratio this difference does not vary with the level of obstruction and that in this general population survey the FEV1/FVC ratio is a reasonable index of airway obstruction.([20](#_ENREF_20)) In clinical practice any discrepancy is likely to be more relevant, but the problem in principle is an issue for any interpretation of the FVC or of the FEV1/FVC ratio. Relative to the error introduced by an assumption that a low FEV1 represents airflow obstruction, air trapping is a minor issue in population surveys.

We have shown that the FEV1-based spirometric criteria are neither sensitive nor specific for identifying levels of obstruction. The new classification of obstruction based on the FEV1/FVC ratio is likely to be independent of ethnic differences and has a similar relation to quality of life, dyspnoea and exacerbation rates as does the FEV1-based classification, but based on evidence from other studies is likely to be less good at predicting mortality. A score based on FVC([6](#_ENREF_6), [7](#_ENREF_7)) or total lung capacity([8](#_ENREF_8)) would probably be even better than the current scores based on FEV1, in predicting mortality. The new index is untested as an indicator of the need to escalate treatment. For the moment at least it should not supplant, though it might supplement, the FEV1 as a biomarker of response in clinical trials.

**Acknowledgments**

FUNDING: This work was funded by the Wellcome Trust (085790/Z/08/Z). Details of other sponsors of the BOLD study and study sites can be found at the BOLD website [http://www.boldstudy.org/sponsors.html]. The sponsors of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

Data for these analyses were contributed by the following investigators: NanShan Zhong (PI), Shengming Liu, Jiachun Lu, Pixin Ran, Dali Wang, Jingping Zheng, Yumin Zhou (Guangzhou Institute of Respiratory Diseases, Guangzhou Medical College, Guangzhou, China); Ali Kocabaş (PI), Attila Hancioglu, Ismail Hanta, Sedat Kuleci, Ahmet Sinan Turkyilmaz, Sema Umut, Turgay Unalan (Cukurova University School of Medicine, Department of Chest Diseases, Adana, Turkey); Michael Studnicka (PI), Torkil Dawes, Bernd Lamprecht, Lea Schirhofer (Paracelsus Medical University, Department of Pulmonary Medicine, Salzburg Austria); Eric Bateman (PI), Anamika Jithoo (PI), Desiree Adams, Edward Barnes, Jasper Freeman, Anton Hayes, Sipho Hlengwa, Christine Johannisen, Mariana Koopman, Innocentia Louw, Ina Ludick, Alta Olckers, Johanna Ryck, Janita Storbeck, (University of Cape Town Lung Institute, Cape Town, South Africa); Thorarinn Gislason (PI), Bryndis Benedikdtsdottir, Kristin Jörundsdottir, Lovisa Gudmundsdottir, Sigrun Gudmundsdottir, Gunnar Gundmundsson, (Landspitali University Hospital, Dept. of Allergy, Respiratory Medicine and Sleep, Reykjavik, Iceland); Ewa Nizankowska-Mogilnicka (PI) , Jakub Frey, Rafal Harat, Filip Mejza, Pawel Nastalek, Andrzej Pajak, Wojciech Skucha, Andrzej Szczeklik,Magda Twardowska, (Division of Pulmonary Diseases, Department of Medicine, Jagiellonian University School of Medicine, Cracow, Poland); Tobias Welte (PI), Isabelle Bodemann, Henning Geldmacher, Alexandra Schweda-Linow (Hannover Medical School, Hannover, Germany); Amund Gulsvik (PI), Tina Endresen, Lene Svendsen (Department of Thoracic Medicine, Institute of Medicine, University of Bergen, Bergen, Norway); Wan C. Tan (PI), Wen Wang (iCapture Center for Cardiovascular and Pulmonary Research, University of British Columbia, Vancouver, BC, Canada); David M. Mannino (PI), John Cain, Rebecca Copeland, Dana Hazen, Jennifer Methvin, (University of Kentucky, Lexington, Kentucky, USA); Renato B. Dantes (PI), Lourdes Amarillo, Lakan U. Berratio, Lenora C. Fernandez, Norberto A. Francisco, Gerard S. Garcia, Teresita S. de Guia, Luisito F. Idolor, Sullian S. Naval, Thessa Reyes, Camilo C. Roa, Jr., Ma. Flordeliza Sanchez, Leander P. Simpao (Philippine College of Chest Physicians, Manila, Philippines); Christine Jenkins (PI), Guy Marks (PI), Tessa Bird, Paola Espinel, Kate Hardaker, Brett Toelle (Woolcock Institute of Medical Research, Sydney, Australia), Peter GJ Burney (PI), Caron Amor, James Potts, Michael Tumilty, Fiona McLean (National Heart and Lung Institute, Imperial College, London), E.F.M. Wouters, G.J. Wesseling (Maastricht University Medical Center, Maastricht, the Netherlands), Cristina Bárbara (PI), Fátima Rodrigues, Hermínia Dias, João Cardoso, João Almeida, Maria João Matos, Paula Simão, Moutinho Santos, Reis Ferreira (The Portuguese Society of Pneumology, Lisbon, Portugal), Christer Janson (PI), Inga Sif Olafsdottir, Katarina Nisser, Ulrike Spetz-Nyström, Gunilla Hägg and Gun-Marie Lund (Department of Medical Sciences: Respiratory Medicine & Allergology, Uppsala University, Sweden), Rain Jõgi (PI), Hendrik Laja, Katrin Ulst, Vappu Zobel, Toomas-Julius Lill (Lung Clinic, Tartu University Hospital), Parvaiz A Koul (PI), Sajjad Malik, Nissar A Hakim, Umar Hafiz Khan (Sher-i-Kashmir Institute of Medical Sciences, Srinagar, J&K, India); Rohini Chowgule (PI)Vasant Shetye, Jonelle Raphael, Rosel Almeda, Mahesh Tawde, Rafiq Tadvi, Sunil Katkar, Milind Kadam, Rupesh Dhanawade, Umesh Ghurup (Indian Institute of Environmental Medicine, Mumbai, India); Imed Harrabi (PI), Myriam Denguezli, Zouhair Tabka, Hager Daldoul, Zaki Boukheroufa, Firas Chouikha, Wahbi Belhaj Khalifa (Faculté de Médecine, Sousse, Tunisia); Luisito F. Idolor (PI), Teresita S. de Guia, Norberto A. Francisco, Camilo C. Roa, Fernando G. Ayuyao, Cecil Z.Tady, Daniel T. Tan, Sylvia Banal-Yang, Vincent M. Balanag, Jr., Maria Teresita N. Reyes, Renato. B. Dantes (Lung Centre of the Philippines, Philippine General Hospital, Nampicuan&Talugtug, Philippines); Sanjay Juvekar (PI), Siddhi Hirve, Somnath Sambhudas, Bharat Chaidhary, Meera Tambe, Savita Pingale, Arati Umap, Archana Umap, Nitin Shelar, Sampada Devchakke, Sharda Chaudhary, Suvarna Bondre, Savita Walke, Ashleshsa Gawhane, Anil Sapkal, Rupali Argade, Vijay Gaikwad (Vadu HDSS, KEM Hospital Research Centre Pune, Pune India); Sundeep Salvi (PI), Bill Brashier, Jyoti Londhe, Sapna Madas (Chest Research Foundation (CRF), Pune India); Mohamed C Benjelloun (PI), Chakib Nejjari, Mohamed Elbiaze, Karima El Rhazi (Laboratoire d’épidémiologie, Recherche Clinique et Santé Communautaire, Fès, Morroco); Daniel Obaseki (PI), Gregory Erhabor, Olayemi Awopeju, Olufemi Adewole (Obafemi Awolowo University, Ile-Ife, Nigeria).

We would like to thank all the study participants who generously gave their time to the study.

**Declaration of Interest statement**

None of the authors have any material conflicts of interest in relation to this paper.

Reference List

1. Swanney MP, Ruppel G, Enright PL, Pedersen OF, Crapo RO, Miller MR, et al. Using the lower limit of normal for the FEV1/FVC ratio reduces the misclassification of airway obstruction. Thorax [Internet]. 2008; 63(12):[1046-51 pp.].

2. Global Initiative for Chronic Obstructive Lung D. Global Strategy for the diagnosis, management and prevention of chronic obstructive pulmonary disease.: Medical Communications Resources, Inc; 2014 [13th October 2014]. 94]. Available from: <http://www.goldcopd.org/uploads/users/files/GOLD_Report_2014_Jun11.pdf>.

3. Celli B, Cote C, Marin J, Casanova C, de Oca M, Mendez R, et al. The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. New England Journal of Medicine. 2004;350(10):1005.

4. Puhan M, Garcia-Aymerich J, Frey M, Riet G, Antó J, Agustí A, et al. Expansion of the prognostic assessment of patients with chronic obstructive pulmonary disease: the updated BODE index and the ADO index. The Lancet. 2009;374(9691):704-11.

5. Miller MR, Pedersen OF, Dirksen A. A new staging strategy for chronic obstructive pulmonary disease. International Journal of COPD. 2007;2(4):657-63.

6. Fried LP, Kronmal RA, Newman AB, Bild DE, Mittelmark MB, Polak JF, et al. Risk factors for 5-year mortality in older adults: the Cardiovascular Health Study. JAMA. 1998;279(8):585-92.

7. Burney PGJ, Hooper R. Forced vital capacity, airway obstruction and survival in a general population sample from the USA. Thorax. 2011;66(1):49-54.

8. Pedone C, Scarlata S, Chiurco D, Conte ME, Forastiere F, Antonelli-Incalzi R. Association of reduced total lung capacity with mortality and use of health services  Chest. 2012;141(4):1025-30.

9. Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, et al. Interpretative strategies for lung function tests. The European respiratory journal. 2005;26(5):948.

10. Quanjer PH. EDITORIAL: Lung function, race and ethnicity: a conundrum. Eur Respir J 2013; 41: 1249–1251. 2013;41:1249-51.

11. Buist AS, Vollmer WM, Sullivan SD, Weiss KB, Lee TA, Menezes AM, et al. The Burden of Obstructive Lung Disease Initiative (BOLD): rationale and design. COPD: Journal of Chronic Obstructive Pulmonary Disease. 2005;2(2):277-83.

12. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al. Standardisation of spirometry. Eur Respir J [Internet]. 2005; 26(2):[153-61 pp.].

13. 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):179-87.

14. Ware JE, Kosinski M, Keller SD. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. . Med Care. 1996;34(3):220-33.

15. Gould W. Interrater Agreement. Stata Tech Bull. 1997;40(November):2-8.

16. Fletcher C, Peto R, Tinker C, Speizer FE. The natural history of chronic bronchitis and emphysema. An eight-year study of early chronic obstructive lung disease in working men in London. Oxford: Oxford University Press; 1976 1976.

17. Burney PGJ, Hooper RJ. The use of ethnically specific norms for ventilatory function in African-American and white populations. Int J Epidemiol. 2012;41(3):782-90.

18. Kannel WB, Lew EA, Hubert HB, Castelli WP. The value of measuring vital capacity for prognositc purposes. Trans Assoc Life Insur Med Dir Am 1980;64:66-83.

19. Amaral AFS, Patel J, Gnatiuc L, Jones M, Burney PGJ. Association of pulse wave velocity with total lung capacity: A cross-sectional analysis of the BOLD London study. Resp Med. 2015;109:1569-75.

20. Bland JM, Altman DG. Measuring agreement in method comparison studies. Statistical Methods in Medical Research. 1999;8:135-60.

**Tables and Figures**

Table 1: Prediction equation coefficients for FEV1/FVC for males and females (9)

|  |  |  |  |
| --- | --- | --- | --- |
|  | Mean Square Error | b0 | b1 |
| Males | 5.883 | 88.066 | -0.2066 |
| Females | 5.954 | 90.809 | -0.2125 |
| b0 – value at age 20 (predicted intercept)  b1 – change in the ratio with each year (coefficient for age) | | | |

Table 2: Prevalence of FEV1-based stages\* by sex in all BOLD sites

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **GOLD staging** | | | | |
| **Sex** | Mild | Moderate | Severe | Very severe | Total |
| Male | 297 (28%) | 536 (51%) | 179 (17%) | 32  (3%) | 1044 (100%) |
| Female | 314 (33%) | 463 (49%) | 147 (15%) | 25  (3%) | 949 (100%) |
| Total | 611 (31%) | 999 (50%) | 326 (16%) | 57  (3%) | 1,993 (100%) |
| \* defined as: mild = post-bronchodilator FEV1>80%, moderate = post-bronchodilator FEV1 (50%, 80%], severe = post-bronchodilator FEV1 (30% ,50%] and very severe = post-bronchodilator FEV1<30% in subjects with post -bronchodilator FEV1/FVC<LLN | | | | | |

Table 3: Cross tabulation between the FEV1/FVC-based staging\* and FEV1-based\*\* staging for chronic airflow obstruction across all BOLD sites in subjects with post bronchodilator FEV1/FVC ratio below the lower limit of normal

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **FEV1/FVC-based Staging** | | | | | | | | | |
| **GOLD Staging** |  | Mild  (n=609) | Moderate  (n=1000) | | | Severe  (n=325) | | Very severe  (n=59) | | Total  (n=1993) |
| Mild (N=609) | 54% | | 28% | | | 2% | | 0% | 31% |
| Moderate (N=1000) | 42% | | 60% | | | 44% | | 2% | 50% |
| Severe (N=325) | 4% | | 12% | | | 47% | | 49% | 16% |
| Very severe N=59) | 0.2% | | 0.2% | | | 8% | | 49% | 3% |
| Total | 100% | | | 100% | 100% | | | 100% | 100% |
| \* FEV1/FVC-based staging of chronic airflow obstruction defined as: mild = post bronchodilator FEV1/FVC<5th percentile; moderate = post-bronchodilator FEV1/FVC< 1.78th percentile; severe = post -bronchodilator FEV1/FVC<0.0027th percentile; and very severe = post-bronchodilator FEV1/FVC<3.0thx10-10 percentile  \*\* FEV1-based stages defined as: mild = post-bronchodilator FEV1>80%, moderate = post-bronchodilator FEV1 (50%, 80%], severe = post-bronchodilator FEV1 (30% ,50%] and very severe = post-bronchodilator FEV1<30% in subjects with post-bronchodilator FEV1/FVC<LLN | | | | | | | | | | |

Table 4: Kappa statistics for agreement between the FEV1-based and the FEV1/FVC-based classifications

|  |  |  |
| --- | --- | --- |
| **Weighting scheme** | **Overall agreement** | **Kappa statistic** |
| No weight | 55% | 0.29 |
| Linear weight | 85% | 0.42 |
| Quadratic weight | 94% | 0.57 |

Table 5: Sensitivity and specificity of the FEV1-based stages compared to the FEV1/FVC-based stages

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Severity cut point** | **Sensitivity** | **95% CI** | **Specificity** | **95% CI** |
| Moderate or greater | 79.5% | (77.3%, 81.6%) | 53.5% | (49.5%, 57.5%) |
| Severe or greater | 61.4% | (56.3%, 66.3%) | 90.8% | (89.3%, 92.2%) |
| Very severe | 50.8% | (37.3%, 64.4%) | 98.6% | (97.9%, 99.0%) |

Table 6: FEV1/FVC-based chronic airflow obstruction staging classification by sex

|  |  |  |
| --- | --- | --- |
|  | **Men** | **Women** |
| Normal | FEV1/FVC ≥ 78.39 - 0.2066 x Age | FEV1/FVC ≥ 81.02 - 0.2125 x Age |
| Mild | FEV1/FVC < 78.39 - 0.2066 x Age | FEV1/FVC < 81.02 - 0.2125 x Age |
| Moderate | FEV1/FVC < 75.71 - 0.2066 x Age | FEV1/FVC < 78.31 - 0.2125 x Age |
| Severe | FEV1/FVC < 64.29 - 0.2066 x Age | FEV1/FVC < 66.75 - 0.2125 x Age |
| Very severe | FEV1/FVC < 47.59 - 0.2066 x Age | FEV1/FVC < 49.85 - 0.2125 x Age |

Figure 1. Coefficients and 95% confidence intervals from the meta-analysis of (a) mental quality of life, (b) physical quality of life, (c) reported dyspnoea, and (d) respiratory exacerbations, with each COPD classification (A: FEV1-based; B: FEV1/FVC-based). In each model no COPD is the reference.

H:\DRAFT PAPERS\BOLD PAPERS\Sonia severity index\APRIL_2017\Figure 1_new (002).tif

Figure 2. Post-bronchodilator FEV1/FVC ratio against age overlaid with the FEV1/FVC-based staging classification by sex

H:\DRAFT PAPERS\BOLD PAPERS\Sonia severity index\APRIL_2017\Figure 2 (004).tif