

Grading the severity of airways obstruction: new wine in new bottles

Philip H. Quanjer Department of Pulmonary Diseases and Department of Paediatrics, Erasmus Medical Centre, Erasmus University, Rotterdam, the Netherlands
pquanjer@ziggo.nl

Jeffrey J. Pretto Department of Respiratory and Sleep Medicine, John Hunter Hospital, and School of Medicine and Public Health, University of Newcastle, Newcastle, New South Wales, Australia
jeff.pretto@hnehealth.nsw.gov.au

Danny J. Brazzale Department of Respiratory and Sleep Medicine, Austin Hospital, and Institute for Breathing and Sleep, Heidelberg, Victoria, Australia
Danny.Brazzale@austin.org.au

Piotr W. Boros Lung Function Lab, National TB & Lung Diseases Research Institute, Warsaw, Poland
piotr.boros@gmail.com

Correspondence and requests for reprints should be addressed to

Philip Quanjer, Kervel 19, 7443 GT Nijverdal, The Netherlands. E-mail: pquanjer@ziggo.nl

Corresponding author: Philip H. Quanjer, pquanjer@ziggo.nl

Running title: Grading the severity of airways obstruction

Abstract: 200 words (200 allowed)

Manuscript: 3095 words (3000 allowed)

Tables: 3

Figures: 4

Abstract

Objective: Redesign current grading of obstructive lung disease so that it is clinically relevant and free of biases related to age, height, sex and ethnic group.

Material: Spirometric records from 17,880 subjects (50.4% females) from hospitals in Australia and Poland, and 21,191 records (53.0% females) from 2 epidemiological studies (age 18-95 years).

Methods: We adopted the American Thoracic Society and European Respiratory Society (ATS/ERS) criteria for airways obstruction based on an FEV₁/(F)VC ratio below the 5th percentile and graded the severity of pulmonary function impairment using z-scores for FEV₁ which signify how many standard deviations a result is from the mean predicted value.

Results: Using the lower limit of normal for FEV₁/(F)VC and z-scores for FEV₁ of -2, -2.5, -3 and -4 to delineate severity grades of airflow limitation leads to close agreement with ATS/ERS severity classifications and removes age, sex and height related bias.

Conclusion: The new classification system is simple, easily memorised and clinically valid. It retains previously established associations with clinical outcomes and avoids biases due to the use of percent predicted FEV₁. Combined with the Global Lung Function prediction equations it provides a world-wide diagnostic standard free of bias due to age, height, sex and ethnic group.

SHORT SENTENCE OF 130 CHARACTERS OR LESS SUMMARISING THE MOST IMPORTANT FINDINGS:

Using FEV₁ z-scores to classify severity of airways obstruction is clinically valid and overcomes bias inherent with %pred.

Introduction

It is accepted that a diagnosis of airflow limitation should be based on an abnormally low ratio of the forced expiratory volume in one second (FEV_1) and vital capacity (expressed as the larger of inspiratory or expiratory vital capacity (VC) or forced expiratory vital capacity (FVC)). Surprisingly there is no universal agreement yet what constitutes a low $FEV_1/(F)VC$ ratio. The European Respiratory Society (ERS) and American Thoracic Society (ATS) [1-3] recommend to regard the 5th percentile of the distribution in a population of healthy lifelong non-smokers as the lower limit of normal (LLN), whereas the Global Initiative on Obstructive Lung Disease (GOLD) [4] recommend to regard a post-bronchodilator ratio of less than 0.70 as indicating persistent airways obstruction. The latter lower limit was arbitrarily selected because of its simplicity, and because it constituted a limit that would not vary with the prediction equation used. Many investigators have documented that the GOLD criterion for defining airway obstruction can result in marked over-estimation of the prevalence of airway obstruction in subjects aged 45 years or older, which may lead to unnecessary medical expenditure, and in under-estimation in younger adults [5]. This is because using a fixed ratio of 0.70 ignores both the natural decline of the $FEV_1/(F)VC$ ratio with age and the sex differences observed in a normal population. The limitations of this approach and the resultant misdiagnoses have been acknowledged by the GOLD group [4]. A standardised approach to interpretation of spirometry utilising the LLN for $FEV_1/(F)VC$ provides a statistically and scientifically more valid approach to defining abnormality in the $FEV_1/(F)VC$ ratio, with consequent reductions in misdiagnoses of pathological airways obstruction and earlier detection of mild disease.

Once a diagnosis of airways obstruction has been made, spirometry values are commonly used to categorise pulmonary function impairment, and the ATS/ERS recommended approach is to use the FEV_1 as a percentage of the predicted value ($FEV_1\%pred$) as the basis for this classification [3]. Current ATS/ERS recommendations defines an $FEV_1\%pred$ of >70% as mild impairment, 60-69% as moderate, 50-59% as moderately severe, 35-49% as severe and <35% as very severe [3]. This approach uses an arbitrary number of severity categories, using arbitrarily decided cut-off values. However, the use of percent predicted in this way leads to a pronounced age-related bias. For example, the LLN for FEV_1 in elderly healthy subjects and in pre-school children may be as low as 65% and 72% of predicted, respectively [6-7]. The result is that the severity of impairment category corresponds with differing degrees of abnormality for different age groups. A more valid approach would be to take into consideration the underlying distribution of normal lung function data when classifying test results into severity category. The z-score (signifying the number of standard deviations a result is from the mean predicted value) provides a metric for achieving this. A similar

approach has previously been shown to provide effective risk stratification for death from respiratory and vascular causes [8].

The objective of the present study is to evaluate an alternative to the ATS/ERS system for grading respiratory impairment to overcome biases related to age, height and sex.

Material

We used adult data from two population based studies: the Health Survey for England (HSE) comprising 9,961 subjects (53.8% female, age 18-93 yr) [9] and the National Health and Nutrition Examination Survey (NHANES III) comprising 11,230 subjects (52.2% female, age 18-90 yr) [10]. All data from these epidemiological studies were included in the analysis, and not just those data used for the derivation of normal predicted values. We also utilised data from three clinical populations: 2,776 records from the Austin Hospital (AH) in Victoria, Australia (49.8% females, age 18-95 yr), 4,258 from the John Hunter Hospital (JHH) in New South Wales, Australia (53.8% females, age 18-95 yr), and 10,846 from the National Research Institute of TB & Lung Diseases (TLDRl) in Warsaw, Poland (47.8% females, age 18-92 yr). The clinical data comprised consecutively collected test results from patients referred for lung function assessment for the typically wide range of clinical purposes seen in large hospital-based clinical laboratories. Tests were conducted between August 2008 and June 2012 (JHH), January 2001 and May 2012 (AH), and April 2009 and June 2012 (TLDRl). All spirometry was performed in accordance with internationally agreed standards applied at the time of data collection [11-12], and only baseline or pre-bronchodilator status data was included in the analysis.

All analyses were limited to those aged 18 years and older and to those of European descent because there were few non-European data in the clinical dataset. This study is a retrospective analysis of de-identified data, obviating the need for approval from local Ethics Committees.

Methods

Predicted values and z-scores were derived for each subject in each dataset using prediction equations from the Global Lung Function Initiative (GLI-2012) [7] using special-purpose software [13]. The z-score indicates the number of standard deviations that a measurement differs from the mean predicted value. Airway obstruction was diagnosed using the ATS/ERS definition [3] of

$FEV_1/(F)VC < LLN$, *i.e.* the z-score was less than -1.645 [1-3,6-7]. Categorisation of severity of airway obstruction was also made using the 5-category scale recommended by ATS/ERS [3].

Data analysis was performed using the statistical software R [Version 3.0.1; R Foundation, <http://www.r-project.org>]. As the relationship between the z-score for FEV_1 and $FEV_1\%pred$ is symmetrically heteroscedastic (figure 1), the mean z-score was derived at ratios of 0.8, 0.7, 0.6, 0.5, 0.4, 0.35 and 0.30, and subsequently described by linear regression. Differences between the ATS/ERS and new grading systems were tested using two-sided t-test and by multinomial logistic regression using sex, age and height categories as covariates.

Results

Table 1 shows the age distribution of the subjects in the five study groups. As expected, the prevalence rate of airways obstruction is much higher in clinical patients than in the population based samples (figure 2).

There is a strong linear relationship between $FEV_1\%pred$ and FEV_1 expressed as a z-score (figure 1) across the $FEV_1\%pred$ range of 30-80% ($z\text{-score } FEV_1 = 5.589 \times FEV_1\%pred - 5.894$, adjusted $R^2 = 0.9997$). This allows reliable reconstruction of ATS/ERS severity classification cutoff values by simple replacement of $FEV_1\%pred$ with z-scores. After accounting for differences in age there were clinically trivial albeit statistically significant differences in intercept: -0.0048 in centre 2 (JHH) and 0.0024 in centre 3 (TLDR).

Table 2 shows the z-score equivalencies to the ATS/ERS schema. The z-scores have been rounded for simplification since it is cumbersome to use z-scores to precise values, and given that the category boundaries were arbitrarily defined originally [3], this rounding will have little overall effect. Z-scores of -2, -2.5, -3 and -4 can be used as cutoff values that faithfully reproduce the ATS/ERS severity categories.

$FEV_1\%pred$ correlated negatively ($R=0.29$, $p<0.001$) with age and sex (higher $FEV_1\%pred$ in women). The relationship between $FEV_1\%pred$ and z-score for FEV_1 was significantly affected by age (figure 1) and sex ($R=0.988$) (data not shown). At $FEV_1 = 70\%pred$ the z-score for FEV_1 varied between -1.53 and -2.59; at a z-score for $FEV_1 = -2$ $FEV_1\%pred$ varied between 68% and 81% (figure 1). In patients with an FEV_1 / FVC ratio below the lower limit of normal the average ATS/ERS grade of 2.79 differed significantly from the 2.69 average grade with the new system (two-sided t-test, $p=0.0006$).

Multinomial logistic regression revealed that allocation to more severe ATS/ERS grades of airways obstruction compared to the new allocation system was significantly related to age, sex and height.

There were 2,097 clinical patients (11.6%) classified as severe or very severe airflow limitation using the GOLD definition of $FEV_1/FVC < 0.70$ and $FEV_1 < 50\%$ of predicted [4]. The proposed new system classified a similar number as severe or very severe ($n=2,079$, 11.5%), however in 624 patients (29.8%) the GOLD and proposed criteria lead to a different classification. The GOLD classification system underestimated severity of airways obstruction in 303 of these patients and overestimated in 321. Figure 3 shows the prevalence of severe and very severe airway obstruction using these two classification systems as related to age.

Table 3 indicates the small differences in severity classification between the ATS/ERS grading system and the new z-score based grading system. The overall prevalence of airways obstruction is identical with old and new systems since they both utilise $FEV_1/FVC < LLN$ to define abnormality.

Discussion

The primary objective of this study was to design a statistically robust system for detecting airways obstruction and for grading the extent of impairment. Our analysis reveals that the ATS/ERS recommended system of using $FEV_1\%pred$ to grade respiratory impairment is prone to biases caused by age, height and sex differences in underlying distribution of lung function. This can be circumvented by the use of z-scores since they fully account for these biases, and that also introduced by ethnicity factors. The bias due to sex and height, although statistically highly significant, is quite small because the coefficient of variation in males and females is virtually the same [7]. To illustrate this point, a 30 year old 170 cm subject with a z-score for FEV_1 of -2 has a $FEV_1\%pred$ of 75.9% and 75.4% for male or female respectively, at age 75 the corresponding values are 65.1% and 64.3%, respectively.

The age bias has created recent discussion and controversy about the criteria used to define airways obstruction [ref 5,14-15,17-20, 24-26]. The ATS and ERS advocate to use the fifth percentile of FEV_1/FVC from a healthy population as the LLN [3], whereas the GOLD group adopted a fixed 0.70 cut-off value for the FEV_1/FVC ratio because of its simplicity and independence of different prediction equations [4]. The GOLD criterion for airway obstruction, by ignoring sex differences and the natural decline of the FEV_1/FVC ratio with age to well below 0.70, leads to more than a 50%

overestimation of the prevalence rate of airway obstruction in subjects aged 45 years or older, and to significant underestimates in younger adults [5]. These findings are corroborated in the present study (figure 2) and have been implicitly acknowledged by the GOLD group [4]. On that account we adopted the LLN for FEV₁/(F)VC as the gold standard for diagnosing airways obstruction. It is important to note that this study is not about diagnosing clinical syndromes such as asthma or COPD (which rely on clinical assessment in addition to spirometric measurement), but to provide a statistically valid system for interpretation of spirometry.

There is obviously some overlap in FEV₁/FVC ratio between those with and those without respiratory disease, and spirometry results cannot be used to categorically separate ill from healthy individuals. How test results are interpreted is clearly important, and the choice between ERS/ATS and GOLD guidelines for this requires some justification. One key difference between the two guidelines is that GOLD identify a higher incidence of airflow obstruction in elderly subjects. One would expect that such subjects in this 'marginal zone' (with FEV₁/FVC <0.70 but >LLN) would develop clinical symptoms and signs of disease, and follow-up studies have shed light on this clinically relevant issue. In longitudinal studies in those in this marginal zone no association has been found with increased risk of all-cause mortality [13,15-17] (except in symptomatic smokers [18]), with development of respiratory symptoms [18], with accelerated decline in FEV₁ [14,18-20], nor with respiratory care utilisation or poorer quality of life scores compared with a reference group [18]. In a review of the literature Hoesein *et al.* [21] could only find one publication by Mannino *et al.* [22] allegedly demonstrating that subjects with an FEV₁/(F)VC ratio <0.70 but >LLN had an increased risk of premature death and hospitalisation for COPD, and on that basis accepted that GOLD grade 1 does represent respiratory disease. However, Mannino *et al.* [22] misrepresented their own findings: the adjusted hazard ratio for death of 1.1 was *not* elevated given the confidence interval for this value was 0.96-1.3. Also the authors conceded that the 'measure of COPD-related hospitalisations was too inclusive' [23]. Thus there is no evidence that this marginal zone represented by GOLD stage 1 indicates respiratory disease. Conversely, individuals with FEV₁/(F)VC <LLN, but not non-smokers with FEV₁/(F)VC <0.70 but >LLN, have an increased risk of all-cause death [24-25], development of respiratory symptoms [24-25], accelerated decline in FEV₁ [18] and hospitalisation for COPD [26]. We therefore believe that there is overwhelming evidence that the LLN, defined as the 5th percentile of the distribution in a reference population, should be used to diagnose pathological airflow limitation.

Another issue of concern is the use of $FEV_1\%$ pred for categorising disease severity. There have been numerous publications, summarised in the GLI-2012 report [7], that show that the underlying scatter in FEV_1 is not proportional to the predicted value, and it varies with age. There is therefore a widening gap between 80% predicted and the LLN for FEV_1 , particularly in adults older than 45 years (figure 4), leading to an age-related over-estimate of the severity of airflow limitation. The modified ATS/ERS criteria (table 2) remove the bias in age, sex and height; they lead to identifying 4% of healthy women and men from the NHANES III and HSE studies aged 18 years and older as having mild airways obstruction.

Using the absolute value of FEV_1/FVC to define disease severity is problematic in that more severe airflow obstruction can lead to a reduction in FVC with a paradoxical increase in FEV_1/FVC and poorer prognosis [27]. For this reason severity scaling using $FEV_1\%$ pred was recommended [3]. FEV_1/FVC ratios as low as 0.31, 0.36, 0.29, 0.21 and 0.15 occurred in ATS/ERS defined mild, moderate, moderately severe, severe and very severe airway obstruction, respectively. Hence we have focused on FEV_1 z-scores.

Correct classification of obstructive spirometry values in the moderate to severe range is of importance in both clinical and research settings. The most recent GOLD guidelines (2011-2013, [4]) introduced categorization for risk of exacerbation, with one condition leading to patients being categorized as 'high risk' if FEV_1 is less than 50 % of predicted. We identified 2097 patients (11.6%) meeting this spirometric criterion. Nearly the same number of patients were identified with severe or very severe airway obstruction according to the new proposed criteria ($FEV_1/FVC < LLN$ and FEV_1 z-score < -3). Although these results appear to be similar, there is actually a 29.8% disagreement in classification. In the younger age groups the GOLD criterion underestimates the real severity of the obstruction in half of these patients, and overestimates it in half of the elderly (figure 3). This may lead to inappropriate management and/or treatment in both situations and introduces a considerable bias in the selection of patients for intervention studies and other research.

Utilising the underlying spread of spirometry data to stratify risk is not new. In 1983 Peto *et al.* described a 10-fold increase in risk of death from respiratory disease over 20 subsequent years in those with FEV_1 more than 2 standard deviations below average [8]. In developing these new classification criteria we have used z-scores as an essentially identical approach to help stratify severity of disease state. The use of z-scores is already standard procedure when evaluating bone mineral density and growth curves [28-29], and recommended for the interpretation of lung

function test results [6-7]. However, recent analyses reveal that neither FEV₁%pred nor z-scores provide the best indices of choice for studying mortality. Expressing lung function as multiples of the minimum FEV₁ considered to be compatible with life (estimated at 400mL and 500mL in female and male adults respectively) may provide better prediction of mortality [30]. In a 175 cm tall male aged 30 year this minimum FEV₁ of 500 mL represents 11.5 %predicted, z-score -6.51, and at age 75 year 17.0 %predicted, z-score -4.17, illustrating that both %predicted and z-score approaches do not account for this age bias. A z-score of -4.17 in the elderly subject is equivalent to -6.51 in the younger subject from a mortality perspective. Hence disease severity should be categorized using our suggested cutoff values but we should not be tempted to use absolute z-scores to further stratify disease severity.

Spirometric indices of pulmonary function correlate poorly with respiratory symptoms, dyspnoea, exacerbations, hospitalisation, exercise limitation and quality of life. Therefore, any system for grading severity of airflow limitation based solely on spirometric test results will inherently fail to satisfactorily capture the complete clinical picture. Rather than being treated as solely a disease of the lungs, COPD is now considered by some as a multifactorial systemic disease [31]. In view of this the GOLD committee introduced a new classification system that combines spirometry with respiratory symptoms in an attempt to quantify future risk, however this system had not been previously validated. Leivseth *et al.* [16] concluded that GOLD grade I was not associated with increased mortality, and that spirometric GOLD grades II and higher predicted mortality better than the new GOLD ABCD groups among people with COPD from a Norwegian general population. Also, survival is better in the more severe COPD group C (low lung function but less dyspnoea) than in the less severe group B (much better lung function but more dyspnoea) [32], and the addition of dyspnoea and exacerbations in severity classification did not add prognostic value on long-term COPD outcomes [33]. It therefore seems that more research is required to clinically phenotype the heterogeneous syndrome of COPD and identify biomarkers of disease, with a view to bringing treatment tailored at the individual patient within reach [34], and improving the classification of respiratory impairment and improving risk assessment.

Adopting the proposed adjustments to ATS/ERS recommendations avoids contaminating research groups with healthy subjects and avoids unnecessary medical interventions. In addition, this proposal would enable younger patients with milder disease, who would not have been identified using the GOLD criteria, to be included in future research initiatives. The all-age prediction equations for different ethnic groups issued by the Global Lung Function Initiative [7], and endorsed by six

large international respiratory societies, have for the first time provided a means for categorising the lower limits of normal and also the severity of pulmonary impairment which is free of bias due to age, height, sex or ethnic group. Thus a diagnosis of airways obstruction and categorisation of its severity no longer needs to be biased and dependent on where measurements were performed.

Conclusion

The redesigned system for diagnosing airways obstruction and for categorising pulmonary function impairment is based on evidence which underpins its clinical validity and retains previously documented associations between severity of respiratory impairment and outcome variables such as all-cause death. Application with the GLI-2012 prediction equations [7], coupled with standardisation in testing methodologies [11], provides the opportunity for a world-wide standard for performance and interpretation of spirometry which is free of major sources of bias.

Acknowledgement

The authors gratefully acknowledge the constructive comments from the reviewers and Prof. Martin R. Miller which helped to substantially improve the manuscript.

References

1. American Thoracic Society. Lung function testing: Selection of reference values and interpretative strategies. *Am Rev Respir Dis* 1991; 144: 1202–1218.
2. Quanjer PH, Tammeling GJ, Cotes JE, *et al.* Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. *Eur Respir J* 1993; 6 (Suppl. 16): 5–40.
3. Pellegrino R, Viegi G, Brusasco V, *et al.* Interpretative strategies for lung function tests. Series "ATS/ERS Task Force: Standardisation of lung function testing". *Eur Respir J* 2005;26:948–968.
4. Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2013. Available from: <http://www.goldcopd.org/>.
5. Quanjer PH, Enright PL, Miller MR, *et al.* The need to change the method for defining mild airway obstruction. *Eur Respir J* 2011; 37: 720-722.
6. Stanojevic S, Wade A, Stocks J, *et al.* Reference ranges for spirometry across all ages. A new approach. *Am J Respir Crit Care Med* 2008; 177: 253–260.
7. Quanjer PH, Hall GL, Stanojevic S, Cole TJ, Stocks J, and the Global Lungs Initiative. Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012; 40: 1324–1343.
8. Peto R, Speizer FE, Cochrane AL, *et al.* The relevance in adults of air-flow 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: 491–500.
9. Falaschetti E, Laiho J, Primatesta P, Purdon S. Prediction equations for normal and low lung function from the Health Survey for England. *Eur Respir J* 2004; 23: 456–463.
10. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general US population. *Am J Respir Crit Med* 1995; 152: 179–187.
11. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, *et al.* Standardisation of spirometry. *Eur Respir J* 2005; 26: 319-338.
12. Standardization of spirometry, 1994 update. American Thoracic Society. *Am J Respir Crit Care Med* 1995; 152: 1107–1136.
13. Quanjer PH, Stanojevic S, Cole TJ, Stocks J. GLI-2012 Data Conversion software. http://www.lungfunction.org/files/InstallGLI2012_DataConversion.EXE.
14. Pedone C, Scarlata S, Sorino C, Forastiere F, Bellia V, Antonelli Incalzi R. Does mild COPD affect prognosis in the elderly? *BMC Pulm Med* 2010; 10: 35.
15. Vaz Fragoso CA, Concato J, McAvay G, *et al.* Chronic obstructive pulmonary disease in older persons: a comparison of two spirometric definitions. *Respir Med* 2010; 104: 1189–1196.

16. Leivseth L, Brumpton BM, Lund Nilsen TI, M Xiao-Mei, Johnsen R, Langhammer A. GOLD classifications and mortality in chronic obstructive pulmonary disease: the HUNT Study, Norway. *Thorax* Online First, published on April 23, 2013 as 10.1136/thoraxjnl-2013-203270.
17. Ekberg-Aronsson M, Pehrsson K, Nilsson J-Å, Nilsson PM, Löfdahl C-G. Mortality in GOLD stages of COPD and its dependence on symptoms of chronic bronchitis. *Respir Res* 2005; 6(1): 98.
18. Bridevaux P-O, Gerbase MW, Probst-Hensch NM, Schindler C, Gaspoz JM, Rochat T. Long-term decline in lung function, utilisation of care and quality of life in modified GOLD stage 1 COPD. *Thorax* 2008; 63: 768–774.
19. Akkermans RP, Berrevoets MA, Smelee IJ, *et al.* Lung function decline in relation to diagnostic criteria for airflow obstruction in respiratory symptomatic subjects. *BMC Pulm Med* 2012; 12: 12.
20. Akkermans RP, Biermans M, Robberts B, *et al.* COPD prognosis in relation to diagnostic criteria for airflow obstruction in smokers. *Eur Respir J* Published on April 5, 2013 as doi: 10.1183/09031936.00158212.
21. Hoesein FAAM, Zanen P, Lammers J-WJ. Lower limit of normal or $FEV_1/FVC < 0.70$ in diagnosing COPD: an evidence-based review. *Respir Med* 2011; 105: 907–915.
22. Mannino DM, Buist AS, Vollmer WM. Chronic obstructive pulmonary disease in the older adult: what defines abnormal lung function? *Thorax* 2007; 62: 237–241.
23. Mannino DM, Buist AS, Vollmer WM. Authors' reply to two letters to the editor. *Thorax* 2007; 62: 1108–1109.
24. Vaz Fragoso C, Gill T, McAvay G, *et al.* Use of lambda-mu-sigma-derived Z score for evaluating respiratory impairment in middle-aged persons. *Respir Care* 2011; 56: 1771–1777.
25. Vaz Fragoso CA, Concato J, McAvay G, *et al.* The ratio of FEV_1 to FVC as a basis for establishing chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2010; 181: 446–451.
26. Vaz Fragoso CA, Concato J, McAvay G, Van Ness PH, Gill TM. Respiratory impairment and COPD hospitalisation in older persons: a competing risk analysis. *Eur Respir J* 2012; 40: 37–44.
27. Casanova C, Cote C, De Torres JP, Aguirre-Jaime A, *et al.* Inspiratory-to-total lung capacity ratio predicts mortality in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2005; 171: 591-597.
28. Cummings SR, Bates D, Black DM. Clinical use of bone densitometry: scientific review. *JAMA* 2002; 288: 1889-1897.
29. World Health Organization child growth standards. World Health Organization 2006. Available from http://www.who.int/childgrowth/standards/Technical_report.pdf.

30. Miller MR, Pedersen OF. New concepts for expressing forced expiratory volume in 1 s arising from survival analysis. *Eur Respir J* 2010; 35: 873-882.
31. Barnes PJ, Celli BR. Systemic manifestations and comorbidities of COPD. *Eur Respir J* 2009; 33: 1165-1185.
32. Lange P, Marott JL, Vestbø J, Olsen KR, Ingebrigtsen TS, Dahl M, Nordestgaard BG. 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: 975-981.
33. Johannessen A, Nilsen RM, Storebø M, Gulsvik A, Eagan T, Bakke P. Comparison of 2011 and 2007 GOLD guidelines for predicting mortality and hospitalization. *Am J Respir Crit Care Med* 2013 Apr 5. [Epub ahead of print].
34. Fingleton J, Weatherall M, Beasley R. Towards individualised treatment in COPD. *Thorax* 2011; 66: 363-364.

Grading the severity of airways obstruction

Table 1 – Number of subjects by age group, in the five populations included in this study.

		Age (yr)								
Group		18-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-95
Females	AH ¹	8	46	91	127	211	334	340	228	8
	JHH ²	29	90	119	187	320	517	486	217	2
	TLDRi ³	34	290	606	600	1412	1485	1029	205	2
	NHANES ⁴	202	1080	925	792	590	779	579	404	19
	HSE ⁵	98	586	925	844	904	662	462	110	7
Males	AH ¹	12	55	108	189	281	351	265	118	4
	JHH ²	26	122	150	253	437	622	479	195	7
	TLDRi ³	33	292	466	600	1360	1354	893	180	5
	NHANES ⁴	236	1144	1117	844	668	745	657	409	40
	HSE ⁵	115	707	1068	1025	1049	719	506	165	9

¹ Austin Hospital, Victoria, Australia; ² John Hunter Hospital, New South Wales, Australia; ³ National Research Institute of TB and Lung Diseases, Warsaw, Poland; ⁴ NHANES III; ⁵ Health Survey for England 1995-1996.

Table 2 – Proposed new grading system for categorisation of airways obstruction. The conventional system is the ATS/ERS suggested system [3] and the new system is based upon defining airways obstruction using the lower limit of normal for FEV₁/(F)VC and using z-scores for FEV₁ to classify severity.

Obstruction	Grade	ATS/ERS 2005	Proposed
		FEV ₁ /(F)VC < LLN and FEV ₁ %pred	FEV ₁ /(F)VC < LLN and FEV ₁ z-score
Mild	1	> 70%	≥ -2
Moderate	2	60%-69%	≥ -2.5 but < -2
Mod. severe	3	50%-59%	≥ -3 but < -2.5
Severe	4	35%-49%	≥ -4 but < -3
Very severe	5	< 35%	< -4
Mean grade		2.79	2.69

Table 3 – Percentage of subjects, separated into clinical and population-based datasets, falling in the respiratory impairment severity categories as defined using the current ATS/ERS guidelines and the proposed new grading system.

Clinical Data		ATS/ERS classification						Total
		No	Mild	Moderate	Mod. Severe	Severe	Very severe	
New Classification	No obstruction	70.6	1.7	1.0	0	0	0	73.3
	Mild	0	5.9	0.1	0	0	0	5.9
	Moderate	0	0.7	2.6	1.0	0	0	4.4
	Moderately severe	0	0	1.2	2.3	1.3	0	4.8
	Severe	0	0	0.1	1.7	4.3	1.4	7.5
	Very severe	0	0	0	0	1.1	3.0	4.1
Total		70.6	8.3	5.0	5.0	6.6	4.4	100.0

Epidemiological Data		ATS/ERS classification						Total
		No	Mild	Moderate	Mod. Severe	Severe	Very severe	
New Classification	No obstruction	90.3	0.7	0.3	0	0	0	91.3
	Mild	0	4.9	0	0	0	0	4.9
	Moderate	0	0.4	0.6	0.2	0	0	1.2
	Moderately severe	0	0	0.4	0.4	0.2	0	0.9
	Severe	0	0	0	0.3	0.6	0.2	1.2
	Very severe	0	0	0	0	0.1	0.3	0.4
Total		90.3	5.9	1.4	0.9	0.9	0.5	100.0

Figure

Grading the severity of airways obstruction

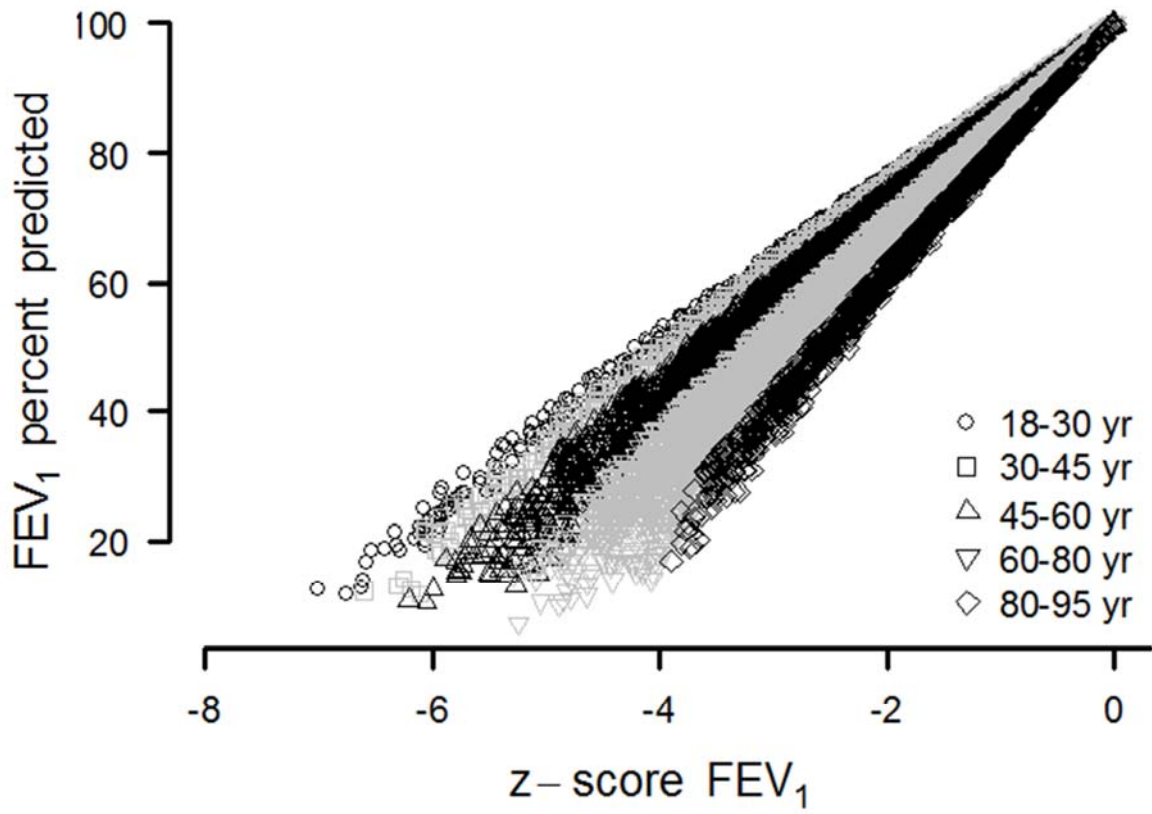


Figure 1

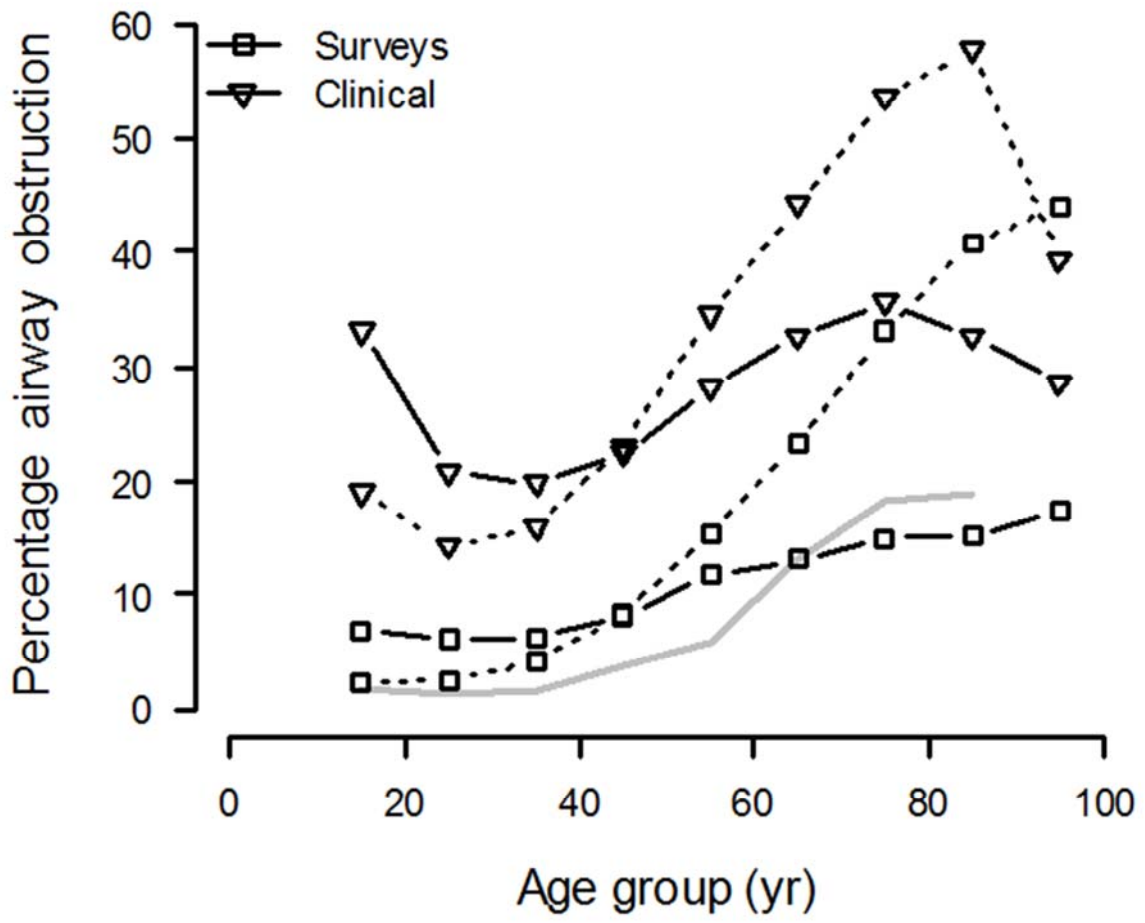


Figure 2

Grading the severity of airways obstruction

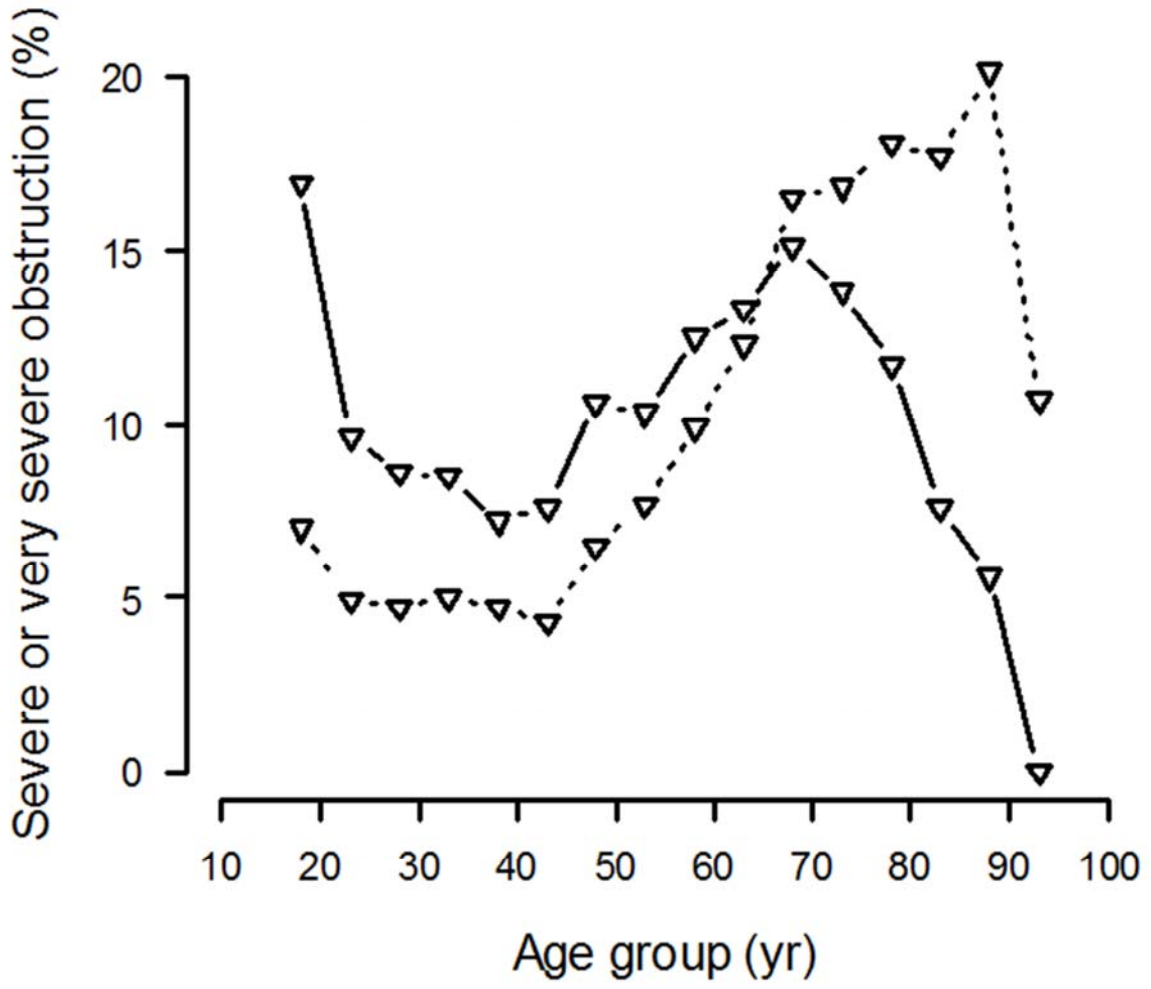


Figure 3

Grading the severity of airways obstruction

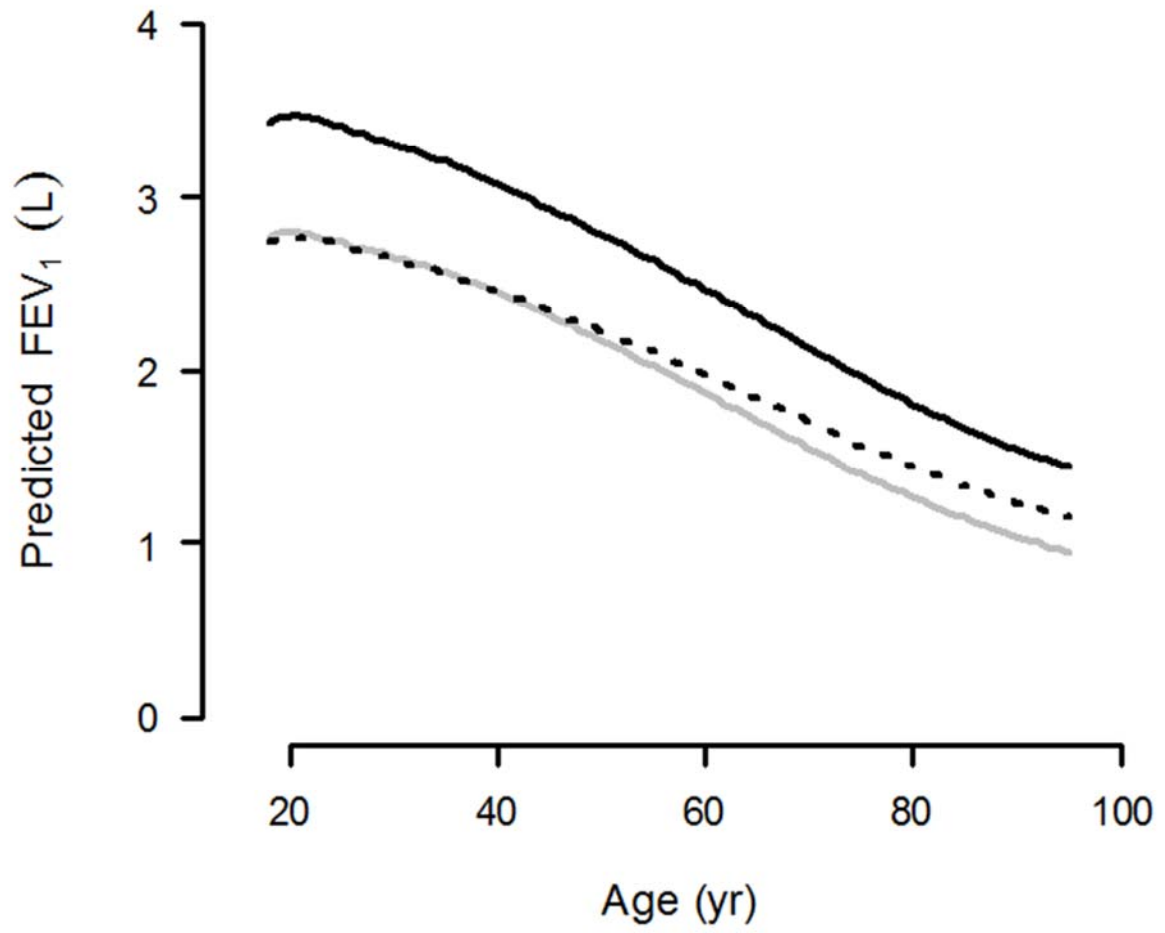


Figure 4