



## Original Article

## GOLD Staging System is Appropriate to Predict Mortality in Older People With Chronic Obstructive Pulmonary Disease

Claudio Pedone<sup>a</sup>, Renato Giua<sup>a,\*</sup>, Nicola Scichilone<sup>b</sup>, Vincenzo Bellia<sup>b</sup>, Raffaele Antonelli-Incalzi<sup>a</sup>

<sup>a</sup> Area di Geriatria, Università Campus Bio-Medico di Roma, Rome, Italy

<sup>b</sup> Dipartimento Biomedico di Medicina Interna e Specialistica (Di.Bi.MIS), Università di Palermo, Palermo, Italy

## ARTICLE INFO

## Article history:

Received 20 September 2017

Accepted 15 January 2018

Available online 9 March 2018

## Keywords:

GOLD  
Chronic obstructive pulmonary disease  
Elderly  
FEV<sub>1</sub>

## ABSTRACT

**Introduction:** In the new GOLD classification the reduction of FEV<sub>1</sub>, expressed as percentage of predicted value (FEV<sub>1</sub>PP), is considered an important prognostic factor. However, the use of FEV<sub>1</sub>PP may introduce bias, especially if based on equations derived from populations different from the one under study. We evaluated how well the GOLD classification stratifies the mortality risk when FEV<sub>1</sub>PP is based on an equation developed in the same population that gave rise to cases, externally developed equations, or as FEV<sub>1</sub> divided by cubed height (FEV<sub>1</sub>/Ht<sup>3</sup>).

**Methods:** We studied 882 participants aged  $\geq 65$  years. Bronchial obstruction was defined using a fixed cut-off of 0.7 for FEV<sub>1</sub>/FVC. Predicted values of FEV<sub>1</sub> were derived from equations based on the same sample of the cases included in this study and from the European Respiratory Society equations. Severity of bronchial obstruction was also classified according to quartiles of FEV<sub>1</sub>/Ht<sup>3</sup>.

**Results:** All the classification systems showed a non-statistically significant linear tendency with 5-years mortality risk. For the 15-years mortality, the linear trend across severity stages is more evident for GOLD classifications, with significant increments in the hazard ratio. Stratification by FEV<sub>1</sub>/Ht<sup>3</sup> could better discriminate the functional status of participants.

**Conclusion:** The severity of bronchial obstruction according to GOLD classes may stratify mortality risk better than quartiles of FEV<sub>1</sub>/Ht<sup>3</sup>, whereas the second seems to be more suited to stratify the risk of clinical outcomes. Concerns about the use of externally developed reference values to calculate FEV<sub>1</sub>PP do not seem confirmed, at least for GOLD classification.

© 2018 SEPAR. Published by Elsevier España, S.L.U. All rights reserved.

### El sistema de estadificación GOLD es adecuado para predecir la mortalidad en personas de edad avanzada con enfermedad pulmonar obstructiva crónica

## RESUMEN

**Introducción:** En la nueva clasificación GOLD, la reducción del FEV<sub>1</sub> expresada como porcentaje del valor predicho (FEV<sub>1</sub>PP) se considera un factor pronóstico importante. Sin embargo, usar el FEV<sub>1</sub>PP puede introducir sesgos, especialmente si se basa en ecuaciones derivadas de poblaciones diferentes de la que se estudia. Se ha evaluado cómo de adecuadamente estratifica GOLD el riesgo de mortalidad cuando el FEV<sub>1</sub>PP se basa en una ecuación desarrollada con la misma población en la que se dieron los casos, usando ecuaciones desarrolladas externamente, o con el FEV<sub>1</sub> dividido por la altura al cubo (FEV<sub>1</sub>/A<sup>3</sup>).

## Palabras clave:

GOLD  
Enfermedad pulmonar obstructiva crónica  
Ancianos  
FEV<sub>1</sub>

**Abbreviations:** BODE index, Body-Mass Index, Airway Obstruction, Dyspnea, and Exercise Capacity Index; COPD, chronic obstructive pulmonary disease; ERS, European Respiratory Society; FEV<sub>1</sub>, Forced Expiratory Volume in 1 second; FEV<sub>1</sub>/Ht<sup>2</sup>, FEV<sub>1</sub> divided by height squared; FEV<sub>1</sub>/Ht<sup>3</sup>, FEV<sub>1</sub> divided by height squared or cubed; FEV<sub>1</sub>PP, FEV<sub>1</sub> expressed as the percent of the predicted value; FEV<sub>1</sub>-ERS, FEV<sub>1</sub> expressed as the percent of the value predicted by ERS equations; FEV<sub>1</sub>-SARA, FEV<sub>1</sub> expressed as the percent of the value predicted by estimating equations derived from SaRA study; FVC, forced vital capacity; FVCP, FVC expressed as the percent of the predicted value; GOLD, Global Initiative for Obstructive Lung Disease; postBD, post-bronchodilator; SaRA, Salute Respiratoria nell'Anziano – Respiratory Health in the Elderly; SGRQ, Saint George Respiratory Questionnaire.

\* Corresponding author.

E-mail address: [renatogiua@alice.it](mailto:renatogiua@alice.it) (R. Giua).

<https://doi.org/10.1016/j.arbres.2018.01.022>

0300-2896/© 2018 SEPAR. Published by Elsevier España, S.L.U. All rights reserved.

**Métodos:** Estudiamos a 882 participantes de edad  $\geq 65$  años. La obstrucción bronquial se definió utilizando un punto de corte fijo de 0,7 para FEV<sub>1</sub>/FVC. Los valores pronosticados de FEV<sub>1</sub> se derivaron de ecuaciones basadas en la misma muestra de los casos incluidos en este estudio y de las ecuaciones de la *European Respiratory Society*. La gravedad de la obstrucción bronquial también se clasificó de acuerdo con los cuartiles de FEV<sub>1</sub>/A<sup>3</sup>.

**Resultados:** Todos los sistemas de clasificación mostraron una tendencia lineal estadísticamente no significativa en el riesgo de mortalidad a 5 años. Para la mortalidad a 15 años, la tendencia lineal a través de los diferentes estadios de gravedad es más evidente para los estadios GOLD, con incrementos significativos en la razón de riesgo. La estratificación por FEV<sub>1</sub>/A<sup>3</sup> podría discriminar mejor el estado funcional de los participantes.

**Conclusión:** La gravedad de la obstrucción bronquial según la estadificación GOLD puede estratificar mejor el riesgo de mortalidad que los cuartiles de FEV<sub>1</sub>/A<sup>3</sup>. Sin embargo, lo segundo parece el método más adecuado para estratificar el riesgo de resultados clínicos. Las reticencias respecto al uso de valores de referencia desarrollados externamente para calcular FEV<sub>1</sub>PP no parecen confirmarse, al menos para la clasificación GOLD.

© 2018 SEPAR. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

## Introduction

For several years, the grading of chronic obstructive pulmonary disease (COPD) severity proposed by the Global Initiative for Obstructive Lung Disease (GOLD) has been based on Forced Expiratory Volume in 1 second (FEV<sub>1</sub>) only. This classification could not adequately predict clinical outcomes<sup>1,2</sup> and therefore other prognostic indices, such as the BODE index<sup>3</sup> have been proposed to stratify the severity of COPD. In 2011, the GOLD proposed a new classification of COPD severity that in addition to the level of FEV<sub>1</sub> reduction also took into account severity of symptoms and frequency of exacerbations,<sup>4</sup> thus generating a panel of four different stages of disease severity (A through D). This new classification, however, does not seem to have better prognostic capacity,<sup>5</sup> and the GOLD group has recently proposed a new classification that merges the GOLD 2007 and GOLD 2011 classes.<sup>6</sup> In this document, the severity of obstruction expressed as reduction of FEV<sub>1</sub> is still considered a very important prognostic factor at the population level.

One potential issue with the GOLD classification is that FEV<sub>1</sub> is expressed as the percent of the value predicted by estimating equations (FEV<sub>1</sub>PP). This approach may introduce bias, and it has been shown that GOLD classification can lead to misclassification of older patients.<sup>7,8</sup>

Due to these limitation, alternative ways of standardizing FEV<sub>1</sub> have been proposed, such as FEV<sub>1</sub> standardized residuals, FEV<sub>1</sub> divided by height squared or cubed (FEV<sub>1</sub>/Ht<sup>3</sup>) or expressed as a function of the sex-specific first percentile.<sup>9</sup> FEV<sub>1</sub>/Ht<sup>3</sup> may be more informative than FEV<sub>1</sub>PP to predict clinical outcomes even in elderly patients.<sup>9–11</sup> Our hypothesis is that part of the suboptimal prognostic capacity of the GOLD stratification comes from the use of equations developed in population that are different from the one that gives rise to the COPD cases. In this study, we evaluated how well the GOLD classification stratifies the mortality risk when FEV<sub>1</sub> is expressed as percentage of the value predicted by an equation developed in the same population that gave rise to cases, externally developed equations, or as FEV<sub>1</sub>/Ht<sup>3</sup>.

## Methods

### Study population

Between January 1996 and July 1999 a total of 1970 participants were recruited within the context of the SaRA (Salute Respiratoria nell'Anziano – Respiratory Health in the Elderly) study. Details on the SaRA project are available elsewhere.<sup>12</sup> This is a multi-center

Italian project investigating various aspects of chronic airway diseases in people  $\geq 65$  years of age attending pulmonary or geriatric outpatient clinics for any reason. Participants were in stable conditions when the spirometry was performed. Enrollment was on a consecutive basis. The study design was approved by the Ethical Committee of the coordinating center (#276/2012). From this dataset, we selected 1296 participants with post-bronchodilator (postBD) spirometry. We then excluded people with a history of asthma ( $N=224$ ). Of the remaining participants, information on vital status as of December 2010 was available for 882; these patients had clinical and spirometric characteristics comparable to those of subjects lost to follow-up. Causes of death were derived from death certificates, and were available for 74% of participants.

### Pulmonary function tests

All the centers were equipped with an identical fully computerized water-sealed Stead-Wells spirometer (Baires System; Biomedin; Padua, Italy) that met the standards of the American Thoracic Society recommendations for diagnostic spirometry.<sup>13</sup> At baseline, tests were performed with a standardized technique in all centers and a quality control process was successfully implemented: all the centers achieved a high quality performance in spirometry.<sup>12</sup> Obstruction was defined using a fixed cut-off of 0.7 for FEV<sub>1</sub>/forced vital capacity (FVC) considering the postBD spirometry. Predicted values of FEV<sub>1</sub> were derived from equations based on the same sample of the cases included in this study (FEV<sub>1</sub>-SARA)<sup>14</sup> and the European Respiratory Society (FEV<sub>1</sub>-ERS).<sup>15</sup> FEV<sub>1</sub>PP was categorized according to the classes proposed by the GOLD guidelines to stratify severity of obstruction (FEV<sub>1</sub>PP  $\geq 80\%$ ,  $80\% < \text{FEV}_{1\text{PP}} \geq 50\%$ ,  $50\% < \text{FEV}_{1\text{PP}} \geq 30\%$ , and  $\text{FEV}_{1\text{PP}} < 30\%$ ). Severity of bronchial obstruction was also classified according to quartiles of FEV<sub>1</sub>/Ht<sup>3</sup>.

### Analytic approach

The demographic and clinical characteristics were compared across COPD severity groups defined using FEV<sub>1</sub>-SARA. We included in this analysis the distance walked in 6 min, expressed as percent predicted,<sup>16</sup> and comorbid diseases such as ischemic heart disease, heart failure, and stroke. Smoking was analyzed as cumulative exposure (pack-years). The multi-dimensional BODE index was also included, as it is able to predict mortality in older people.<sup>17</sup> The SaRA questionnaire did not include a specific item on exacerbations, therefore we combined two of the Saint George Respiratory Questionnaire (SGRQ) items (“During the past 3 months how many

severe or very unpleasant attacks of chest trouble have you had?” and “How long did the worst attack of chest trouble last?”) to define exacerbations as an attack that lasted for at least 3 days. The risk of dying was estimated using the Kaplan–Meier method, differences in the survival risk was evaluated using the log-rank test. The relative hazard of dying was estimated using a Cox proportional hazard model. The assumptions of such a model were checked by inspecting the distribution of the Schoenfeld residuals over time.

To estimate the discriminative capacity of each risk stratification tool, we estimated the relative increase in risk of each stratum compared to the preceding stratum, i.e. GOLD I vs. not obstructed, GOLD II vs. GOLD I and so on. The goodness of fit of these models was evaluated using the likelihood ratio test, and the overall diagnostic performance of these models was evaluated using the C-statistic. The analyses were performed for both medium-term (5 years) and very long-term (15 years) mortality.

## Results

The mean age of our sample was 73.2 (SD: 6), men were 57.8%. Characteristics of the population according to GOLD stages determined using the FEV1–SARA are reported in Table 1. Participants with bronchial obstruction were predominantly males, especially in classes III and IV. Beside having a higher BODE index, participants in these classes tended to have lower body mass index, higher exposure to cigarette smoking, and higher prevalence of cerebro-vascular disease. We did not observe a linear relationship between GOLD stages and prevalence of comorbidities; for example

the prevalence of ischemic heart disease, heart failure, and stroke was higher in GOLD class III than in GOLD class IV.

The FEV1 predicted by the ERS equation was on average 150 ml smaller compared to the SARA equation. Consequently, 21 patients classified in class I, 8 patients classified in class II, and 3 patients classified in class III using the ERS equation were classified in class II, III, and IV, respectively, according to the SARA equation. This notwithstanding, the general pattern across GOLD classes based on FEV1–ERS was not different from that observed using FEV1–SARA. Likewise, no clear specific relationships in clinical characteristics were found after stratification by quartiles of FEV1/Ht<sup>3</sup> (Table 2).

In participants with COPD, the average distance walked in 6' was 67.6 m. In the same group, 15% reported at least one exacerbation in the preceding 3 months, and 24.5% had a modified Medical Research Council dyspnea index >2. The global SGRQ score was 39. We found no differences in the capacity of the different severity grading methods to stratify exercise capacity expressed by the distance walked in 6' (Fig. 1, panel A). We found that the other important clinical outcomes (exacerbations, dyspnea, and quality of life) differed only across the first three GOLD classes, while participants in class IV did not differ from patients in class III. Using quartiles of FEV1/Ht<sup>3</sup>, instead, the linear association was evident across all categories (Fig. 1, panels B–D).

Data on vital status were gathered after a median of 13 years (range: 0.5–15). The risk for mortality was 19.5% at 5 years and 53.7% at the end of follow-up. Fig. 2 shows the Kaplan–Meier curves for 15-year mortality risk stratified by stages. Considering the 5-years mortality, all the classification systems showed a fairly

**Table 1**  
Characteristics of the sample according to GOLD classification (predicted value obtained with SARA equation<sup>14</sup>).

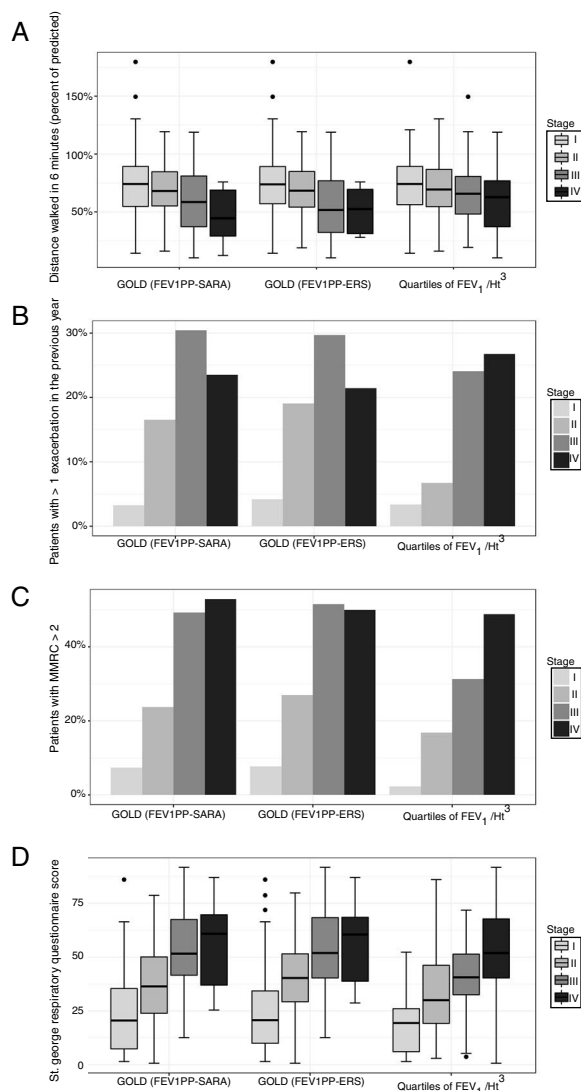
	Not obstructed N: 535	GOLD 1 N: 122	GOLD 2 N: 139	GOLD 3 N: 69	GOLD 4 N: 17
Mean age (SD)	73 (6)	74 (6)	73 (5)	73 (6)	72 (6)
Sex (men)	44	78	75	87	94
Body mass index	26.4 (4)	25.7 (3.3)	26.2 (4.3)	24.9 (4)	23.5 (3.1)
FVCP postBD	98.9 (17.8)	95.7 (11.6)	64.5 (8.5)	40.8 (6.2)	25.7 (3.4)
FEV1PP postBD	88.2 (15.6)	100.6 (12.7)	79.6 (13.5)	67.3 (13.7)	57.4 (15)
BODE index	1 (1)	1 (1)	2 (2)	5 (2)	7 (2)
Pack/year	12 (22)	35 (34)	38 (37)	48 (37)	37 (33)
Ischemic heart disease	5.4	7.4	9.4	11.6	5.9
Heart failure	3.9	5.7	8.6	13	11.8
History of stroke	6.5	2.5	2.9	7.2	0
Diabetes mellitus	12.1	9.8	11.5	8.7	17.6
Peripheral artery disease	4.9	9.8	6.5	2.9	5.9
Malignancies	3.9	5.7	7.2	2.9	0

FEV1PP postBD: forced expiratory volume in 1 second of post-bronchodilator spirometry; FVCP postBD: forced vital capacity of post-bronchodilator spirometry; SD: standard deviation.

**Table 2**  
Characteristics of the sample according to quartiles of FEV1 divided by squared height.

	Not obstructed N: 535	Qtl. I N: 89	Qtl. II N: 89	Qtl. III N: 83	Qtl. IV N: 86
Mean age (SD)	73 (6)	73 (6)	74 (5)	73 (6)	74 (6)
Sex (men)	44	90	75	70	81
Body mass index	26.4 (4)	26.1 (2.9)	25.7 (4)	26.4 (4.6)	24.5 (3.8)
FVCP postBD	98.9 (17.8)	97.6 (11.6)	77.6 (9.6)	59.9 (12.1)	38.6 (9.5)
FEV1PP postBD	88.2 (15.6)	101.1 (12.1)	88.5 (12.9)	77.2 (16.1)	65.8 (15.2)
BODE index	1 (1)	1 (1)	2 (2)	3 (2)	5 (2)
Pack/year	12 (22)	39 (32)	37 (37)	37 (39)	43 (37)
Ischemic heart disease	5.4	4.5	15.7	4.8	10.5
Heart failure	3.9	2.2	10.1	10.8	11.6
History of stroke	6.5	3.4	3.4	2.4	4.7
Diabetes mellitus	12.1	9	13.5	10.8	9.3
Peripheral artery disease	4.9	10.1	10.1	3.6	3.5
Malignancies	3.9	1.1	14.6	3.6	2.3

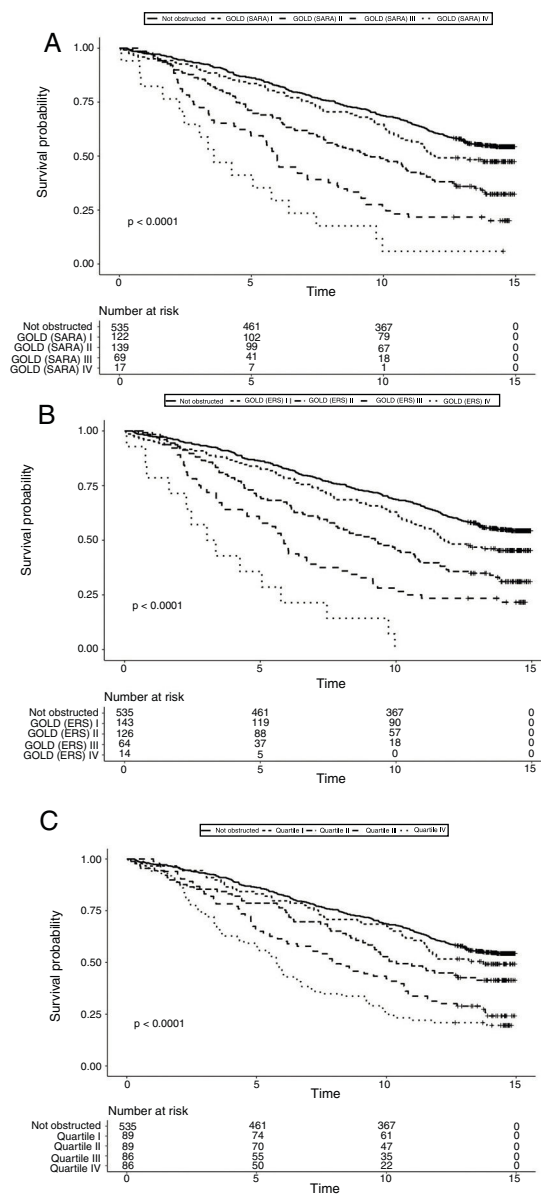
FEV1PP postBD: forced expiratory volume in 1 second of post-bronchodilator spirometry; FVCP postBD: forced vital capacity of post-bronchodilator spirometry; SD: standard deviation.



**Fig. 1.** Distribution of walked distance (A), rate of exacerbations (B), dyspnea (C), and St. George Respiratory Questionnaire (D) according to COPD severity rated by GOLD or quartiles of FEV<sub>1</sub>/Ht<sup>3</sup>.

linear relationship with mortality. Nonetheless, as reported in Table 3, the incremental changes in risk associated with each class were not statistically significant, except for the increase in risk from stage II to stage III of both GOLD classifications. The goodness of fit of the quartiles of FEV<sub>1</sub>/Ht<sup>3</sup> was somewhat worse compared to the GOLD classifications, but the overall discriminative capacity expressed by the c statistics was 0.6 for all models. When the models were adjusted for age and sex, the incremental association between the GOLD stages defined using the SARA predicting equation was only marginally affected, while for GOLD stages defined using the ERS predicting equation the linear association was more evident, with significant increases of hazard ratio in group II vs. I, group III vs. II and group IV vs. III. No significant incremental risk was observed for quartiles of FEV<sub>1</sub>/Ht<sup>3</sup>.

Considering the 15-years mortality risk (Fig. 2), the linear association across severity stages is more evident for the two GOLD classifications, as also confirmed by the significant increments in the hazard ratio, shown in Table 3. Although the linear association was evident also for quartiles of FEV<sub>1</sub>/Ht<sup>3</sup>, the increment in the hazard ratio was significant only for stage III compared to stage II. The goodness of fitness of the models obtained with the GOLD classification system was marginally better compared to that of



**Fig. 2.** Kaplan–Meier curves for 15-years risk of mortality, according to COPD severity rated by GOLD, considering FEV<sub>1</sub>-SARA (panel A) and FEV<sub>1</sub>-ERS (panel B), or quartiles of FEV<sub>1</sub>/Ht<sup>3</sup> (panel C).

the model including quartiles of FEV<sub>1</sub>/Ht<sup>3</sup>, but the overall discriminative capacity did not change between the three classification systems. After correction for age and sex, no substantial changes were observed.

Finally, when only deaths from respiratory causes were taken into account, the overall pattern did not change, although the estimates were more unstable due to the relatively low number of events (46 overall, 28 in the first 5 years).

## Discussion

Our data indicate that the GOLD classification system is superior to a stratification based on quartiles of FEV<sub>1</sub> standardized by height in predicting mortality in older persons with COPD. This finding is partly in contrast with other reports<sup>9–11</sup> showing that FEV<sub>1</sub>/Ht<sup>3</sup> is superior to FEV<sub>1</sub>PP in stratifying the mortality risk. One possible explanation for this discrepancy is that previous studies on this issue have compared groups of FEV<sub>1</sub>PP and FEV<sub>1</sub>/Ht<sup>3</sup> based on

**Table 3**

Hazard ratio for mortality according to GOLD stages and quartiles of FEV1 divided by squared height. Each category is contrasted with the preceding category (e.g. GOLD I vs. not obstructed, GOLD II vs. GOLD I, etc.).

	GOLD (FEV1-SARA)	GOLD (FEV1-ERS)	Quartiles of FEV1/Ht <sup>3</sup>
<b>5-years mortality</b>			
Not obstructed	1	1	1
Stage I	1.22 (0.74–1.99)	1.25 (0.79–1.98)	1.24 (0.71–2.16)
Stage II	1.84 (1.08–3.15)	1.88 (1.13–3.13)	1.33 (0.68–2.63)
Stage III	1.56 (0.96–2.53)	1.56 (0.95–2.56)	1.61 (0.9–2.88)
Stage IV	1.79 (0.87–3.69)	2.04 (0.96–4.34)	1.38 (0.84–2.27)
Log-likelihood	49.2	52.2	44.2
C-statistic	0.625	0.626	0.623
<b>15-years mortality</b>			
Not obstructed	1	1	1
Stage I	1.24 (0.94–1.63)	1.31 (1.02–1.69)	1.17 (0.85–1.6)
Stage II	1.52 (1.11–2.09)	1.5 (1.1–2.03)	1.29 (0.87–1.93)
Stage III	1.61 (1.16–2.25)	1.52 (1.07–2.16)	1.52 (1.05–2.2)
Stage IV	1.89 (1.08–3.29)	2.53 (1.4–4.59)	1.37 (0.97–1.93)
Log-likelihood	81.8	82.5	76.5
C-statistic	0.598	0.598	0.598

FEV1/Ht<sup>3</sup>: forced expiratory volume in 1 second divided by height cubed.

percentiles, and not GOLD classes vs. percentiles of FEV1/Ht<sup>3</sup>. One exception is a study by Miller et al.<sup>18</sup> that compared GOLD classes vs. arbitrarily defined groups of FEV1 divided by height squared (FEV1/Ht<sup>2</sup>), in which the authors found that FEV1/Ht<sup>2</sup> could better stratify mean survival compared to GOLD classes. In this study, however, the cut-off to define bronchial obstruction was set at FEV1/FVC < 0.89, and the authors did not provide information on the mortality risk across groups.

A possible explanation of our findings is that the GOLD classification uses pragmatic cut-off of FEV1PP based on consensus, clinical experience, and previous epidemiological data,<sup>19</sup> and in this study we compared it with a distribution-based classification. The use of arbitrary cut-off (as in the GOLD classification) may better fit the risk of mortality compared to distribution-based cut-off (i.e. quartiles) if the relationship between FEV1 and risk was not linear, while the opposite would be true if the relationship was linear. An alternative explanation is that FEV1PP, on which the GOLD classification is based, better captures the underlying pathophysiological changes compared to FEV1/Ht<sup>3</sup>. Several studies, however, starting from the seminal observation by Fletcher and Peto,<sup>20</sup> indicate that FEV1 standardized using height is a good indicator of reduction of pulmonary function with respect to the mortality risk. Further studies investigating different cut-off for FEV1/Ht<sup>3</sup> are needed to better explore this issue.

To our knowledge, this is the first study comparing the discriminative capacity of GOLD classification and quartiles of FEV1/Ht<sup>3</sup> with respect to walking speed, dyspnea, rate of exacerbations, and disease-specific quality of life in COPD. In contrast to what we observed for mortality, quartiles of FEV1/Ht<sup>3</sup> could better stratify more specific clinical indicators (dyspnea, exacerbations, disease-specific quality of life) compared to the GOLD classification, while no grading system could stratify an overall indicator of exercise capacity such as the distance walked in 6'. The same observation made above about mortality risk may apply for these findings: the distribution-based classification may have a better fit because the relationship is linear. Furthermore, these outcomes were evaluated at the same time of the spirometry, therefore the discriminative capacity of the classification systems are relevant to the actual clinical conditions, not to a future outcome such as mortality. Based on these data, the GOLD classification seems better suited to identify people at generic risk for mortality, but not patients with worse health status related to pulmonary problems. In keeping with this hypothesis, the discriminative capacity of GOLD stages is better for very long-term compared to medium-term survival, and there was

no difference in the discriminative capacity for mortality from any cause or mortality from pulmonary causes.

Results obtained using predicting equations developed in the same population that gave rise to the cases differed marginally from those obtained using equation developed in a different population. Thus, although guidelines recommend the use of internally developed standards,<sup>21</sup> the use of externally developed standards seem to be acceptable. Indeed, we found a relatively small discrepancy between GOLD classes assigned by FEV1-SARA or FEV1-ERS.

Limitations of this study deserve consideration. An important limitation is that we do not have follow-up data on pulmonary and physical function and symptoms, therefore we cannot evaluate which classification system better identifies people with worse prognosis in terms of clinical course of disease. At the same time we cannot discriminate if some subjects has changed GOLD class during the follow-up. We used measured height, and this could introduce a bias as vertebral fractures are frequent in the older population (especially women) and COPD.<sup>22</sup> Due to this problem, people with vertebral fractures might have been misclassified as having a FEV1/ht<sup>3</sup> higher than the real. Since vertebral fractures are a risk factor for adverse outcomes in COPD,<sup>23</sup> this bias may have caused a reduction in the discriminative capacity of FEV1/ht<sup>3</sup>. Furthermore, even if a classification basing on quartiles of FEV1/Ht<sup>3</sup> is statistically logical considering the absence of indications from epidemiological studies, it may occur that this would not be the best method.

The observed frequency of exacerbation refers to the last 3 months period and, thus, cannot be compared to that reported in other studies. For example by Hurst et al.<sup>24</sup> Indeed, we based our definition on two SGRQ items that have a time frame of only 3 months; therefore our data are not directly comparable to those reporting yearly exacerbation rates. Finally, the loss of many patients to follow-up may have biased our results on mortality. However, patients followed up and patients lost to follow up had comparable clinical characteristics.

## Conclusion

In conclusion, our study indicates that the severity of bronchial obstruction according to GOLD classes may stratify mortality risk better than quartiles of FEV1/Ht<sup>3</sup>, whereas the second seems to be more suited to stratify clinical outcomes, such as dyspnea, walking speed and quality of life. Concerns about the appropriateness of using externally developed reference values to calculate FEV1PP do not seem to be confirmed, at least for GOLD classification.

## Authors' contribution

Each have substantially contributed to data collection, manuscript drafting and revision. All authors read the final manuscript and approved it for publication.

## Conflict of interest

The authors declare they have no conflict of interest.

## References

1. Papaioannou AI, Loukides S, Gourgoulis KI, Kostikas K. Global assessment of the COPD patient: time to look beyond FEV1? *Respir Med.* 2009;103:650–60.
2. Soriano JB, Lamprecht B, Ramirez AS, Martinez-Cambor P, Kaiser B, Alfageme I, et al. Mortality prediction in chronic obstructive pulmonary disease comparing the GOLD 2007 and 2011 staging systems: a pooled analysis of individual patient data. *Lancet Respir Med.* 2015;3:443–50.
3. Celli BR, Cote CG, Marin JM, Casanova C, Montes de Oca M, Mendez RA, et al. The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. *N Engl J Med.* 2004;350:1005–12.

4. Vestbo J, Hurd SS, Agustí AG, Jones PW, Vogelmeier C, 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.* 2013;187:347–65.
5. Johannessen A, Nilsen RM, Storebø M, Gulsvik A, Eagan T, Bakke P. Comparison of 2011 and 2007 Global Initiative for Chronic Obstructive Lung Disease guidelines for predicting mortality and hospitalization. *Am J Respir Crit Care Med.* 2013;188:51–9.
6. Global Initiative for Chronic Obstructive Lung Disease. *Glob Initiati Chronic Obstr Lung Dis – GOLD s.d.*
7. Vaz Fragoso CA, Concato J, McAvay G, Van Ness PH, Rochester CL, Yaggi HK, et al. Chronic obstructive pulmonary disease in older persons: a comparison of two spirometric definitions. *Respir Med.* 2010;104:1189–96.
8. Fragoso CAV. Epidemiology of chronic obstructive pulmonary disease (COPD) in aging populations. *COPD.* 2016;13:125–9.
9. Miller MR, Pedersen OF. New concepts for expressing forced expiratory volume in 1 s arising from survival analysis. *Eur Respir J.* 2010;35:873–82.
10. Pedone C, Scarlata S, Scichilone N, Forastiere F, Bellia V, Antonelli-Incalzi R. Alternative ways of expressing FEV1 and mortality in elderly people with and without COPD. *Eur Respir J.* 2013;41:800–5.
11. Pedone C, Scarlata S, Zito A, Forastiere F, Scichilone N, Battaglia S, et al. Alternative ways of expressing forced expiratory volume in the first second and long-term mortality in elderly patients with asthma. *Ann Allergy Asthma Immunol Off Publ Am Coll Allergy Asthma Immunol.* 2013;111:382–6.
12. Bellia V, Pistelli R, Catalano F, Antonelli-Incalzi R, Grassi V, Melillo G, et al. Quality control of spirometry in the elderly. The SA.RA. study. *Salute Respirazione nell'Anziano = Respiratory Health in the Elderly.* *Am J Respir Crit Care Med.* 2000;161:1094–100.
13. Standardization of Spirometry, 1994 Update. American Thoracic Society. *Am J Respir Crit Care Med.* 1995;152:1107–36.
14. Pistelli R, Bellia V, Catalano F, Antonelli-Incalzi R, Scichilone N, Rengo F. Spirometry reference values for women and men aged 65–85 living in southern Europe: the effect of health outcomes. *Respir Int Rev Thorac Dis.* 2003;70:484–9.
15. Quanjer PH, Tammeling GJ, Cotes JE, Pedersen OF, Peslin R, Yernault JC. Lung volumes and forced ventilatory flows. *Eur Respir J.* 1993;6 Suppl. 16:5–40.
16. Casanova C, Celli BR, Barria P, Casas A, Cote C, de Torres JP, et al. The 6-min walk distance in healthy subjects: reference standards from seven countries. *Eur Respir J.* 2011;37:150–6.
17. Stolz D, Meyer A, Rakic J, Boeck L, Scherr A, Tamm M. Mortality risk prediction in COPD by a prognostic biomarker panel. *Eur Respir J.* 2014;44:1557–70.
18. Miller MR, Pedersen OF, Dirksen A. A new staging strategy for chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis.* 2007;2:657–63.
19. Pauwels RA, Buist AS, Calverley PM, Jenkins CR, Hurd SS. GOLD Scientific Committee, global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop summary. *Am J Respir Crit Care Med.* 2001;163:1256–76.
20. Fletcher C, Peto R. The natural history of chronic airflow obstruction. *Br Med J.* 1977;1:1645–8.
21. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al. Standardisation of spirometry. *Eur Respir J.* 2005;26:319–38.
22. de Vries F, van Staa TP, Bracke MSGM, Cooper C, Leufkens HGM, Lammers J-WJ. Severity of obstructive airway disease and risk of osteoporotic fracture. *Eur Respir J.* 2005;25:879–84.
23. Pascual-Guardia S, Badenes-Bonet D, Martín-Ontiyuelo C, Zuccarino F, Marín-Corral J, Rodríguez A, et al. Hospital admissions and mortality in patients with COPD exacerbations and vertebral body compression fractures. *Int J Chron Obstruct Pulmon Dis.* 2017;12:1837–45.
24. Hurst JR, Vestbo J, Anzueto A, Locantore N, Müllerova H, Tal-Singer R, et al. Susceptibility to exacerbation in chronic obstructive pulmonary disease. *N Engl J Med.* 2010;363:1128–38.