Characteristics, stability and outcomes of the 2011 GOLD COPD groups in the ECLIPSE cohort

Alvar Agusti[1], Lisa D. Edwards[2], Bartolomé Celli[3], William MacNee[4], Peter M. A. Calverley[5], Hana Müllerova[6], David A. Lomas[7], Emiel Wouters[8], Per Bakke[9,10], Steve Rennard[11], Courtney Crim[2], Bruce E. Miller[12], Harvey O. Coxson[13], Julie C. Yates[2], Ruth Tal-Singer[12] and Jørgen Vestbo[14,15] on behalf of the ECLIPSE investigators[16]

Affiliations: [1]Thorax Institute, Hospital Clinic, IDIBAPS, Universitat de Barcelona and CIBER Enfermedades Respiratorias, FISIB, Mallorca, Spain. [2]GlaxoSmithKline, Research Triangle Park, NC, [3]Dept of Respiratory Medicine; Brigham and Women's Hospital, Boston, MA, [11]University of Nebraska Medical Center, Omaha, NE, and [12]GlaxoSmithKline, King of Prussia, PA, USA. [4]University of Edinburgh & Royal Infirmary, Edinburgh, [5]Dept of Respiratory Medicine, University Hospital Aintree, Liverpool, [6]Epidemiology, GlaxoSmithKline R&D, Middlesex, [7]Dept of Medicine, University of Cambridge, Cambridge Institute for Medical Research, Cambridge, and [15]Manchester Academic Health Sciences Centre, University of Manchester, Manchester, UK. [8]Dept of Respiratory Medicine, Maastricht University Medical Centre, Maastricht, The Netherlands. [9]Institute of Medicine, University of Bergen, Bergen, and [10]Dept of Thoracic Medicine, Haukeland University Hospital, Bergen, Norway. [13]Dept of Radiology, University of British Columbia, Vancouver General Hospital, Vancouver, BC, Canada. [14]Dept of Respiratory Medicine, Odense University Hospital/University of Southern Denmark, Denmark. [16]For a complete list of principal investigators and centres participating in ECLIPSE see Acknowledgements.

Correspondence: A. Agusti: Institut del Tòrax. Hospital Clínic, Universitat de Barcelona. Villarroel 170, Escala 3, Planta 5. 08036 Barcelona, Spain: E-mail: alvar.agusti@clinic.ub.es

ABSTRACT   The 2011 Global Initiative for Chronic Obstructive Lung Disease (GOLD) classifies patients with chronic obstructive pulmonary disease (COPD) into four groups (A to D).

We explored the characteristics, stability and relationship to outcomes of these groups within the ECLIPSE study (Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points) (n=2101).

Main results showed that: 1) these groups differed in several clinical, functional, imaging and biological characteristics in addition to those used for their own definition; 2) A and D groups were relatively stable over time, whereas groups B and C showed more temporal variability; 3) the risk of exacerbation over 3 years increased progressively from A to D, whereas that of hospitalisation and mortality were lowest in A, highest in D and intermediate and similar in B and C, despite the former having milder airflow limitation. The prevalence of comorbidities and persistent systemic inflammation were highest in group B.

The different longitudinal behaviour of group A *versus* B and C *versus* D (each pair with similar forced expiratory volume in1 s (FEV1) values supports the 2011 GOLD proposal of assessing COPD patients by more than FEV1 only. However the assumption that symptoms do not equate to risk appears to be naïve, as groups B and C carry equally poor clinical outcomes, though for different reasons.

@ERSpublications
Exploring distribution, characteristics, temporal stability and relationship with 2011 GOLD group outcomes in ECLIPSE http://ow.ly/mjJ8p

Introduction

Chronic obstructive pulmonary disease (COPD) is a heterogeneous disease, and the forced expiratory volume in the 1 s (FEV1) does not describe this complexity fully [1]. The recently released 2011 Global Initiative for Chronic Obstructive Lung Disease (GOLD) recommendations recognise this limitation and propose a novel multidimensional approach to assess COPD patients, which includes the level of symptoms (as assessed by either the modified Medical Research Council (mMRC) score or the chronic obstructive pulmonary disease (COPD) Assessment Test (CAT), the FEV1 value (expressed as a percentage of predicted value) and the individual history of previous exacerbations [2]. As a result, COPD patients can now be categorised and treated according to four groups (A, B, C and D) [2].

Very recently, three studies have used existing cohorts of COPD patients to explore different aspects of this new GOLD assessment proposal. HAN *et al.* [3] used the COPDGene cohort [4] to explore the influence of the instrument used to assess the level of symptoms (mMRC *versus* the St. George's Respiratory Questionnaire (SGRQ), as a surrogate for CAT) on group assignment and risk of exacerbations during follow-up. LANGE *et al.* [5] investigated the ability of the 2011 GOLD proposal to predict the clinical course of the disease in two Danish general population cohorts combined [5]. Finally, SORIANO *et al.* [6] determined the distribution and validity of the four GOLD 2011 groups as predictors of mortality in eleven small Spanish cohorts merged for the analysis.

The ECLIPSE study(Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points) is a large, observational, longitudinal, multicentre, international study that includes a very detailed, clinical, functional imaging, and biological characterisation of COPD patients with a 3-year follow-up[1, 7]. Therefore, it offers a unique opportunity to reproduce some of these previous analyses and to explore other novel aspects of potential relevance. Specifically, in this study we used the ECLIPSE cohort to: 1) compare the 2007 and 2011 GOLD recommendations, in terms of patient distribution and assessment of disease severity; 2) characterise in detail the four 2011 GOLD groups; 3) explore their stability after a 3-year follow-up; and 4) investigate their relationship with clinical relevant outcomes (exacerbations, hospitalisations and mortality), and identify significant risk factors.

Methods

Study design

The design of the ECLIPSE study has been previously published [7]. In brief the ECLIPSE study is an observational, longitudinal, controlled study that recruited patients who mostly attended the outpatient clinics of secondary or tertiary care hospitals and, occasionally, primary care. After a baseline visit, participants were evaluated at 3 months, 6 months and then every 6 months for 3 years. ECLIPSE complies with the Declaration of Helsinki and Good Clinical Practice Guidelines, and has been approved by the ethics committees of the participating centres. All participants provided written informed consent.

Population

The original ECLIPSE cohort included 2164 patients with COPD, as previously published [1]. Since the mMRC dyspnoea score was missing in 62 patients and one patient had an unknown GOLD stage, the present analysis includes 2101 GOLD stage I–IV patients (97% of the original ECLIPSE cohort) with full 2011GOLD data. These 2101 patients were categorised into one of four groups (A, B, C or D) according to the 2011 GOLD recommendations [2] using the mMRC score (the CAT had not been published [8] when the ECLIPSE study was designed [7]) and the FEV1 value, both of which were measured at recruitment, as well as the number of exacerbations retrospectively recorded from the year before inclusion in the ECLIPSE study.

Measurements

The methodology used in ECLIPSE has also been reported in detail elsewhere [1, 7]. We used standardised questionnaires to record dyspnoea (mMRC), comorbidities (American Thoracic Society and Division of

Lung Disease of the National Institute of Health questionnaire (ATS-DLD-78-adult)), health-status (St. George's Respiratory Questionnaire for COPD patients (SGRQ-C)) and depression (Center for Epidemiological Studies Depression Scale (CES-D)) [1, 7]. Nutritional status was assessed by the body mass (BMI) and fat-free mass indexes (FFMI) by bioelectrical impedance [7]. COPD exacerbations requiring treatment with antibiotics, oral corticosteroids and/or hospitalisation in the year prior to the study (and during follow-up) were also recorded. Spirometry and the 6-min walking distance (6MWD) were performed according to international guidelines [9, 10]. Spirometric reference values were those of the European Community for Coal and Steel (ECCS) [11]. The BMI, airflow obstruction, dyspnoea, exercise capacity (BODE) index was calculated according to CELLI *et al.* [12]. All subjects underwent a low-dose computed tomography (CT) scan of the chest acquired using multi-detector-row CT scanners (GE Healthcare or Siemens Healthcare), as described elsewhere [1]. Hospitalisations and all-cause mortality were recorded during a 3-year follow-up [13, 14]. Finally, the annual rate of FEV_1 decline in each patient [15] and the proportion of patients with persistent systemic inflammation at baseline [14] were calculated as reported previously and included in the current analysis.

Statistical analysis

Results are shown as mean ± SD or proportion, as appropriate. Jonckheere–Terpstra tests were used to assess differences in subject characteristics among groups with ordering of A ⩽ B ⩽ C ⩽ D [16]. Kaplan–Meier curves were constructed to describe the occurrence of exacerbations, hospitalisations and all-cause mortality over the study period for both the 2007 [17] and 2011 GOLD [2] classifications. Concordance probabilities and their associated 95% confidence intervals were calculated for Cox regression models for exacerbations, hospitalisations and all-cause mortality to evaluate the predictive accuracy of the 2007 and 2011 GOLD classifications [18]. A Cox proportional hazards regression model was used to identify factors associated with increased risk of mortality and hospitalisation, A p-value <0.05 (two sided) was considered statistically significant.

Results

Patient distribution and assessment of disease severity by the 2007 and 2011 GOLD recommendations

Figure 1 shows the distribution of the COPD patients included in the analysis according to the 2007 (fig 1a) and 2011 (fig 1b) GOLD recommendations. The latter resulted in a rightward shift of the distribution, indicating that more patients were being classified in more severe groups. The mean ± SD 2007 GOLD stage was not different between groups A (2.0 ± 0.0) and B (2.0 ± 0.1), and only slightly worse (albeit statistically significantly p<0.001) between groups C (3.0 ± 0.5) and D (3.2 ± 0.6). Figure 1c also shows the cumulative distribution of the four 2011 GOLD groups within each 2007 GOLD stage. GOLD stage II includes a large proportion of group A (54%) and B (32%) patients but also some group C (7%) and D patients (8%). Since, by definition, GOLD stage II patients must have an FEV_1 ⩾50% pred [17], these latter two groups were classified as high risk groups (C or D) exclusively, because of their past exacerbation history (⩾2 per year.). Alternatively, also by definition, GOLD stages III and IV cannot include group A or B patients, where FEV_1 must be ⩾50% of reference [2]. There were more group D patients in GOLD stage IV than GOLD stage III (fig. 1c). The frequency distribution of these four groups was not different in males and females (online supplementary table 1E).

Characterisation of the 2011 GOLD groups

The application of the 2011 GOLD assessment criteria resulted in the expected differences in the mMRC scores, FEV_1 and exacerbation rates between the four groups (table 1). In addition, these four groups differed also in many of the other characteristics studied (tables 1 and online supplementary table 2E). It is of note that some variables were particularly abnormal in the two high symptoms groups (B and D; SGRQ-C and CES-scores, many comorbid conditions, exercise capacity and most inflammatory biomarkers), whereas others were mostly abnormal in the two high risk groups (C and D; amount of emphysema present and arterial oxygenation saturation) (tables 1 and online supplementary table 2E). Finally, of note, age, gender, cumulative smoking exposure, FEV_1 reversibility and the annual rate of FEV_1 decline were either not different between groups or differences were of unlikely clinical relevance (tables 1 and online supplementary table 2E).

As in two previous studies [4, 5], we stratified the C and D groups further according to the specific criteria that determined the inclusion of patients in these high risk groups (FEV_1<50% reference (C1 or D1), exacerbation history ⩾2 years. (C2 or D2) or both (C3 or D3)). An FEV_1 <50% of reference alone was the most frequent reason to classify a patient as high risk in either the C or D groups (70% and 63% of cases, respectively), whereas the presence of frequent exacerbations alone was responsible for a much lower
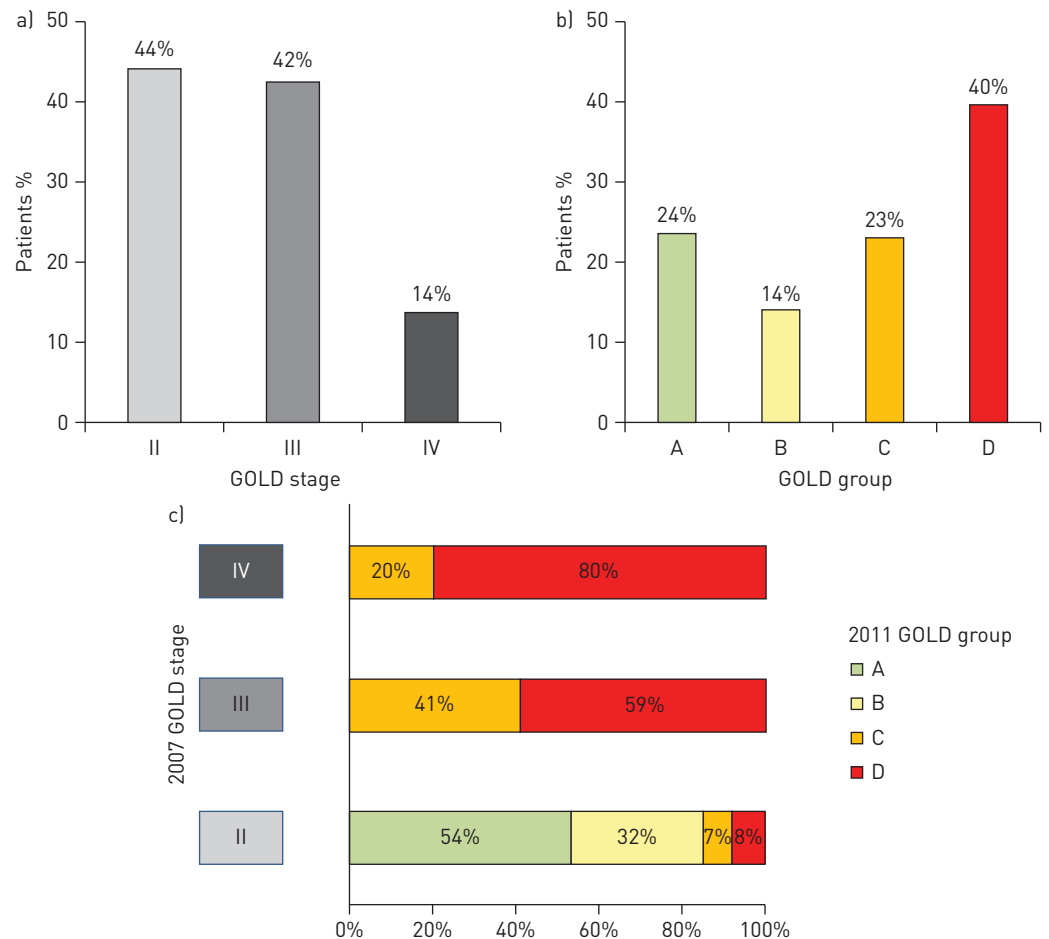
FIGURE 1 Frequency distribution of the chronic obstructive pulmonary patients studied here according to the 2007 (a) or 2011 (b) Global Initiative for Chronic Obstructive Lung Disease (GOLD) recommendations. Two subjects were classified as having GOLD stage I but were not included in a) or b). c) Shows the cumulative distribution of the four patient groups for the 2011 GOLD classification within each 2007 GOLD stage.

proportion of patients (13% and 9%, respectively); the remaining patients (18% and 28%, respectively) were classified as C or D patients because of the presence of both low FEV1 and frequent exacerbations. Online supplementary tables 3E and 4E present the main characteristics of these three C and D subgroups. By definition, exacerbation frequency was particularly high in C2/D2 and C3/D3.

Temporal stability of the 2011 GOLD groups during the 3-year follow-up
Figure 2 presents the distribution of the four 2011 GOLD groups both at recruitment and 3 years later. However, it is important to note that not all patients studied at recruitment contributed data during the 3-year follow-up, due to a variety of reasons that include death, drop-out, lost for follow-up and/or missing data. This effect was less marked in group A, worse in group D and intermediate in groups B and C. With this caveat in mind, it is interesting to see that the relative majority of A and D patients do remain in their original classification, whereas B and C patients have greater temporal variability (fig. 2). B patients either remained in the same category (B), improved (A) or deteriorated (D) in roughly similar proportions. By contrast C patients, either remained stable (C) or deteriorated (D) and only a relatively small fraction improved (B or A). Finally, it is of note that the proportion of patients that moved to group A from B, C or D was lowest in D, intermediate in C and largest in B (fig. 2).

Relationship with long-term outcomes: comparison of the 2007 and 2011 GOLD recommendations
Figure 3 presents Kaplan–Meier curves that describe the occurrence of exacerbations (a, d and g), hospitalisations (b, e and h) and all-cause mortality (c, f and i) over the study period for 2007 GOLD stages in all patients (a–c), 2011 GOLD groups in all patients (d–f) and 2011 GOLD groups for patients with a FEV1<50% reference (GOLD II). Within 2007 GOLD stages, each of the pairwise comparisons are significantly different for exacerbations, hospitalisations (p<0.001) and all-cause mortality (p⩽0.002).

TABLE 1 Outcomes of the four groups of patients as per the 2011 Global Initiative for Chronic Obstructive Lung Disease (GOLD) recommendations [2].

Distribution	2011 GOLD groups				p-value
	A	B	C	D	
Patients	495 (24)	293 (14)	483 (23)	830 (40)	
GOLD 2011 assessment variables					
mMRC dyspnoea score	0.6 ± 0.5	2.3 ± 0.6	0.8 ± 0.4	2.6 ± 0.7	<0.001
FEV1 % predicted	64.2 ± 8.1	61.6 ± 8.2	42.2 ± 11.5	37.2 ± 11.2	<0.001
Exacerbations previous 12 months	0.2 ± 0.4	0.3 ± 0.5	1.0 ± 1.4	1.3 ± 1.4	<0.001
Demographics					
Age years	63.2 ± 7.2	64.1 ± 7.3	62.9 ± 7.2	63.6 ± 6.9	0.078
Female	191(39)	106(36)	149(31)	280(34)	0.07
Body mass index kg·m^{-2}	26.7 ± 4.8	28.5 ± 6.1	25.2 ± 4.9	26.5 ± 6.1	<0.001
Fat-free mass index kg·m^{-2}	17.3 ± 2.5	18.0 ± 3.5	16.7 ± 2.5	17.1 ± 2.9	<0.001
Smoking, pack-years	48.0 ± 29.4	47.5 ± 25.8	45.2 ± 23.2	51.1 ± 27.7	<0.001
Current smoker	188 (38)	101 (34)	207 (43)	267 (32)	0.001
Symptoms					
Chronic bronchitis	131 (26)	104 (35)	182 (38)	318 (38)	<0.001
SGRQ-C total score	31.9 ± 17.3	55.3 ± 15.5	44.6 ± 16.8	62.2 ± 15.6	<0.001
CESD total score	8.3 ± 8.0	13.8 ± 9.4	9.5 ± 8.2	13.5 ± 9.7	<0.001
Self-reported comorbidities					
With any comorbidity	383 (77)	248 (85)	350 (72)	690 (83)	<0.001
Number of comorbidities	2.0 ± 1.7	2.8 ± 2.2	1.7 ± 1.7	2.3 ± 1.9	<0.001
Physiology					
FEV1 % reversibility	11.4 ± 13.2	11.8 ± 12.9	10.3 ± 13.5	10.2 ± 14.5	0.035
S_{aO_2} %	95.7 ± 2.1	95.1 ± 2.5	94.4 ± 3.2	93.7 ± 3.2	<0.001
6-min walking distance m	440 ± 106	360 ± 101	408 ± 106	307 ± 114	<0.001
BODE index	0.8 ± 0.8	2.7 ± 1.2	2.6 (1.2	5.1 ± 1.6	<0.001
Imaging					
% LAA -950 HU	11.8 ± 9.7	12.5 ± 9.9	18.9 ± 11.5	22.5 ± 12.8	<0.001
Outcomes during 3-year follow-up					
ECOPD per year	0.6 ± 0.8	0.9 ± 1.0	1.3 ± 1.4	1.7 ± 1.7	<0.001
⩾1 hospitalisation	56 (11)	71 (25)	143 (30)	378 (46)	<0.001
Mortality rate	22 (4)	28 (10)	41 (8)	114 (14)	<0.001
Mortality rate per 100 patient-years	1.6 ± 7.4	3.7 ± 11.5	3.2 ± 10.5	5.5 ± 13.7	<0.001
Rate of annual FEV1 decline mL per year	-33.4 (46.2)	-38.0 (47.0)	-30.2 (45.5)	-31.9 (34.6)	0.157

Data are presented as n (%) or mean ± SD, unless otherwise stated. mMRC: dyspnoea score according to the modified Medical Research Council questionnaire; FEV1; forced expiratory volume in 1 s; SGRQ-C: St George's Respiratory Questionnaire for chronic obstructive pulmonary disease (COPD) patients; CESD: Center for Epidemiological Studies Depression Scale; S_{aO_2}; arterial oxygen saturation; BODE: body mass index, airflow obstruction, dyspnoea, exercise capacity index; % LAA: percentage of low attenuation areas (i.e. emphysema) in the computed tomography; ECOPD: exacerbation of COPD.

Similarly, within 2011 GOLD groups (all patients), each of the pairwise comparisons are also significantly different for exacerbations (p<0.001), but not for hospitalisations or mortality, which were different in A and D (p⩽0.010), but similar in B and C patients, despite group B having better lung function and fewer exacerbations; yet, group B also had, at baseline, the highest prevalence of comorbidities (table 1), particularly cardiac comorbidities (online supplementary table 2E), and of patients with persistent systemic inflammation (fig. 4), defined as we described previously [14].

Compared with the 2007 GOLD classification, the 2011 GOLD classification had a higher concordance probability (95% CI) for predicting exacerbations (0.591 (0.577–0.605) versus 0.570 (0.558–0.583); p=0.003) and hospitalisations (0.659 (0.640–0.679) versus 0.628 (0.610–0.646); p=0.001); both classifications were similar for the prediction of all-cause mortality (0.596 (0.562–0.630) versus 0.614 (0.576–0.653); p=0.32).

Because there is limited information in the ECLIPSE database on specific causes of death [1, 7], we used a Cox proportional hazards regression model to identify factors associated with increased risk of mortality. These included more severe airflow limitation, high age, low BMI, high FFMI, increased number of comorbidities, and lower exercise capacity (table 2). Interestingly, GOLD B patients had less airflow
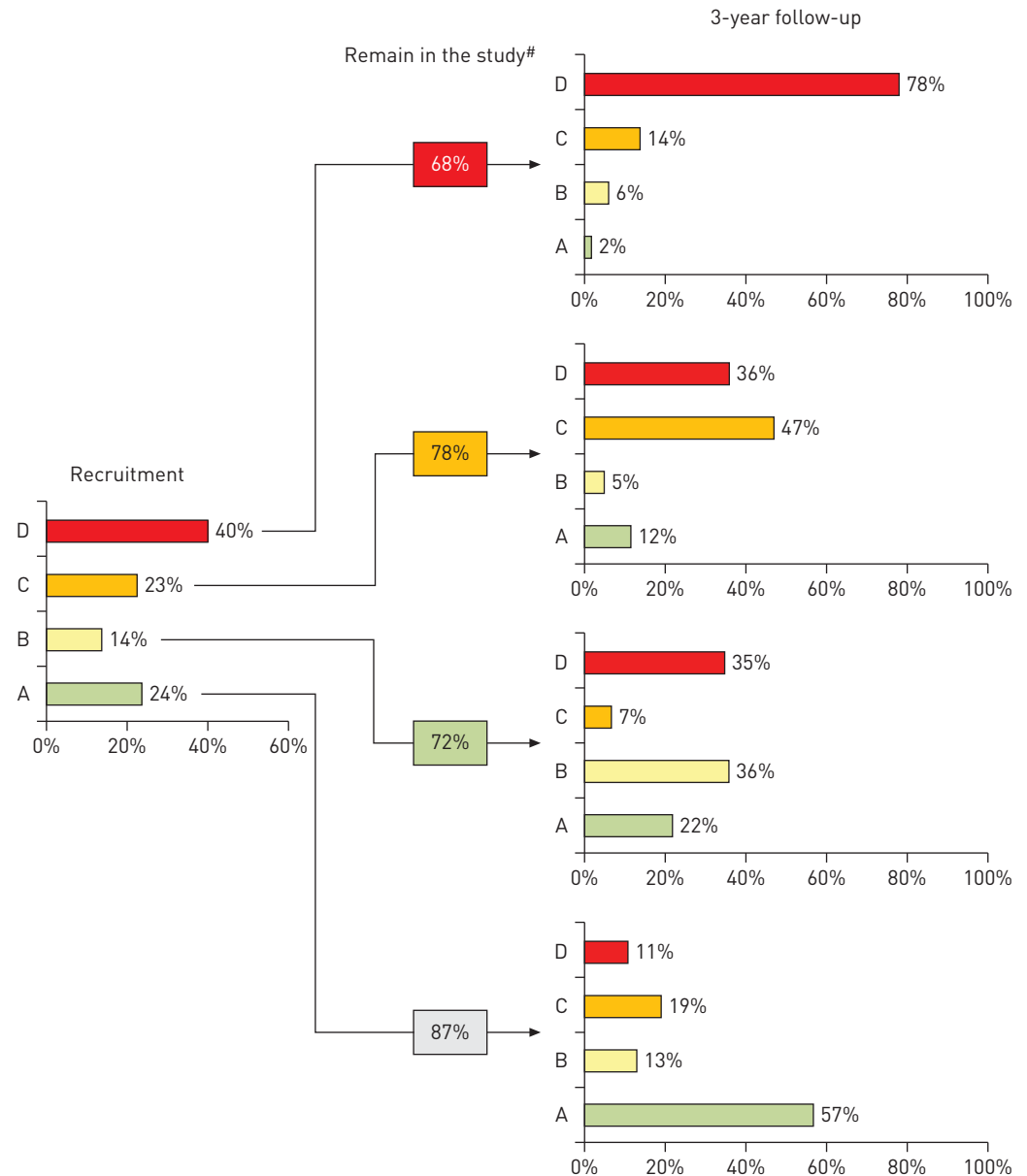
FIGURE 2 Temporal variation of the four 2011 Global Initiative for Chronic Obstructive Lung Disease groups at recruitment and during the 3-year follow-up. Boxes with percentages indicate the proportion of baseline patients available for examination at the 3-year follow-up. #: death/drop-out/lost for follow-up/missing data.

limitation and higher BMI than GOLD C patients, but were older, had a higher FFMI, more comorbidities, and poorer exercise capacity (table 1). These differences may contribute to their similar mortality during follow-up (fig. 3). A similar set of covariates (FEV_1 % pred, age, BMI, number of comorbidities, prior exacerbation history, arterial oxygen saturation and dyspnoea) were associated with increased risk of hospitalisation (table 2). For each of these covariates, differences were also noted between GOLD B and GOLD C patients (table 1).

Two final aspects of our analysis deserve comment. First, there were no significant differences between gender across the four groups for mortality or hospitalisation during follow up (online supplementary table 1E) but females had a higher exacerbation rate than males ($p \leqslant 0.017$). Secondly, if only 2007 GOLD stage II patients (FEV_1 >50% of reference) are considered (fig. 1c), the 2011 GOLD proposal lacks discriminant power to predict exacerbations in C2 and D2 patients, the risk of hospitalisation is similar and intermediate in C2 and B patients and, interestingly, mortality during follow-up was zero in C2 patients (fig. 3g–i).
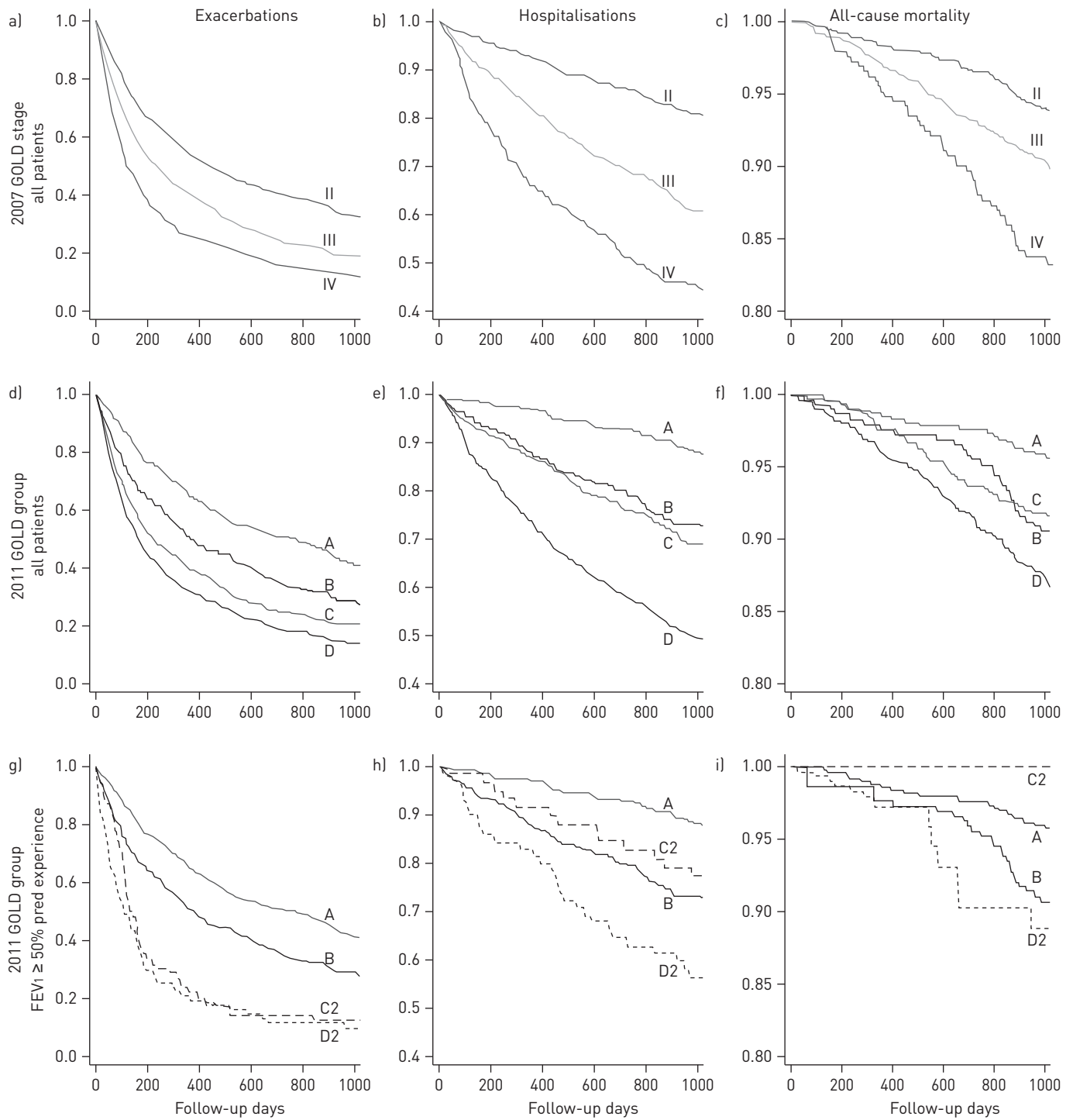
FIGURE 3 Kaplan–Meier curves for exacerbations (a, d and g), hospitalisations during follow-up (b, e and h) and all-cause mortality (c, f and i) for the 2007 Global Initiative for Chronic Obstructive Lung Disease (GOLD) recommendations for all patients (a–c), 2011 GOLD recommendations for all patients (d–f) and 2011 GOLD recommendations for patients with forced expiratory volume in 1 s (FEV1) ⩾50% of the reference (GOLD stage II) (g–i).

Discussion

The 2011 GOLD recommendations have proposed a novel, multidimensional strategy to assess disease severity and guide therapy in COPD [2]. Because this proposal was mostly based on expert opinion [2], several recent studies have used existing COPD cohorts to test it [3, 5, 6]. Our study confirms some of their results and extends them by presenting novel aspects not previously explored.
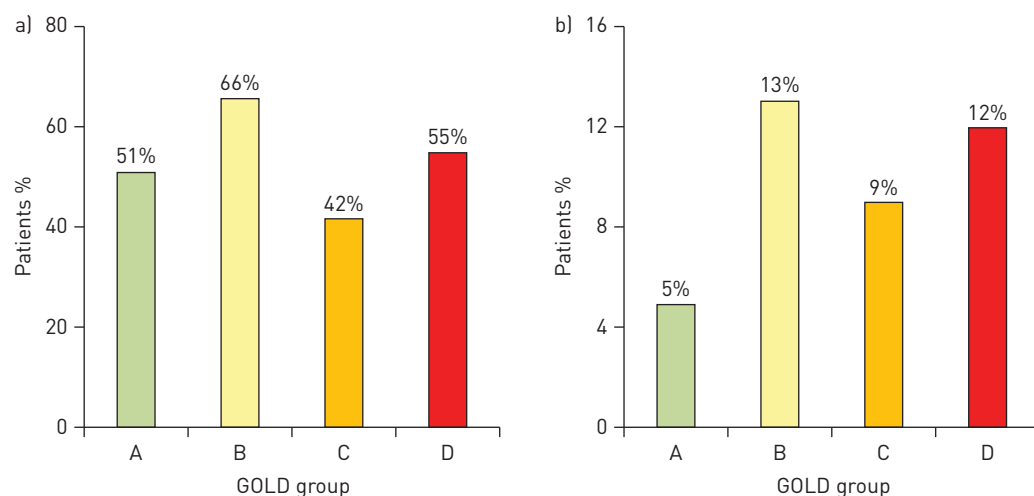
FIGURE 4 Percentage of patients with at least one self-reported comorbidity (a) or persistent, systemic inflammation (b), as described by AGUSTI *et al.* [14], in the four groups of the 2011 Global Initiative for Chronic Obstructive Lung Disease (GOLD) patient classification. a) $p<0.05$ for all comparisons except for group A *versus* group B. b) $p<0.05$ for all comparisons except for group B *versus* group D. It is of note that both were highest in B patients.

Previous findings

Our results reproduce and support five previous observations. 1) In keeping with LANGE *et al.* [5] we found that the new 2011 GOLD classification assesses disease severity differently from the previous 2007 proposal, since more individuals are categorised in the more severe group (fig. 1a and b). This is not surprising if the multidimensional nature of the 2011 GOLD proposal is considered [2]. 2) The prevalence of the four 2011 GOLD groups in our study (fig. 1b) was similar to that reported in previous COPD cohorts recruited from hospital clinics [3, 6], but clearly different from that described in COPD patients identified in the general population, where group A clearly predominates [5]. 3) In agreement with previous reports [3, 5] we found that a low FEV1 (<50% reference) was the main reason for classifying patients as high risk (C and D groups). 4) We reproduced previous observations (fig 3d–f) indicating that the new 2011 GOLD assessment proposal predicts future exacerbations appropriately [3, 5]. Likewise, in keeping with LANGE *et al.* [5], we

TABLE 2 Cox proportional hazards regression analyses for all-cause mortality and hospitalisations during follow-up

Independent variable	Estimate	p-value
Mortality		
Forced expiratory volume in 1 s % predicted	0.982 (0.969–0.996)	0.011
Age years	1.052 (1.021–1.085)	<0.001
Body mass index kg·m^{-2}	0.931 (0.882–0.983)	0.01
Fat-free mass index kg·m^{-2}	1.146 (1.036–1.267)	0.008
Number of comorbidities	1.100 (1.001–1.210)	0.048
6-min walking distance m	0.997 (0.996–0.999)	0.002
Hospitalisations		
Modified Medical Research Council dyspnoea score	1.246 (1.119–1.387)	<0.001
Forced expiratory volume in 1 s % predicted	0.974 (0.966–0.982)	<0.001
Exacerbations in previous 12 months	1.195 (1.123–1.270)	<0.001
Age years	1.021 (1.005–1.038)	0.01
Body mass index kg·m^{-2}	0.979 (0.960–0.999)	0.04
Hypertension	0.669 (0.526–0.851)	0.001
Angina	0.565 (0.371–0.858)	0.007
Osteoarthritis	0.695 (0.489–0.987)	0.042
Diabetes	0.604 (0.392–0.932)	0.023
Depression requiring treatment	0.657 (0.473–0.912)	0.012
Number of comorbidities	1.209 (1.113–1.314)	<0.001
Arterial oxygen saturation %	0.929 (0.902–0.956)	<0.001

Data are presented as hazard ratio (95% confidence interval), unless otherwise stated.

observed that mortality was similar in groups B and C (fig. 4), despite the former having, by definition, better lung function and fewer exacerbations [2]. We think that this is a clear example of how the new 2011 GOLD assessment proposal goes beyond FEV1 and identifies a group of patients who, despite the presence of moderate airflow limitation, are at high risk of mortality. SORIANO et al. [6] reported a similar observation over 3- but not at 10-year follow-up, raising the possibility of different causes of death at different time-points during follow-up. The specific causes of death were not investigated by SORIANO et al. [6], but LANGE et al. [5] reported that B patients often died of cardiovascular disease and cancer. In keeping with this observation, we also found that B patients showed the highest prevalence of comorbidities and included more patients with persistent systemic inflammation (fig. 4, table 1). The former is known to have a direct and significant impact on survival in COPD [19] and the latter has been recently shown to increase mortality six-fold, irrespective of the severity of the pulmonary abnormalities present [14]. These observations may have clinical implications for the practitioner, since B patients appear to be a unique group that deserve special attention in order to understand the origin of their greater symptoms, which are likely to relate to the presence of comorbidities, which may occur in patients with moderate airflow limitation [1], and need to be actively looked for and treated adequately if present [2]. 5) The observation that despite having frequent exacerbations and requiring frequent hospitalisations, C2 patients have a mortality of zero during follow-up (fig. 3) is also of clinical interest and in keeping with the observations by LANGE et al. [5].

Novel observations

Our study extends previous reports [3, 5, 6] by providing potentially relevant information on six aspects not investigated before. 1) We found that the four 2011 GOLD groups differed in many clinical, functional, imaging and biological characteristics in addition to those used for their categorisation (tables 1 and online supplementary table 2E). Interestingly, some of them (comorbidities and systemic inflammation) are more prevalent in the two high symptom groups (B and D), whereas others (emphysema and low arterial oxygen saturation) are more prevalent in the two high risk groups (C and D) (tables 1 and online supplementary table 2E). Although both HAN et al. [3] and LANGE et al. [5] identified some of these differences, the level of characterisation of the participants in ECLIPSE is much more detailed so, this information constitutes an important repository of data (tables 1 and online supplementary tables 2E, 3E and 4E) of potential use for a better understanding of the disease and for the design of future studies. 2) Our analysis is the first to investigate the temporal stability of the four 2011 GOLD groups. We observed that a relative majority of A and D patients do actually remain in their original classification, whereas B and C patients present higher temporal variability (fig. 2). These changes can be due to either disease progression, or the beneficial effect of therapy but, in any case, this observation provides evidence of the degree of heterogeneity of the natural history of the disease and may highlight a potential window of opportunity for therapeutic intervention. 3) Also not explored previously [3, 5, 6], we found that the 2011 GOLD assessment proposal does not separate the COPD cohort into groups that differ in their rate of FEV1 decline (table 1). This suggests that further segmentation of the COPD population beyond that proposed by 2011 GOLD assessment may be required to fully describe COPD heterogeneity, since several recent studies [20, 21], including an analysis of the ECLIPSE cohort [15], have clearly shown that the rate of FEV1 decline varies among COPD patients. 4) No previous study has investigated the relationship of the four 2011 GOLD groups with hospitalisations during follow-up. We found that, similarly to mortality, hospitalisations were less prevalent in A, most in D and intermediate and similar in the B and C groups (fig. 3). As in the case of mortality discussed above, this observation emphasises a, sometimes overlooked, fact; i.e. that patients with COPD may require hospitalisation (and/or die) for reasons not directly related to COPD (comorbidities). Because they are amenable to therapy [22], this observation supports the emphasis made by the 2011 GOLD document on the importance of actively looking for, and adequately treating if present, the comorbidities known to frequently occur in COPD patients [2]. 5) Since ECLIPSE does not have information on specific causes of death, we explored factors associated with increased risk of mortality (and hospitalisations) using a Cox proportional hazards regression model. Interestingly the identified risk factors were similar for mortality and hospitalisations (table 2). The fact that they were distributed relatively evenly between B and C patients (tables 1 and online supplementary table 2E) may in part explain the relatively similar outcomes exhibited by these two groups despite their different severities of airflow limitation and/or past history of exacerbations, and may help clinicians to assess the risk of individual patients. 6) We explored, for the first time, the potential gender differences among 2011 GOLD groups. We found that their frequency distribution was not different in males or females (online supplementary table 1E). Furthermore, we observed that there were no significant gender differences for mortality or hospitalisations although, interestingly, females had a higher (p⩽0.017) exacerbation rate during follow-up than males (online supplementary table 1E).

Potential limitations

The following potential limitations of our study deserve comment. 1) Like all previous studies [3, 5, 6], our analysis took advantage of an existing cohort assembled for different purposes, so it might not mirror the population of COPD patients at large. 2) Because patients enrolled in the ECLIPSE study were mostly recruited from hospital clinics and were treated according to the attending physician, our results may not be directly generalised to untreated patients or patients with milder disease (*i.e.* GOLD stage I). However, it is of interest that LANGE *et al.* [5] reported similar observations in patients recruited from the general population with a skewed distribution towards milder disease (77% classified in the A group). 3) Because of the inclusion criteria used in the ECLIPSE study (age 40–75 years, baseline post-bronchodilator FEV_1 $<$80% of the predicted value, baseline post-bronchodilator FEV_1/FVC $\leqslant$0.7 and a smoking history of $\geqslant$10 pack-years [7]), it is possible that a small fraction of A and B patients analysed here may be above the lower limit of the normal. Yet, all of them were diagnosed of COPD by their attending physicians considering their exposure to risk factors and current clinical symptomatology. 4) Comorbidities were self-reported and we lack information on specific causes of death, so some of our observations and interpretations require prospective validation in future studies.

Conclusions

The different longitudinal behaviour of group A *versus* B, and C *versus* D (both pairs with similar FEV_1 values) supports the 2011 GOLD proposal of assessing COPD patients by more than just the degree of airflow limitation. However the assumption that symptoms do not equate to risk appears to be naïve, as groups B and C carry equally poor clinical outcomes though for different reasons. The latter may explain why group B patients are those most likely to switch group over time.

References

1 Agusti A, Calverley P, Celli B, *et al.* Characterisation of COPD heterogeneity in the ECLIPSE cohort. *Respir Res* 2010; 11: 122.

2 Vestbo J, Hurd SS, Agusti AG, *et al.* Global strategy for the diagnosis, management and prevention of chronic obstructive pulmonary disease, GOLD executive summary. *Am. J Respir Crit Care Med* 2013; 187: 347–365.

3 Han MK, Mullerova H, Curran-Everett D, *et al.* GOLD 2011 disease severity classification in COPDGene: a prospective cohort study. *Lancet Respi Med* 2012; 1: 1–8.

4 Regan EA, Hokanson JE, Murphy JR, *et al.* Genetic epidemiology of COPD (COPDGene) study design. *COPD* 2010; 7: 32–43.

5 Lange P, Marott JL, Vestbo J, *et al.* Prediction of the clinical course of chronic obstructive pulmonary disease, using the new GOLD classification. *Am J Respir Crit Care Med* 2012; 186: 975–981.

6 Soriano JB, Alfajame I, Almagro P, *et al.* Distribution and prognostic validity of the new global initiative for chronic obstructive lung disease grading classification. *Chest* 2013; 143: 694–702.

7 Vestbo J, Anderson W, Coxson HO, *et al.* Evaluation of COPD longitudinally to identify predictive surrogate end-points (ECLIPSE). *Eur Respir J* 2008; 31: 869–873.

8 Jones PW, Harding G, Berry P, *et al.* Development and first validation of the COPD Assessment Test. *Eur Respir J* 2009; 34: 648–654.

9 Standardization of spirometry, 1994 update. American Thoracic Society. *Am J Respir Crit Care Med* 1995; 152: 1107–1136.

10 ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med* 2002; 166: 111–117.

11 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.

12 Celli BR, Cote CG, Marin JM, *et al.* The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. *N Eng J Med* 2004; 350: 1005–1012.

13 Hurst JR, Vestbo J, Anzueto A, *et al.* Susceptibility to exacerbation in chronic obstructive pulmonary disease. *N Eng J Med* 2010; 363: 1128–1138.

14 Agustí A, Edwards LD, Rennard SI, *et al.* Persistent systemic inflammation is associated with poor clinical outcomes in COPD: a novel phenotype. *PLoS ONE* 2012; 7: e37483.

15 Vestbo J, Edwards LD, Scanlon PD, *et al.* Changes in forced expiratory volume in 1 second over time in COPD. *N Eng J Med* 2011; 365: 1184–1192.

16 Jonckheere AR. A distribution-free k-sample test against ordered alternatives. *Biometrika* 1954; 41: 133–145.

17 Rabe KF, Hurd S, Anzueto A, *et al.* Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* 2007; 176: 532–555.

18 Gonen M, Heller G. Concordance probability and discriminatory power in proportional hazards regression. *Biometrika* 2005; 92: 965–970.

19 Mannino DM, Thorn D, Swensen A, *et al.* Prevalence and outcomes of diabetes, hypertension and cardiovascular disease in COPD. *Eur Respir J* 2008; 32: 962–969.

20 Casanova C, de Torres JP, Aguirre-Jaíme A, *et al.* The progression of chronic obstructive pulmonary disease is heterogeneous: the experience of the BODE cohort. *Am J Respir Crit Care Mede* 2011; 184: 1015–1021.

21 Nishimura M, Makita H, Nagai K, *et al.* Annual change in pulmonary function and clinical phenotype in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2012; 185: 44–52.

22 Postma D, Anzueto A, Calverley P, *et al.* A new perspective on optimal care for patients with COPD. *Prim Care Respir J* 2011; 20: 205–209.