



FAQs about the GOLD 2011 assessment proposal of COPD: a comparative analysis of four different cohorts

Alvar Agusti^{1,2}, Suzanne Hurd³, Paul Jones⁴, Leonardo M. Fabbri⁵, Fernando Martinez⁶, Claus Vogelmeier⁷, Jørgen Vestbo^{8,9,10} and Robert Rodriguez-Roisin¹

Affiliations: ¹Thorax Institute, Hospital Clinic, IDIBAPS, Universitat de Barcelona, Barcelona, and ²FISIB, CIBER Enfermedades Respiratorias (CIBERES), Mallorca, Spain. ³Global Initiative for Chronic Obstructive Lung Disease, Vancouver, WA, and ⁴University of Michigan School of Medicine, Ann Arbor, MI, USA. ⁵St George's Hospital Medical School, London, and ⁶Manchester Academic Health Sciences Centre, South Manchester University Hospital NHS Foundation Trust, Manchester, UK. ⁷University of Modena and Reggio Emilia, Modena, Italy. ⁸Dept of Respiratory Medicine, University of Marburg, Marburg, Germany. ⁹Odense University Hospital, Odense, and ¹⁰University of Southern Denmark, Odense, Denmark.

Correspondence: A. Agusti, Thorax Institute, Hospital Clínic, Villarroel 170, 08036 Barcelona, Spain.
E-mail: Alvar.Agusti@clinic.ub.es

ABSTRACT Since the publication of the new Global Initiative for Chronic Obstructive Lung Disease (GOLD) proposal for the assessment of chronic obstructive lung disease (COPD), four studies have used existing cohorts to explore the characteristics, temporal variability and/or relationship with outcomes of the four resulting patient categories (A, B, C and D).

Here, we compare their results and address a number of frequently asked questions (FAQs) on the topic.

The most salient findings were that: 1) the prevalence of these four groups depends on the specific population studied, C being the least prevalent; 2) comorbidities are particularly prevalent in the two “high-symptom” groups (B and D); 3) patients classified as A or D tend to remain in the same group over time, whereas those classified as B or C change substantially during follow-up; 4) mortality at 3 years was lowest in A and worst in D but surprisingly similar (and intermediate) in B and C; and 5) the incidence of exacerbations during follow-up increases progressively from A to D but that of hospitalisations behave similarly to mortality.

These results identify several strengths and shortcomings of the new GOLD assessment proposal, particularly that group B is associated with more morbidity and high mortality.



@ERSpublications

GOLD group B COPD is associated with more morbidity and high mortality <http://ow.ly/nGkbP>

For editorial comments see pages 1163 and 1166.

The following article was not considered here as it was published after the submission of the present review: Johannessen A, Nilsen RM, Storebø M, *et al.* 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: 51–59.

Received: Feb 27 2013 | Accepted after revision: April 21 2013 | First published online: May 03 2013

Conflict of interest: Disclosures can be found alongside the online version of this article at www.erj.ersjournals.com

Copyright ©ERS 2013

The Global Initiative for Chronic Obstructive Lung Disease (GOLD) has recently proposed a new multidimensional system for the assessment and management of patients with chronic obstructive pulmonary disease (COPD) that combines the impact of the disease as perceived by the patient with the severity of airflow limitation and the past history of exacerbations [1]. As a result, COPD patients are now classified in four categories or groups (A, B, C and D) (fig. 1) that, along with the assessment of potential comorbidities, can assist in guiding their therapy [1].

This proposal was based on the recognition that COPD is a complex and heterogeneous disease and that the severity of airflow limitation (assessed by measuring forced expiratory volume in 1 s (FEV₁)) is poorly related to many other clinically relevant aspects of the disease [7]. In addition, recent trials have shown that the arbitrary cut-off values for FEV₁ did not match the application of pharmacological treatment. However, the revised GOLD strategy is an empirical proposal mostly based on expert opinion, as available evidence refers to management but not assessment of the disease and the studies providing evidence have classified patients according to the previous GOLD recommendations, which were based almost exclusively on FEV₁ [8] with or without a history of exacerbations. Not surprisingly, therefore, soon after the release of this new GOLD proposal, a number of investigators rushed to explore, in their existing cohorts (*i.e.* COPDGene [2], Copenhagen [3], Collaborative Cohorts to assess Multicomponent Indices of COPD in Spain [4] and ECLIPSE (Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints) [5]), the distribution, characteristics, temporal stability and/or relationship with long-term outcomes of these four patient categories. In this article, we: 1) review and compare the results of these four studies (table 1); and 2) list a number of frequently asked questions (FAQs) about this GOLD assessment proposal and provide some answers based on these published results. Overall, this review provides a global view of the strengths and limitations of the new GOLD recommendations [1].

The GOLD 2011 assessment proposal

The GOLD 2011 assessment proposal of COPD includes two dimensions: the impact of the disease as perceived by the patient and the risk of future exacerbations (fig. 1) [1]. To classify patients into the “low-symptom” or “high-symptom” groups, GOLD 2011 recommends the use of the modified Medical Research Council (mMRC) scale of breathlessness (<2 or ≥2, respectively) or the COPD Assessment Test (CAT) (<10 or ≥10, respectively). To assess the risk of future exacerbations, GOLD 2011 suggests the use of two different (and poorly correlated) criteria: the severity of airflow limitation, as assessed by FEV₁ (<50% predicted), and the previous individual history of exacerbations (two or more exacerbations per year, or one or more hospitalisation per year) [1]. Whenever there is a discrepancy between the risk category, as assessed by the FEV₁, and/or the exacerbation history, the variable indicating the highest risk should be used [1]. Notably, very recently (February 2013), the annually updated version of the GOLD *Global Strategy for Diagnosis, Management, and Prevention of COPD* [6] has proposed that the Clinical COPD Questionnaire (CCQ) [9] can also be used as an alternative to CAT to assess the symptomatic impact of the disease and

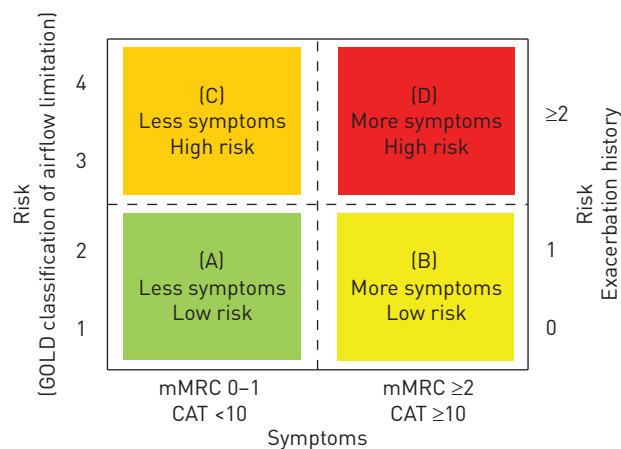


FIGURE 1 Four patient categories (or groups) resulting from the application of the 2011 Global Initiative for Chronic Obstructive Lung Disease (GOLD) assessment proposal [1]. For further details, see the main text. Although this is the classification used for the four studies reviewed here [2–5], it is of note that the 2013 update [6] proposes to add (not included in the figure) the Clinical COPD Questionnaire among the instruments that can be used to assess symptoms and to consider as high risk any patient who has had at least one hospitalisation due to chronic obstructive pulmonary disease (COPD) during the past year (also not included in the figure). mMRC: modified Medical Research Council dyspnoea score; CAT: COPD Assessment Test score. Reproduced and modified from [6] with permission from the publisher.

that one hospitalisation due to an exacerbation of COPD suffices to classify the patient into the “high-risk” groups (C and D). Yet, these changes were not considered by the four papers reviewed herein [2–5] because they were published before the GOLD 2013 update release [6].

According to the results of this assessment, four patient groups (or categories) can be identified (fig. 1) [1]. It is important to note, however, that groups A and B are defined unequivocally according to risk (FEV₁ ≥ 50% pred and no more than one exacerbation per year), so any given patient must fulfil these two criteria to be classified in one of these two groups (fig. 1). This contrasts with groups C and D, which admit several potential combinations. For instance, a given patient can be classified in group C because the presence of fewer symptoms (by definition), and either FEV₁ < 50% pred, or two or more exacerbations/year (or more than one hospitalisation per year in the GOLD 2013 update [6]), or both (the same occurs with patients in group D) (fig. 1). In fact, some of the studies discussed here [2, 5], but not the original GOLD document [1], have named these three subgroups as C₁, C₂ and C₃, and D₁, D₂ and D₃, respectively (see later).

This combined assessment proposal [1] is expected to reflect the complexity of COPD better than the unidimensional (FEV₁) analysis previously used for staging the disease [8] and is therefore a potentially important step forward towards a more personalised approach to COPD [10]. However, to some extent, it is an empirical proposal based upon expert opinion and, hence, it requires experimental validation [1]. Besides, the categorisation is constructed to reflect different needs for management and, in contrast to the old staging, categorisation may not necessarily reflect severity.

Published results on existing cohorts

COPDGene

The COPDGene study is a multicentre, observational investigation aimed at identifying genetic variants related to COPD and thoracic computed tomography-defined COPD phenotypes [11]. An initial group of 4000 smokers (with and without COPD), including both non-Hispanic white American and African-American individuals across all severities of airflow limitation, were recruited and assessed. Findings will be verified in an additional 2000 COPD cases and 2000 smoking control subjects [11]. For the comparative analysis discussed here, it is important to note that this study included participants not previously diagnosed with COPD [11] and that such patients may have lower symptom levels, as assessed by instruments such as the St George’s Respiratory Questionnaire (SGRQ), than patients diagnosed with COPD [12].

HAN *et al.* [2] used information from 4484 COPD patients (table 1) included in the COPDGene study to investigate: 1) whether or not the two instruments proposed by the GOLD 2011 revision to assess the level of symptoms of the patients (mMRC or CAT) produced similar results on patient group assignment (the COPDGene study did not have CAT data, so the SGRQ (≥ 25 *versus* < 25) was used as a surrogate for the CAT threshold (≥ 10 *versus* < 10, respectively) proposed by GOLD to classify patients in the “less” or “more” symptoms groups); and 2) the relationship of the four patient groups (A, B, C and D) with the frequency of exacerbation recorded prospectively during telephonic follow-up (1.4 years).

Main results were as follows. 1) 33.6% *versus* 29.4% of patients were assigned to group A when the mMRC or the SGRQ (as a surrogate of CAT) criteria were used, respectively (fig. 2). These percentages were 20.5% *versus* 24.7% in group B, 7.9% *versus* 4.9% in group C, and 38.0% *versus* 41.0% in group D (κ-coefficient

TABLE 1 Main demographic, clinical and functional variables of the chronic obstructive pulmonary disease (COPD) patients included in each of the four studies analysed here

	COPDGene [2]	Copenhagen [3]	Cocomics [4]	ECLIPSE [5]
Subjects n	4484	6628	3163	2101
Recruitment source	Hospitals	General population	7 cohorts merged	Hospitals
Country	USA	Denmark	Spain	Europe/USA
Age years	65	68	66	64
Males %	53	52	93	65
BMI kg·m⁻²	27	25	28	27
FEV₁ % pred	40	40	54	50
Follow-up years	1.7	4.3	10	3

For further explanation, see the main text. Cocomics: Collaborative Cohorts to assess Multicomponent Indices of COPD in Spain; ECLIPSE: Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints; BMI: body mass index; FEV₁: forced expiratory volume in 1 s; % pred: % predicted.

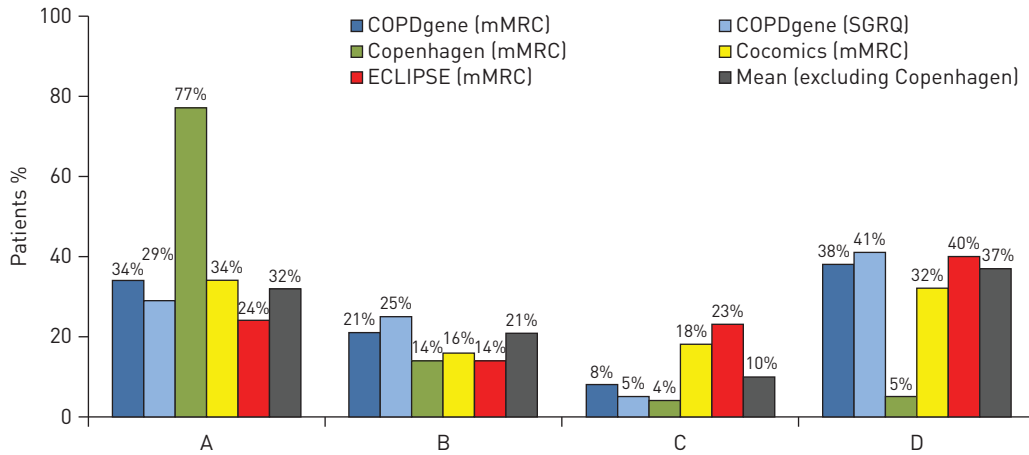


FIGURE 2 Frequency distribution of the four patient groups observed in the four studies compared here [2–5]. The right-hand columns indicate the mean value (excluding data from the Copenhagen study, the only one that recruited patients from the general population [3]). For further explanation, see the main text. mMRC: modified Medical Research Council dyspnoea scale; ECLIPSE: Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints; SGRQ: St George’s Respiratory Questionnaire; Cocomics: Collaborative Cohorts to assess Multicomponent Indices of COPD in Spain.

for agreement 0.77) (fig. 2). 2) Depending on the specific risk factor used to determine category assignment in the high-risk groups (FEV₁ only (C₁ and D₁), previous exacerbation history only (C₂ and D₂) or both (C₃ and D₃)) (fig. 3), significant heterogeneity in prospective exacerbation rates was observed, particularly in group D. In this latter group, the mean (95% CI) annual rate of exacerbations observed during follow-up in the D₁, D₂ and D₃ subgroups was, respectively, 0.89 (0.78–1.00), 1.34 (1.0–1.6) and 1.86 (1.6–2.1) (p<0.0001).

From these results, HAN *et al.* [2] concluded that: 1) the instrument used to measure symptoms influences category assignment; 2) the prevalence of group C patients (low symptoms and high risk) is low; and 3) exacerbation rates appear different depending on whether or not high-risk categorisation was based on lung function, exacerbation history or both.

Copenhagen study

For this analysis, LANGE *et al.* [3] pooled data from two similar but independent general population studies: the fourth examination of the Copenhagen City Heart Study (CCHS) from 2001–2003 and the examination

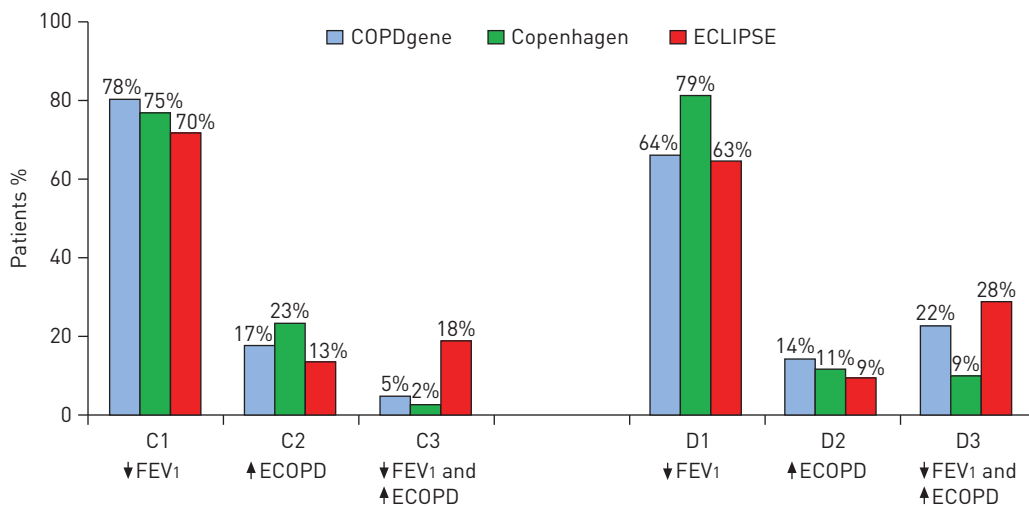


FIGURE 3 Prevalence of group C and D subtypes in three out of the four studies compared here [2, 3, 5]. In all of them, the main reason for categorisation as a high-risk patient (groups C and D) is a forced expiratory volume in 1 s (FEV₁) <50% predicted alone. For further explanation, see the main text. ECOPD: exacerbations of chronic obstructive pulmonary disease (COPD); ECLIPSE: Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints.

of the Copenhagen General Population Study (CGPS) from 2003–2010. The CCHS is a prospective epidemiological study that started in 1976–1978 and selected a random sample of subjects from the national Danish Civil Registration System. A total of 14 223 residents of inner Copenhagen participated in the initial survey [13], 6237 of whom attended the fourth survey in 2001–2003 [14]. The CGPS is a prospective epidemiological study, the design of which is almost identical to the CCHS, aiming to recruit >100 000 individuals representative of the general population, and to collect genotypic and phenotypic data of relevance to a wide range of health-related problems; recruitment began in 2003 and is still in progress [15]. Merging data from these two general cohorts, LANGE *et al.* [3] identified 6628 COPD patients (table 1) who were monitored for >4 years for COPD exacerbations, hospital admissions and mortality. In this convenience cohort, they sought to investigate the ability of the new GOLD 2011 assessment proposal (using the mMRC criteria to categorise symptoms) to predict the clinical course of COPD [3].

The main results showed that: 1) the vast majority of patients in this study belonged to group A (fig. 2), probably reflecting that this cohort of patients was identified from the general population; 2) the proportion of patients experiencing a COPD exacerbation during the first year of observation increased progressively from groups A to B to C to D (2.2%, 5.8%, 25.1% and 28.6%, respectively); and 3) at the 3-year follow-up, mortality rates were 3.8%, 10.6%, 8.2% and 20.1% in groups A, B, C and D, respectively (fig. 4) [3]. Notably, groups B and D, both characterised by a higher degree of dyspnoea, had five to eight times higher mortality from cardiovascular disease and cancer than groups A and C.

From these results, LANGE *et al.* [3] concluded that: 1) the new GOLD 2011 assessment proposal performs well by identifying individuals at risk of exacerbations; and 2) despite being classified by GOLD as low risk, patients in group B had significantly poorer survival than those in group C (classified as high risk). They suggested that the poor prognosis of group B patients can be related to comorbidities, such as cardiovascular disease or cancer, hence requiring special assessment and treatment [3]. An important aspect of this study to consider is the fact that whereas GOLD recommendations relate to patients with a clinical diagnosis of COPD, epidemiological studies rely heavily on spirometry.

Cocomics study

The Cocomics study is a pooled analysis of individual patient data (age, sex, mMRC score, post-bronchodilator spirometry and all-cause mortality) from 11 COPD cohorts recruited in seven different cities in Spain (Galdakao, Pamplona, Requena, Seville, Tenerife, Terrassa and Zaragoza) for different purposes. In this pooled population (table 1), SORIANO *et al.* [4] sought to: 1) determine the distribution of the four groups proposed by the GOLD 2011 revision; and 2) compare its validity in predicting all-cause mortality up to 10 years, as compared to the previous GOLD staging system based mostly on the FEV₁ value alone. Other outcomes, namely exacerbations, were not reliably recorded and were not included in the analysis [4].

The main results were as follows. 1) Of the 3633 patients included in the analysis, 1064 (33.6%) were classified as group A, 515 (16.3%) as group B, 561 (17.7%) as group C and 1023 (32.3%) as group D (fig. 2). This distribution, however, varied significantly (Chi-squared $p < 0.01$) between the 11 cohorts pooled [4]. 2) The ability of GOLD 2007 to predict mortality was not different from that of GOLD 2011 at 1 year (0.635 *versus* 0.639, $p = 0.53$), 3 years (0.637 *versus* 0.645, $p = 0.21$) or 10 years (0.639 *versus* 0.642, $p = 0.76$). Again, however, survival varied greatly between cohorts [4].

From these results, SORIANO *et al.* [4] concluded that: 1) the GOLD 2011 assessment proposal results in an uneven split of the COPD population; and 2) its capacity to predict mortality is no different from the old GOLD staging approach based in FEV₁ only. It should be remembered, however, that the multidimensional GOLD assessment proposal was put forward to provide a structured assessment to guide treatment directed towards symptoms and exacerbations, not to provide a prognostic marker of mortality.

ECLIPSE study

ECLIPSE is a multicentre, international, longitudinal study aimed at identifying clinically relevant COPD subtypes (phenotypes) and the genetic factors and/or biomarkers that correlate with them and predict disease progression [16]. ECLIPSE included 2164 clinically stable COPD patients, 337 smokers with normal lung function and 245 never-smokers, who were extensively characterised and followed up for 3 years [7]. AGUSTI *et al.* [5] explored the distribution, characteristics, temporal stability and relationship with outcomes (exacerbations, hospitalisations and all-cause mortality) of 2101 COPD patients with complete mMRC, spirometry and previous exacerbation history according to the 2011 GOLD assessment proposal (table 1).

The main results were as follows. 1) In addition to the expected differences between groups in the three variables that define them (mMRC, FEV₁ and previous exacerbations), the four groups differed also in many other clinical characteristics studied (table 2). Hence, pulmonary emphysema and arterial

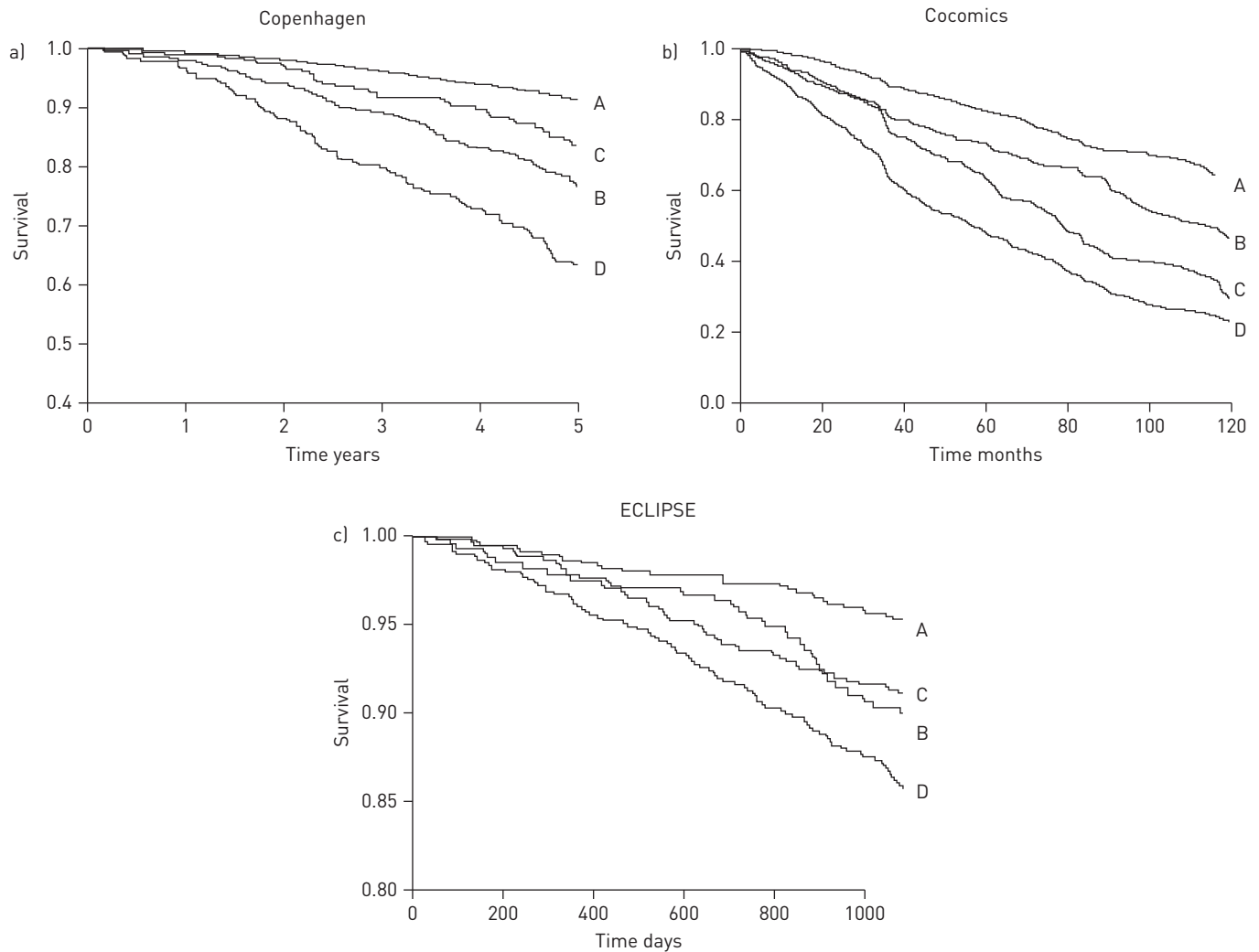


FIGURE 4 Kaplan–Meier survival curves (all-cause mortality) in three of the studies compared here [3–5]. For further explanations, see the main text. Cocomics: Collaborative Cohorts to assess Multicomponent Indices of COPD in Spain; ECLIPSE: Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints.

oxygenation impairment were particularly prevalent in the two high-risk categories (groups C and D) whereas comorbidities and persistent systemic inflammation were worse in the two highly symptomatic categories (groups B and D). By contrast, age, sex, FEV₁ reversibility and FEV₁ decline were not different between groups, although the latter was numerically higher in patients in the B category (table 2) [5]. 2) An FEV₁ <50% pred was the most frequent determinant of being classified as a group C (C₁, 70%) or D (D₁, 63%) patient, whereas a history of frequent exacerbations exclusively was the least prevalent (subgroups C₂ (13%) and D₂ (9%)) (fig. 3). 3) Group A and D patients were relatively stable over time, whereas those in groups B and C showed marked variability, some patients improving and others deteriorating during follow-up [5]. 4) Finally, the incidence of exacerbations during follow-up increased progressively ($p < 0.001$) from groups A to B to C to D [5]. By contrast, and similar to the findings of LANGE *et al.* [3] in the Copenhagen study, hospitalisations (fig. 4) and all-cause mortality were lowest in group A, highest in group D ($p \leq 0.010$), and intermediate but similar in groups B and C [5]. Notably, in this context, group B patients had less severe airflow limitation than group C patients but the highest prevalence of comorbidities (table 2) and persistent systemic inflammation [7].

From these results, AGUSTI *et al.* [5] concluded that: 1) there is significant clinical heterogeneity across the four GOLD 2011 groups; 2) the two extreme categories (A and D) were relatively stable over time but the two intermediate ones (groups B and C) show greater temporal variability, probably in relation to disease progression and/or response to therapy; 3) the new GOLD assessment proposal is feasible and valid for assessing the risk of future exacerbations (which increase steadily from groups A to B to C to D) but does

TABLE 2 The main clinical characteristics and associated outcomes of the four Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2011 categories in the ECLIPSE cohort [7]

Distribution	GOLD 2011 groups				p-value
	A	B	C	D	
GOLD 2011 assessment variables					
mMRC dyspnoea score	0.6	2.3	0.8	2.6	<0.001
FEV ₁ % pred	64.2	61.6	42.2	37.2	<0.001
Exacerbations previous 12 months n	0.2	0.3	1.0	1.3	<0.001
Demographics					
Age years	63.2	64.1	62.9	63.6	NS
Females %	39	36	31	34	NS
BMI kg·m ⁻²	26.7	28.5	25.2	26.5	<0.001
Smoking pack-years	48.0	47.5	45.2	51.1	<0.001
Current smokers %	38	34	43	32	0.001
Symptoms and comorbidities					
Chronic bronchitis %	26	35	38	38	<0.001
SGRQ-C total score	31.9	55.3	44.6	62.2	<0.001
Comorbidities n	2.0	2.8	1.7	2.3	<0.001
Physiology and imaging					
FEV ₁ reversibility %	11.4	11.8	10.3	10.2	NS
SaO ₂ %	95.7	95.1	94.4	93.7	<0.001
6MWD m	440	360	408	307	<0.001
Emphysema CT LLA [#] %	11.8	12.5	18.9	22.5	<0.001
Outcomes during follow-up[¶]					
ECOPDs per year n	0.6	0.9	1.3	1.7	<0.001
≥1 hospitalisation %	11	25	30	46	<0.001
Mortality rate %	4	10	8	14	<0.001
Rate of FEV ₁ decline mL·year ⁻¹	-33.4	-38.0	-30.2	-31.9	NS

Data are presented as mean values. For further explanation, see the main text. COPD: chronic obstructive pulmonary disease; ECLIPSE: Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints; mMRC: modified Medical Research Council; FEV₁: forced expiratory volume in 1 s; % pred: % predicted; BMI: body mass index; SGRQ-C: St George's Respiratory Questionnaire for COPD Patients; SaO₂: arterial oxygen saturation; 6MWD: 6-min walk distance; CT: computed tomography; LAA: low-attenuation area; ECOPD: exacerbation of COPD; NS: nonsignificant. #: -950 HU; ¶: 3 years. Reproduced and modified from [5].

not discriminate the risk of future hospitalisations and all-cause mortality for group B and C patients, probably highlighting the importance of symptom (and comorbidity) assessment in clinical practice.

Frequently asked questions

The comparative analysis of these four studies [2–5] helps us to respond some FAQs regarding the GOLD assessment proposal but poses other questions that will inevitably require further research.

What is the prevalence of these four groups?

The answer depends on the population studied (fig. 2). In the general population, the most prevalent group is A (77%) [3], whereas in patients recruited from secondary and tertiary care, it is D (37%) followed closely by, surprisingly, A (32%) [2, 4, 5]; the prevalence of group B is ~20% and, clearly, group C is the least prevalent (10%), although it is also the one with the highest variability between studies (fig. 2).

Does the instrument used to determine the level of symptoms matter?

HAN *et al.* [2] compared the mMRC and SGRQ scores (as a surrogate of CAT) in COPDGen and concluded that concordance between the two was good ($\kappa=0.77$) but the distribution of the four groups (fig. 2) was significantly different. In a more recent European study in primary care, a direct comparison of mMRC and CAT showed lower concordance ($\kappa=0.63$) between the two instruments [17]; yet, agreement improved ($\kappa=0.79$) if mMRC ≥ 1 was used as the cut-point instead of the threshold suggested by GOLD (≥ 2) [1]. However, if a broader perspective is taken (fig. 2), the impact of the specific instrument used to determine the level of symptoms may be no greater than other potential confounders, such as the specific type of cohort studied. In any case, a standardised approach would be desirable if future clinical studies are to be compared. In this context, it is worth noting that, although historically, there are a lot of data using

the Medical Research Council scale, this instrument reflects only one aspect of COPD (*i.e.* breathlessness), whereas the CAT and the CCQ (now included in the GOLD 2013 revision [6]) are more comprehensive.

Do these groups differ in other clinical characteristics than those used for their own definition?

Three studies [2, 3, 5] identified differences between groups besides those used for their own definition (table 2). For instance, in ECLIPSE, the prevalence of comorbidities and systemic inflammation was highest in the high-symptom groups (B and D), whereas the presence of emphysema and lower arterial oxygen saturation was highest in the high-risk groups (C and D) [5]. By contrast, other potentially relevant clinical variables, including age, sex and FEV₁ reversibility, were not different across groups [5]. The precise nature of these comorbidities in ECLIPSE has been described in detail elsewhere [7] but includes the most prevalent ones, such as cardiovascular disease, metabolic syndrome, osteoporosis and depression [18].

Are there distinct subgroups in high-risk categories?

Three studies [2, 3, 5] explored this question and came to very similar conclusions. The most frequent reason to categorise a given patient as high risk (*i.e.* groups C and D) was an FEV₁ <50% pred (C₁ or D₁) (fig. 3). By contrast, that a patient was considered high risk based only on the previous history of exacerbations (*i.e.* with an FEV₁ >50% pred; C₂ or D₂) was relatively rare (fig. 3). The implications of these differences in terms of prognosis or optimal care deserve future research.

Do these four groups of patients remain stable over time?

Only the ECLIPSE study addressed this question [5]. The results showed that the relative majority of patients assigned to groups A and D continue to belong to these same two groups 3 years later, whereas those classified in groups B or C tend to vary more (fig. 5). The mechanisms explaining these temporal changes (or lack of changes) also deserve future research but they are probably related to either so-called disease progression or response to therapy. Importantly, these findings show that there is also heterogeneity in the disease progression, which might present opportunities for therapeutic intervention [6].

Do these four groups predict mortality?

The GOLD assessment proposal was not designed to predict mortality. Yet, the four GOLD categories do indeed relate to mortality, albeit in a surprising manner. Both LANGE *et al.* [3], in the general population recruited in Copenhagen, and AGUSTI *et al.* [5], in the ECLIPSE cohort, found that survival at 3 years was better for patients in group A, worse for group D, and intermediate and similar for groups B and C (fig. 3). This is consistent with previous studies on the prognostic importance of symptom perception [19]. In Cocomics, survival at 3 years was also similar in groups B and C, but results changed after 10 years' follow-up (fig. 4) [4], perhaps indicating a survival bias.

With respect to the precise causes of death in these four patient categories, LANGE *et al.* [3] showed that groups B and D (both "high symptoms") had five to eight times higher mortality from cardiovascular disease and lung cancer than the fewer-symptoms groups (A and C). Specific causes of death were not investigated in COPDGene [2] or Cocomics [4], whereas in ECLIPSE, Cox proportional hazards regression analysis identified a number of factors associated with increased risk of mortality (severity of airflow limitation, older age, lower body mass index, increased number of comorbidities and lower exercise capacity) [5]. This information may help clinicians to assess the risk of death in individual patients.

Do these four groups of patients predict exacerbations/hospitalisations?

All four studies showed that the incidence of exacerbations increases progressively from A to B to C to D [2–5]. This is not surprising, as the history of previous exacerbation, the best current predictor of future exacerbations [20], is one of the variables used to define these four groups [1]. By contrast, the capacity to predict hospitalisations was explored in ECLIPSE only [5] and, like mortality (see earlier), hospitalisations during follow-up were scant in group A, frequent in group D, and intermediate (and similar) in groups B and C patients. Furthermore, identified risk factors for hospitalisations were similar to those identified for mortality (see earlier) [5].

Why do groups B and C appear to have similar risk of death/hospitalisation?

Both LANGE *et al.* [3] and AGUSTI *et al.* [5] suggested that high symptoms in group B may originate from comorbidities and not from airflow limitation. In fact, LANGE *et al.* [3] showed in the general population that group B patients frequently die of cardiovascular disease or lung cancer. Likewise, in the ECLIPSE cohort, group B patients were also characterised by the highest prevalence of comorbidities and presence of persistent systemic inflammation [5]. The former is known to have a direct and significant impact on survival in COPD [21] while the latter has been recently shown to increase mortality six-fold irrespective of the severity of the underlying pulmonary abnormalities [22]. All in all, these observations highlight that

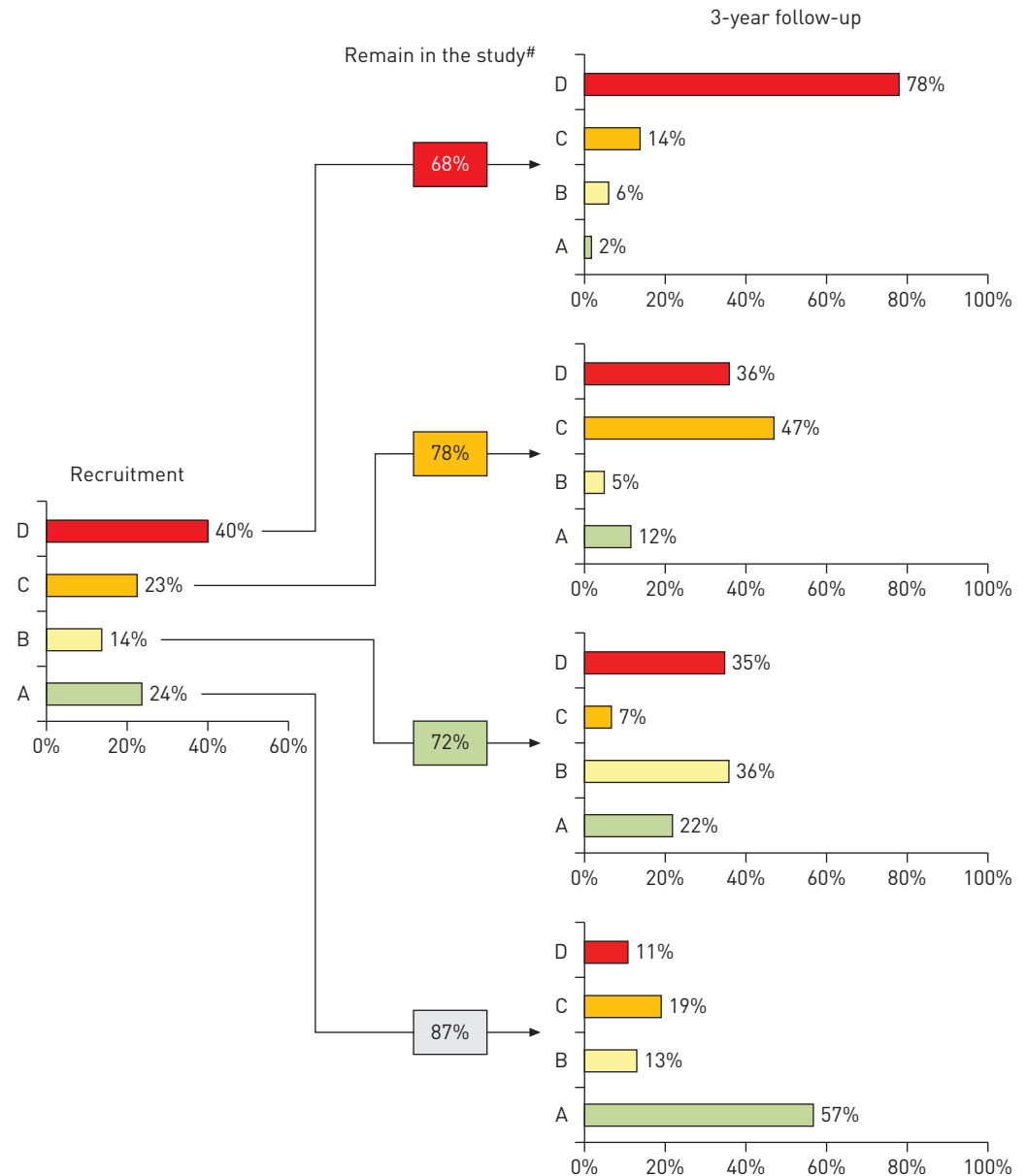


FIGURE 5 Temporal stability of the four patient categories in the Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints. Boxes with percentages indicate the proportion of baseline patients available for examination at the 3-year follow-up. For further explanation, see the main text. #: death/drop-out/lost for follow-up/missing data. Reproduced from [5].

COPD patients may need hospitalisation (or die) during follow-up for reasons not directly related to their lung function (*e.g.* comorbidities) and support the GOLD recommendation on the importance of actively assessing (and treating if present) the most frequent comorbidities in these patients [1].

Do these four groups have different rates of annual lung function decline?

ECLIPSE addressed this question and showed that there were no significant differences in the rate of lung function decline between the four groups (table 2) [5]. This lack of differences suggests that further segmentation of the COPD population beyond that proposed by GOLD may be required to better describe COPD heterogeneity in full [5].

Do these groups have different patterns and/or severity of comorbidities?

Comorbidities appear to be more prevalent in group B patients [3, 5] and, as discussed above, they may contribute to the increased risk of hospitalisations and death in this group (fig. 4). Notably, however, in the

four studies reviewed here [2–5], comorbidities were either identified from medical records or self-reported, but not actively identified by clinical investigations [23]. This may be clinically relevant, as most comorbidities are undiagnosed. Despite that the new GOLD recommendations [1] clearly highlight the important role of comorbidities in COPD [21, 23, 24], a current challenge is how to incorporate them explicitly in the proposed assessment diagram. Ongoing discussion is addressing the importance of comorbidities for the definition of COPD severity and how to incorporate them in the A–D diagram or a different diagram [25].

Do these four groups of patients require different therapeutic strategies?

To date, there are no studies where patients have been investigated on the basis of the A–D categorisation as inclusion criteria; the four studies reviewed here [2–5] are observational and none tested the effect of any particular therapeutic intervention. Therefore, from this comparative analysis, no specific therapeutic recommendations can be made to support or refute those provided by GOLD, which are the result of expert consensus based on previous randomised clinical trials (RCTs). Hopefully, this much-needed evidence will become available in the near future either by retrospective analysis of already published RCTs or, ideally, by new prospective ones.

Strengths and limitations of this analysis

As in any analysis, the current one has strengths and shortcomings. Among the former, this is the first paper comparing four recent studies [2–5] that, in turn, explore several relevant aspects related to the potential validations of the recent GOLD recommendations [1]. Hence, the results of the current analysis can be of importance for the translation of the GOLD recommendations into practice, for the design of future studies and for the eventual refinement of the GOLD document itself. However, several limitations need to be acknowledged. Firstly, the analysis compares four studies, all of which use existing cohorts of COPD patients that were established for different purposes, so results will have to be replicated prospectively in other populations. Secondly, all of the included studies used observational cohorts, so in none of them was treatment controlled; rather, it was dictated by the local physician taking care of the patient [2–5]. More importantly, all patients were treated, so these are not studies describing the “natural history of COPD” but the natural history of treated COPD.

Conclusions

This comparative analysis identifies a number of similarities and differences across four recently published studies [2–5] that investigated the clinical implications of the new GOLD recommendations [1]. From this analysis, a number of FAQs can be addressed but other new questions arose that may require further research. Altogether, though, the results discussed herein constitute a repository of potentially useful information for practitioners, investigators, administrators, regulators, stakeholders and guideline writers. However, it is now time for new studies to be designed and performed prospectively in these four patient groups, both to relate them to severity and prognosis and/or comorbidities, and to assess their response to treatment.

References

- 1 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.
- 2 Han MK, Muellerova H, Curran-Everett D, *et al.* GOLD 2011 disease severity classification in the COPDGene study: a prospective cohort study. *Lancet Respir Med* 2013; 1: 43–50.
- 3 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.
- 4 Soriano JB, Alfajame I, Almagro P, *et al.* Distribution and prognostic validity of the new GOLD grading classification. *Chest* 2013; 143: 694–702.
- 5 Agusti A, Edwards L, Celli B, *et al.* Characteristics, stability and outcomes of the GOLD 2011 COPD groups in the ECLIPSE cohort. *Eur Respir J* 2013; 42: 637–647.
- 6 Global Initiative for Chronic Obstructive Lung Disease. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. www.goldcopd.org/uploads/users/files/GOLD_Report_2013_Feb20.pdf Date last updated: February 20, 2013.
- 7 Agusti A, Calverley P, Celli B, *et al.* Characterisation of COPD heterogeneity in the ECLIPSE cohort. *Respi Res* 2010; 11: 122–136.
- 8 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.
- 9 Reda AA, Kotz D, Kocks JW, *et al.* Reliability and validity of the clinical COPD questionnaire and chronic respiratory questionnaire. *Respir Med* 2010; 104: 1675–1682.
- 10 Agusti A, Sobradillo P, Celli B. Addressing the complexity of chronic obstructive pulmonary disease: from phenotypes and biomarkers to scale-free networks, systems biology, and P4 medicine. *Am J Respir Crit Care Med* 2011; 183: 1129–1137.
- 11 Regan EA, Hokanson JE, Murphy JR, *et al.* Genetic epidemiology of COPD (COPDGene) study design. *COPD* 2010; 7: 32–43.

- 12 Miravitlles M, Soriano JB, Garcia-Rio F, *et al.* Prevalence of COPD in Spain: impact of undiagnosed COPD on quality of life and daily life activities. *Thorax* 2009; 64: 863–868.
- 13 Appleyard M, Hansen AT, Schnohr P, *et al.* The Copenhagen City Heart Study: a book of tables with data from the first examination (1976–78) and a five-year follow-up (1981–83). *Scand J Soc Med* 1989; 170: 1–160.
- 14 Lange P, Mogelvang R, Marott JL, *et al.* Cardiovascular morbidity in COPD: a study of the general population. *COPD* 2010; 7: 5–10.
- 15 Stender S, Frikke-Schmidt R, Nordestgaard BG, *et al.* Sterol transporter adenosine triphosphate-binding cassette transporter G8, gallstones, and biliary cancer in 62,000 individuals from the general population. *Hepatology* 2011; 53: 640–648.
- 16 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.
- 17 Jones P, Adamek L, Nadeau G, *et al.* Comparisons of health status scores with MRC grades in a primary care COPD population: implications for the new GOLD 2011 classification. *Eur Respir J* 2013; 42: 647–654.
- 18 Fabbri LM, Luppi F, Beghe B, *et al.* Complex chronic comorbidities of COPD. *Eur Respir J* 2008; 31: 204–212.
- 19 Nishimura K, Izumi T, Tsukino M, *et al.* Dyspnea is a better predictor of 5-year survival than airway obstruction in patients with COPD. *Chest* 2002; 121: 1434–1440.
- 20 Hurst JR, Vestbo J, Anzueto A, *et al.* Susceptibility to exacerbation in chronic obstructive pulmonary disease. *N Engl J Med* 2010; 363: 1128–1138.
- 21 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.
- 22 Agusti 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.
- 23 Vanfleteren LE, Spruit MA, Groenen M, *et al.* Clusters of comorbidities based on validated objective measurements and systemic inflammation in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2013; 187: 728–735.
- 24 Divo M, Cote C, de Torres JP, *et al.* Comorbidities and risk of mortality in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2012; 186: 155–161.
- 25 Decramer M, Janssens W. Chronic obstructive pulmonary disease and comorbidities. *Lancet Respir Med* 2013; 1: 73–83.