ERJ Express. Published on May 3, 2013 as doi: 10.1183/09031936.00036513

April 23, 2013

FREQUENTLY ASKED QUESTIONS (FAQ's) ABOUT THE GOLD 2011 ASSESSMENT PROPOSAL OF COPD

A comparative analysis of four different cohorts

Alvar Agusti[1,2], Suzanne Hurd[3], Paul Jones[4], Leonardo Fabbri[5], Fernando Martinez[6], Claus Vogelmeier[7], Jørgen Vestbo[8] and Robert Rodriguez-Roisin[1]

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

Correspondence: Dr. Alvar Agustí. Thorax Institute. Hospital Clínic. Villarroel 170, 08036 Barcelona, Spain. Tel: +34 93 227 1701; Fax: +34 93 227 9868; e-mail: Alvar.Agusti@clinic.ub.es

Running title: FAQ's about GOLD COPD assessment

Key words: Assessment − Chronic Obstructive Pulmonary Disease (COPD) − Management − Outcomes − Pharmacology − Therapeutic Strategy

Conflict of interests: All authors are members of the Scientific Committee of GOLD.

Main text word count (excluding abstract, references, Tables and Figure legends):

4,195 words

References: 25

Tables: 2

Figures: 5

ABSTRACT

Since the publication of the new GOLD proposal for the assessment of 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, 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, being C the less prevalent one; *(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 three years was lowest in A and worse 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 hospitalizations 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.

Abstract word count: 191 words

INTRODUCTION

The Global Initiative for Chronic Obstructive Lung Disease (GOLD) strategy 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) (**Figure 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 (FEV_1) is poorly related to many other clinically relevant aspects of the disease [2]. In addition, recent trials have shown that the arbitrary cut-off values for FEV_1 did not match the application of pharmacological treatment. However, the revised GOLD strategy is an empirical proposal mostly based on expert opinion, since 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 the FEV_1 [3], 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 [4], Copenhagen [5], Cocomics [6] and ECLIPSE [7]) the distribution, characteristics, temporal stability and/or relationship with long-term outcomes of these four patient categories. In this paper, we: *(1)* review and compare the results of these four studies (**Table 1**); and, *(2)* list a number of frequently asked questions (FAQ's) about this GOLD assessment proposal and provide

some answers based on these published results. Overall, this paper 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 (**Figure 1**) [1]. To classify patients in the 'Low or High Symptom' groups, GOLD 2011 recommends the use of the modified British Medical Research Council scale of breathlessness (mMRC; $<$ or ≥ 2, respectively) or the COPD Assessment Test (CAT; $<$ 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 the FEV1 (< 50% predicted), and the previous individual history of exacerbations (≥2 exacerbation/yr. or ≥1 hospitalization/yr.) [1]. Whenever there is a discrepancy between the risk category, as assessed by the FEV_1 and/or the exacerbation history, the variable indicating the highest risk should be used [1]. Of note, very recently (February 2013) the annually updated version of the GOLD document has proposed that the Clinical COPD Questionnaire (CCQ) questionnaire [8] can also be used alternatively to the CAT questionnaire to assess the symptomatic impact of the disease, and that one hospitalization due to an exacerbation of COPD suffices to classify the patient into the 'high risk' groups (C or D) [9]. Yet, these changes were not considered by the four papers herein reviewed because they were published [4-7] before the GOLD 2013 update release [9].

According to the results of this assessment, four patient groups (or categories) can be identified (**Figure 1**) [1]. It is important to note, however, that Groups A and B are

defined unequivocally according to risk ($FEV_1 \geq 50\%$ predicted AND 0-1 exacerbation/year), so any given patient must fulfil these two criteria to be classified in one of these two groups (**Figure 1**). This contrasts with the C and D groups, which admit several potential combinations. For instance, a given patient can be classified in the C group because the presence of less symptoms (by definition) AND EITHER $FEV_1 < 50\%$ predicted OR ≥ 2 exacerbations/year (or >1 hospitalization/yr. in the GOLD 2013 Update [9]) OR both (the same occurs with patients in group D) (**Figure 1**). In fact, some of the studies discussed below [4,7], but not the original GOLD document [1], have named these three subgroups as C_1, C_2 and C_3 or D_1, D_2 and D_3, respectively (see below).

This combined assessment proposal is expected to reflect the complexity of COPD better [1] than the uni-dimensional (FEV_1) analysis previously used for staging the disease [3] and is therefore a potentially important step forward towards a more personalized 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 (CT)-defined COPD phenotypes [11]. An initial group of 4,000 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 2,000 COPD cases and 2,000 smoking control subjects [11]. For the comparative analysis below, 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 SGRQ than patients diagnosed with COPD [12].

Han *et al.* used information from 4,484 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. Yet, the COPDgene study did not have CAT data, so the St George's Respiratory Questionnaire (≥ 25 vs. < 25) was used as a surrogate for the CAT threshold (≥ 10 vs. < 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 or D) with the frequency of exacerbation recorded prospectively during telephonic follow-up (1.4 years) [4].

Main results showed that: *(1)* 33.6% vs. 29.4% of patients were assigned to Group A when the mMRC or the SGRQ (as a surrogate of CAT) criteria were used, respectively (**Figure 2**). These percentages were 20.5% vs. 24.7% in Group B, 7.9% vs. 4.9% in Group C, and 38.0% vs. 41.0% in Group D (Kappa coefficient for agreement, 0.77) (**Figure 2**); and, *(2)* depending on the specific risk factor that determined category assignment in high risk groups (FEV_1 only (C_1, D_1), previous exacerbation history only (C_2, D_2), or both (C_3, D_3)) (**Figure 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 D_1, D_2 and D_3 subgroups were, 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.* 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' categorization was based on lung function, exacerbation history, or both [4].

Copenhagen Study

For this analysis, Lange *et al.* [5] pooled data from two similar but independent, general population studies: the fourth examination of the Copenhagen City Heart Study (CCHS) from a survey in 2001–2003, and the examination of the Copenhagen General Population Study (CGPS) from that in 2003–2010. The CCHS is a prospective epidemiologic 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], 6,237 of whom attended the fourth survey in 2001–2003 [14]. On the other hand, the CGPS is a prospective epidemiologic study, whose design is almost identical to the CCHS, which aims to recruit more than 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, a recruitment that began in 2003 and is still in progress [15]. Merging data from these two general cohorts, Lange *et al.* identified 6,628 COPD patients (**Table 1**) who were monitored for more than 4 years regarding 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 categorize symptoms) to predict the clinical course of COPD [5].

Main results showed that [5]: *(1)* the vast majority of patients in this study belonged to the A group (**Figure 2**), likely 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 3 years follow-up, mortality rates were 3.8%, 10.6%, 8.2% and 20.1% in Groups A, B, C and D, respectively (**Figure 4**). Of note, Groups B and D, both characterized by a higher degree of dyspnoea, had 5-8 times higher mortality from cardiovascular disease and cancer than Groups A and C.

From these results, Lange *et al.* concluded that [1]: *(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 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 [5]. An important aspect to consider in this study 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 (*COllaborative COhorts to assess Multicomponent Indices of COPD in Spain*) study is a pooled-analysis of individual patient-data (age, gender, mMRC scale, post-bronchodilator spirometry and all-cause mortality) from eleven COPD cohorts recruited in seven different cities in Spain (Galdakao, Pamplona, Requena, Sevilla, Tenerife, Terrassa and Zaragoza) for different purposes. In this pooled population (**Table 1**), Soriano *et al.* [6] sought to: *(1)* determine the distribution of the four groups proposed by the GOLD 2011 Revision; and, *(2)* compare its validity to predict all-cause mortality up to ten years, as compared to the previous GOLD staging system based mostly on the FEV_1 value alone. Other outcomes, namely exacerbations, were not reliably recorded and were not included in the analysis [6].

Main results showed that [6]: *(1)* of the 3,633 patients included in the analysis, 1,064 (33.6%) were classified in Group A, 515 (16.3%) in Group B, 561 (17.7%) in Group C and 1,023 (32.3%) in Group D (**Figure 2**). This distribution, however, varied significantly (χ^2 p value < 0.01) between the eleven cohorts pooled [6]; and, *(2)* the ability of GOLD 2007 to predict mortality at 1 year was not different from that of GOLD 2011 (0.635 vs. 0.639, p=0.53), 3 years (0.637 vs. 0.645, p=0.21) or 10 years (0.639 vs. 0.642, p=0.76). Again, however, survival varied greatly between cohorts [6].

From these results, Soriano *et al.* 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 FEV_1 only [6]. It should be remembered, however, that the multi-dimensional 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

The ECLIPSE (*Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points*) study 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 2,164 clinically stable COPD patients, 337 smokers with normal lung function and 245 never smokers, who were extensively characterized and followed up for 3 years [2]. Agusti *et al.* explored the distribution, characteristics, temporal stability and relationship with outcomes (exacerbations, hospitalizations and all-cause mortality) of 2,101 COPD patients with complete mMRC, spirometry and previous exacerbation history according to the 2011 GOLD assessment proposal (**Table 1**) [7].

Main results showed that [7]: *(1)* in addition to the expected differences between groups in the three variables that define them (mMRC, FEV_1 and previous exacerbations), the four groups differed also in many other clinical characteristics studied (**Table 2**). Hence, the amount of pulmonary emphysema and arterial 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, gender, FEV_1 reversibility or FEV_1 decline were not different between groups, although the latter was numerically higher in patients in the B category (**Table 2**) [7]; *(2)* an $FEV_1 < 50\%$ predicted 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 less prevalent one (subgroups C2 (13%) and D2 (9%)) (**Figure 3**); *(3)* Groups 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 [7]*;* and, finally, *(4)* the incidence of exacerbations during follow-up increased progressively ($p < 0.001$) from Groups A to B to C and to D [7]. By contrast, and similar to the findings of Lange *et al.* in the Copenhagen study [5], hospitalizations (**Figure 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 patients [7]. Of note 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.* 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, likely in relation to disease progression and/or response to therapy; *(3)* the new GOLD assessment proposal is feasible and valid to assess the risk of future exacerbations (which increase steadily from Groups A to B to C to D), but does not discriminate the risk of future hospitalizations and all-cause mortality for Groups B and C patients, likely highlighting the importance of symptom (and co-morbidity) assessment in clinical practice [7].

FREQUENTLY ASKED QUESTIONS (FAQs)

The comparative analysis of these four studies [4-7] helps to respond some FAQs regarding the GOLD assessment proposal, but it opens others that will inevitably require further research.

What is the prevalence of these four groups?

The answer depends on the population studied (**Figure 2**). In the general population the most prevalent group is A (77%) [5], whereas in patients recruited from secondary and tertiary care it is D (37%) followed closely by, surprisingly, A (32%) [4,6,7]; the prevalence of Group B is about 20% and, clearly, Group C is the less prevalent (10%), although it is also the one with the highest variability between studies (**Figure 2**).

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

Han *et al.* compared the mMRC and SGRQ scores (as a surrogate of CAT) in COPDgene and concluded that concordance between the two was good (kappa 0.77) but the distribution of the four groups (**Figure 2**) was significantly different [4]. 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 $\geq$1 was used as the cut-point instead of the threshold suggested by GOLD ($\geq$2) [1]. However, if a broader perspective is taken (**Figure 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 is a lot of data using the MRC, this instrument reflects only one aspect of COPD (i.e.,

breathlessness), whereas the CAT and the CCQ (now included in the GOLD 2013 revision [9]) are more comprehensive.

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

Three studies [4,5,7] 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) [7]. By contrast, other potentially relevant clinical variables, including age, gender or FEV_1 reversibility were not different across groups [7]. The precise nature of these comorbidities in ECLIPSE has been described in detail elsewhere [2] but includes the most prevalent ones such as cardiovascular disease, metabolic syndrome, osteoporosis and depression, among others [18].

Are there distinct subgroups in 'high risk' categories (C and D)?

Three studies [4,5,7] explored this question and came to very similar conclusions. The most frequent reason to categorize a given patient as a 'high risk' one (i.e. to classify him/her in Groups C and D) was an FEV_1 value < 50% of predicted (C_1 or D_1) (**Figure 3**). By contrast, that a patient was considered 'high risk' based only on the previous history of exacerbations (i.e., with an FEV_1 > 50%; C_2 or D_2) was relatively rare (**Figure 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 [7]. Results showed that the relative majority of patients assigned to Groups A and D continue to belong to these same two groups three years later, whereas those classified in Groups B or C tend to vary more. The mechanisms explaining these temporal changes (or lack of changes) also deserve future research, but they are likely related to either disease progression or response to therapy. Importantly, these findings show that there is also heterogeneity in the so-called "disease progression" which might open opportunities for therapeutic intervention [7].

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.* in the general population recruited in Copenhagen [5] as well as Agusti *et al.* in the ECLIPSE cohort [7] found that survival at 3 years was better for patients Group A, worse for Group D and intermediate and similar for Groups B and C (**Figure 3**). This is consistent with previous studies on the prognostic importance of symptom perception [19]. In Cocomics, survival at three years was also similar in Groups B and C, but results changed after 10 years follow-up (**Figure 4**) [6], indicating perhaps a survival bias.

With respect to the precise causes of death in these four patient categories, Lange *et al.* showed that Groups B and D (both 'High Symptoms') had 5-8 times higher mortality from cardiovascular disease and lung cancer than the 'Less Symptoms' groups (A and C) [5]. Specific causes of death were not investigated in COPDgene [4] or Cocomics [6], 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 (BMI), increased number of comorbidities, and lower exercise capacity) [7]. This information may help clinicians to assess the risk of death in individual patients.

Do these four groups of patients predict exacerbations/hospitalizations?

All four studies showed that the incidence of exacerbations increases progressively from A to B to C to D [4-7]. This is not surprising since 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 hospitalizations was explored in ECLIPSE only [7] and, alike mortality (see above), hospitalizations during follow-up were scanty in Group A, frequent in D and intermediate (and close!) in Groups B and C patients [7]. Further, identified risk factors for hospitalizations were similar to those identified for mortality (see above) [7].

Why do Groups B and C appear to have similar risk of death/hospitalization?

Both Lange *et al.* [5] and Agusti *et al.* [7] suggested that 'high symptoms' in Group B may originate from comorbidities and not from airflow limitation. In fact, Lange *et al.* showed in the general population that Group B patients frequently die of cardiovascular disease or lung cancer [5]. Likewise, in the ECLIPSE cohort, Group B patients were also characterized by the highest prevalence of comorbidities and presence of persistent systemic inflammation [7]. 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 COPD patients may need hospitalization (or die)

during follow-up for reasons not directly related to their lung function (for instance, comorbidities) and supports 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 [7] (**Table 2**). 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 [7].

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

Comorbidities appear to be more prevalent in B patients [5,7] and, as discussed above, they may contribute to the increased risk of hospitalizations and death in this group (**Figure 4**). Of note, however, in the four studies reviewed here [4-7] comorbidities were either identified from medical records or self-reported, but not actively identified by clinical investigations [23]. This may be clinically relevant since most co-morbidities are undiagnosed. Despite that the new GOLD recommendations [1] clearly highlighted in the text the important role of comorbidities in COPD [21,23,24], a pending challenge is how to incorporate them explicitly in the proposed assessment diagram. Ongoing discussion is addressing the importance of co-morbidities for the definition of COPD severity and how to incorporate them in the A,B,C,D diagram or in 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 [4-7] 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 (RCT's). Hopefully, this much needed evidence will become available in the near future either by retrospective analysis of already published RCT's or, ideally, by new prospective ones.

Strengths and limitations of this analysis

As any analysis, the current one has strengths and shortcomings. Among the former, this is the first paper comparing four recent studies [4-7] 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 utilizes existing cohorts of COPD patients that were established for different purposes, so results will have to be replicated prospectively in other populations. And, secondly, all of them are observational cohorts, so that in none of them, treatment was controlled and was dictated by the local physician taking care of the patient [4-7]. 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 [4-7] that investigated the clinical implications of the new GOLD recommendations [1]. From this analysis, a number of FAQs could be addressed but other new questions, that may require further research, arose. Altogether, though, the results discussed herein constitute a repository of potentially useful information for practitioners, investigators, administrators, regulators, stakeholders and guideline writers. Yet, 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. Agusti A, Calverley P, Celli B et al. Characterisation of COPD heterogeneity in the ECLIPSE cohort. *Respiratory Research* 2010; 11: 122-136.

3. 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.

4. Han MK, Mullerova H, Curran-Everett D et al. GOLD 2011 disease severity classification in COPDGene: a prospective cohort study. *The Lancet Respiratory Medicine* 2012; 1: 1-8.

5. Lange P, Marott JL, Vestbo J+ et al. Prediction of the Clinical Course of Chronic Obstructive Pulmonary Disease, Using the New GOLD Classification. *American Journal of Respiratory and Critical Care Medicine* 2012; 186: 975-981.

6. Soriano JB, Alfajame I, Almagro P et al. Distribution and prognostic validity of the new GOLD grading classification. *Chest* 2012;

7. 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;

8. Reda AA, Kotz D, Kocks JW, Wesseling G, Van Schayck CP. Reliability and validity of the clinical COPD questionniare and chronic respiratory questionnaire. *Respir Med* 2010; 104: 1675-1682.

9. Vestbo J, Hurd SS, Agusti AG et al. Global Initiative for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease. Updated 2013. *www. goldcopd. org* 2013;

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. *American Journal of Respiratory and Critical Care Medicine* 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, Jensen G, Nyboe J. 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, Vestbo J, Jensen JS. Cardiovascular morbidity in COPD: A study of the general population. *COPD* 2010; 7: 5-10.

15. Stender S, Frikke-Schmidt R, Nordestgaard BG, Tybjaerg-Hansen A. 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, Banik N. Comparisons of health status scores with MRC grades in a primary care COPD population: implications for the new GOLD 2011 classification. *Eur Respir J* 2012;

18. Fabbri LM, Luppi F, Beghe B, Rabe KF. Complex chronic comorbidities of COPD. *Eur Respir J* 2008; 31: 204-212.

19. Nishimura K, Izumi T, Tsukino M, Oga T. 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. *New England Journal of Medicine* 2010; 363: 1128-1138.

21. Mannino DM, Thorn D, Swensen A, Holguin F. 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;

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. *The Lancet Respiratory* 2013; 1: 73-83.

Table 1. Main demographic, clinical and functional variables of the COPD patients included in each of the four studies analysed here [4-7]. For further explanations, see text.

	COPDGene	**Copenhagen CS**	**Cocomics**	**ECLIPSE**
N	4,484	6,628	3,163	2,101
Recruitment source	Hospitals	General population	7 cohorts	Hospitals
Country	USA	Denmark	Spain	Europe/USA
Age, yrs.	65	68	66	64
Male, %	53	52	93	65
BMI, Kg/m^2	27	25	28	27
FEV_1, % predicted	40	40	54	50
Follow-up, yrs.	1.7	4.3	10	3

Table 2. Heat map representation of the main clinical characteristics and associated outcomes of the four GOLD 2011 categories in the ECLIPSE cohort [7]. White rows indicate lack of statistically significant differences between them, whereas a colour code is used to highlight differences in severity (green (low), orange (moderate) or red (high)). For further explanations, see text.

	GOLD 2011 groups				
Distribution	**Group A**	**Group B**	**Group C**	**Group D**	**p value**
GOLD 2011 assessment variables					
mMRC dyspnoea score	0.6	2.3	0.8	2.6	<0.001
FEV_1 % predicted	64.2	61.6	42.2	37.2	<0.001
Exacerbations previous 12 months	0.2	0.3	1.0	1.3	<0.001
Demographics					
Age, years	63.2	64.1	62.9	63.6	NS
Female, %	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 smoker, %	38%	34%	43%	32%	0.001
Symptoms & Comorbidities					
Chronic bronchitis, %	26%	35%	38%	38%	<0.001
SGRQ-C Total score	31.9	55.3	44.6	62.2	<0.001
Number of comorbidities	2.0	2.8	1.7	2.3	<0.001
Physiology & Imaging					
FEV_1 % reversibility	11.4%	11.8%	10.3%	10.2%	NS
SaO2, %	95.7%	95.1%	94.4%	93.7%	<0.001
6MWD, m	440	360	408	307	<0.001
Emphysema CT, %LAA -950HU	11.8%	12.5%	18.9%	22.5%	<0.001
Outcomes during follow-up (3 yrs.)					
ECOPD, per year	0.6	0.9	1.3	1.7	<0.001
≥1 Hospitalization, %	11%	25%	30%	46%	<0.001
Mortality rate, %	4%	10%	8%	14%	<0.001
Rate of annual FEV_1 decline, ml/yr.	-33.4	-38.0	-30.2	-31.9	NS

FIGURE LEGENDS

Figure 1. Four patient categories (or groups) resulting from the application of the 2011 GOLD assessment proposal (modified from Reference [1]). For further details, see text. Although this is the classification used for the four studies reviewed here [4-7] , it is of note, the 2013 update [9] proposes to add (not included in the figure) the CCQ 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 hospitalization because of COPD during the past year (also not included in the figure) (www.goldcopd.org).

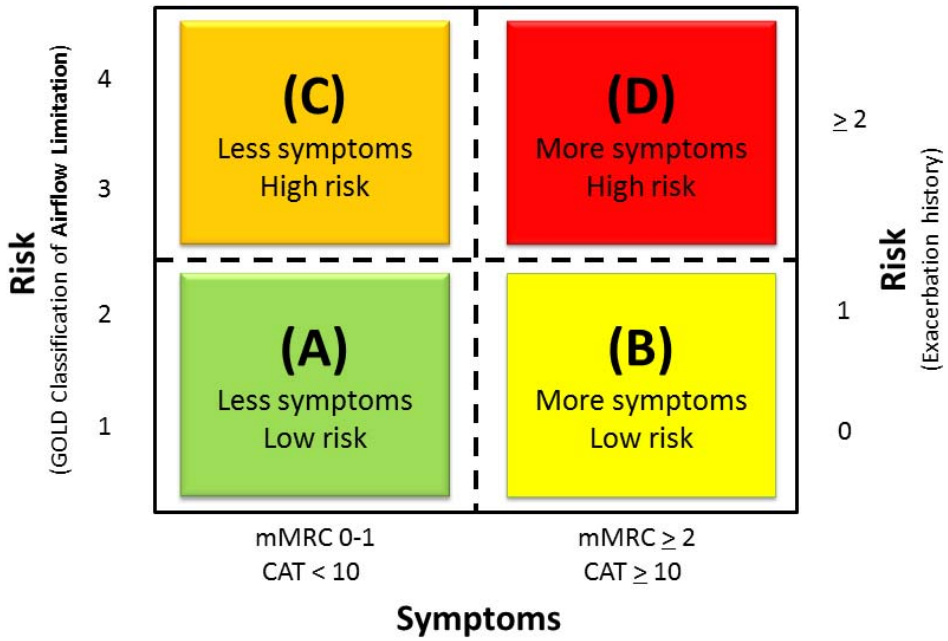

Figure 1

Figure 2. Frequency distribution of the four patient groups observed in the four studies compared here [4-7]. Grey columns indicate the mean value (excluding data from the

Copenhagen study, the only one that recruited patients from the general population [5]). For further explanations, see text.

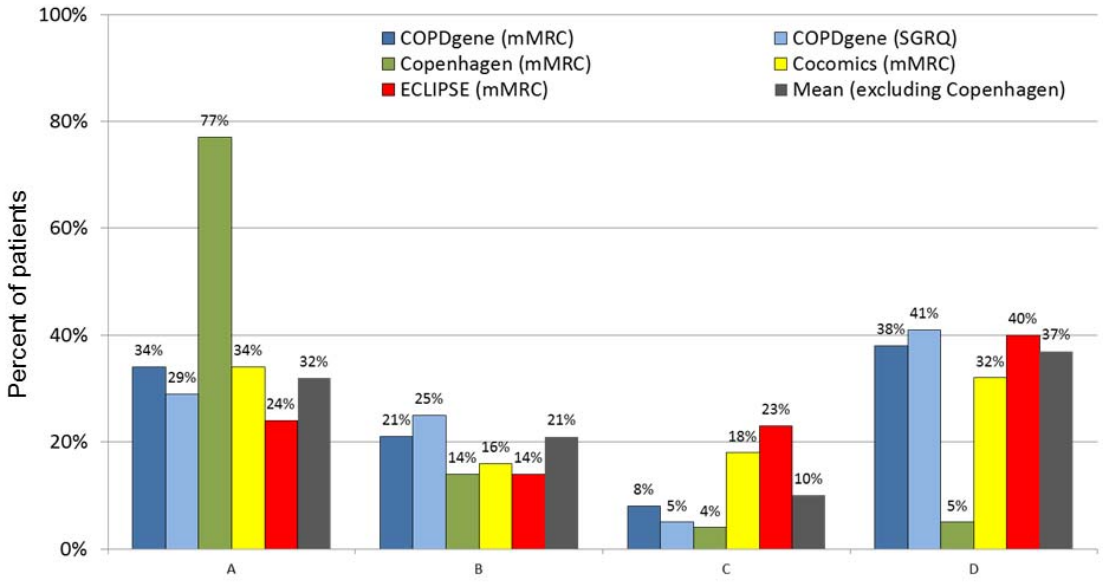

Figure 2

Figure3. Prevalence of C and D subtypes in three out of the four studies compared here [4,5,7]. In all of them, the main reason to be categorized as a high risk patient (C or D groups) is an FEV1 < 50% of the reference value alone. For further explanations, see text.

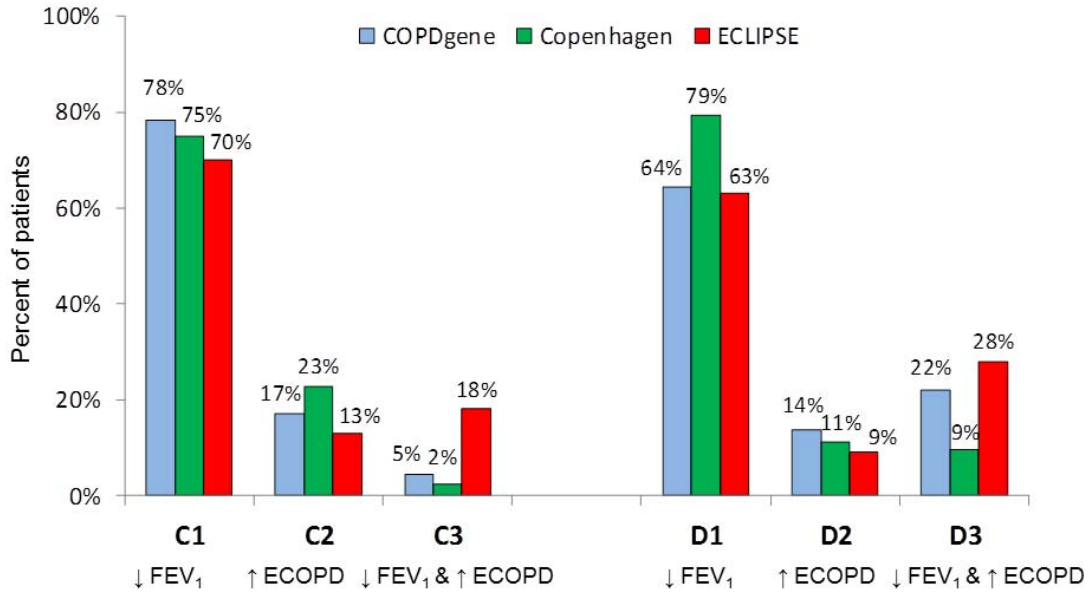

Figure 3

Figure 4. Kaplan–Meier survival curves (all-cause mortality) in three of the studies compared here [5-7]. For further explanations, see text.

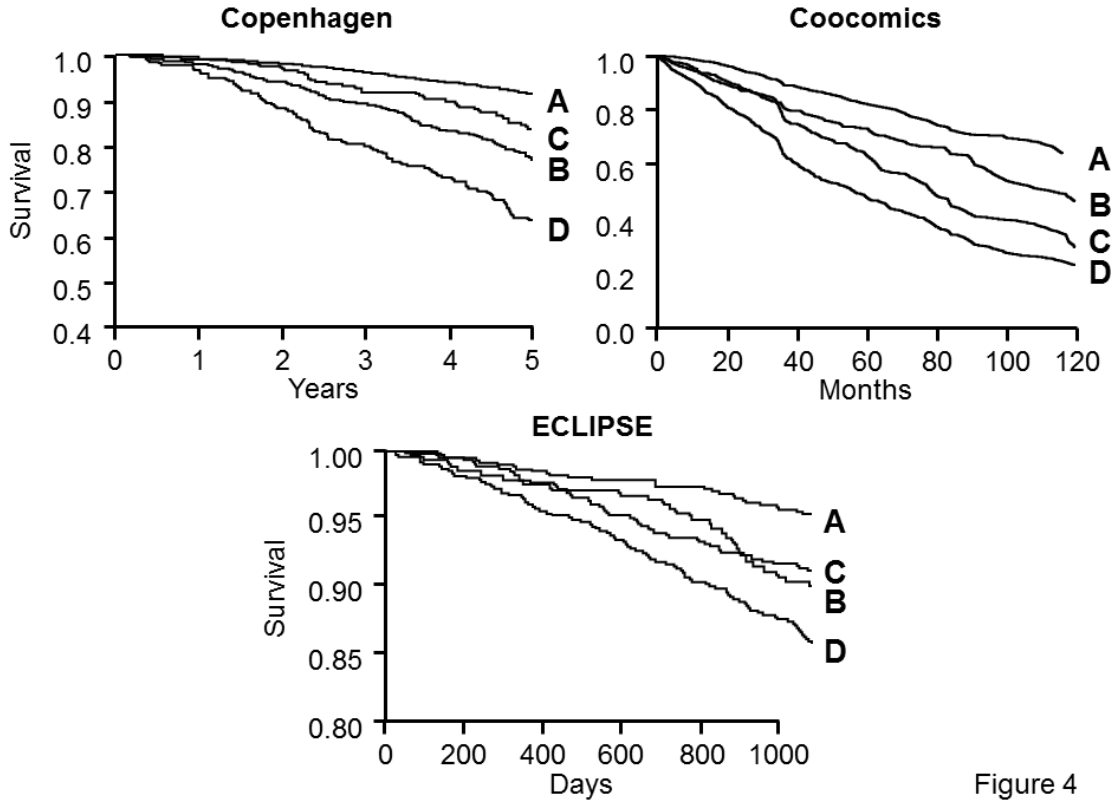

Figure 5. Temporal stability of the four patient categories in the ECLIPSE study (reproduced with permission from reference [7]). For further explanations, see text.

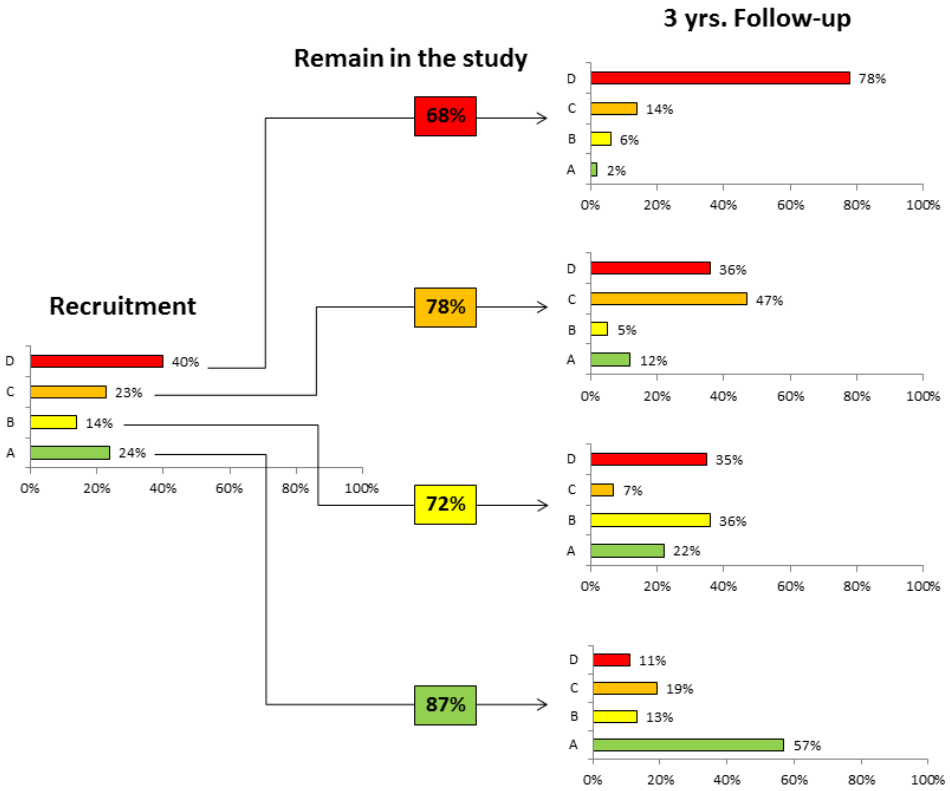
Recruitment
D 40%
C 23%
B 14%
A 24%
0% 20% 40% 60% 80% 100%
Remain in the study
68%
78%
72%
87%
3 yrs. Follow-up
D 78%
C 14%
B 6%
A 2%
D 36%
C 47%
B 5%
A 12%
D 35%
C 7%
B 36%
A 22%
D 11%
C 19%
B 13%
A 57%
0% 20% 40% 60% 80% 100%

Figure 5