The distribution of COPD in UK general practice using the new GOLD classification

John Haughney¹, Kevin Gruffydd-Jones², June Roberts³, Amanda J Lee⁴, Alison Hardwell⁵, Lorcan McGarvey⁶

¹Centre of Academic Primary Care, University of Aberdeen, Aberdeen, UK
²Box surgery, Wiltshire, UK
³Salford Royal NHS Foundation Trust, Salford, UK
⁴Medical Statistics Team, University of Aberdeen, UK
⁵National Services for Health Improvement, Dartford, UK
⁶Centre of Infection and Immunity, The Queen’s University of Belfast, UK

Correspondence: John Haughney; Centre of Academic Primary Care, University of Aberdeen, Aberdeen, UK. Email: j.haughney@abdn.ac.uk

Word count: 2982
ABSTRACT
The new Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2011 document recommends a combined assessment of COPD based on current symptoms and future risk. A large database of primary care COPD patients across the UK was used to determine COPD distribution and characteristics according to the new GOLD classification. Eighty general practices provided patients with a Read Code diagnosis of COPD. Electronic and hand searches of patient medical records were undertaken, optimising data capture.
Data for 9219 COPD patients were collected. For the 6283 patients with both FEV$_1$ and mMRC score (mean (SD) age 69.2 (10.6) years, BMI 27.3 (6.2) kg/m2), GOLD 2011 group distributions were: A (low risk, less symptoms): 36.1%; B (low risk, more symptoms): 19.1%; C (high risk, less symptoms): 19.6%; D (high risk, more symptoms): 25.3%, contrasting with GOLD 2007 classification: I (mild): 17.1%; II (moderate): 52.2%; III (severe): 25.5%; IV (very severe): 5.2%. Of patients with FEV$_1$ % predicted $\geq$ 50%, 20% had $\geq$2 exacerbations in previous 12 months; of patients with FEV$_1$ < 50% predicted, 70% had <2 exacerbations in previous 12 months.
This database, representative of UK primary care COPD patients, identified greater proportions of patients in the mildest and most severe categories comparing 2011 versus 2007 GOLD classifications. Discordance between airflow limitation severity and exacerbation risk was observed.
INTRODUCTION
Chronic obstructive pulmonary disease (COPD) is a multi-component disease that shows marked heterogeneity in terms of clinical outcomes, disease severity and progression. Traditionally the severity of COPD has been assessed by the degree of airflow obstruction, using a staging system based on forced expiratory volume in one second (FEV\textsubscript{1}) [1]. It is now recognised that FEV\textsubscript{1} alone fails to represent the complexity of COPD and is poorly related to important patient factors including breathlessness, health status, level of comorbidities and exacerbation risk [2,3]. This deficiency prompted the UK National Institute for Clinical Excellence (NICE) to recommend the multidimensional assessment of COPD severity as a key area for implementation (NICE 2010) [4]. Most recently, the Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2011 strategy document proposed a combined assessment of COPD using the modified Medical Research Council (mMRC) dyspnoea grade, current health status assessed by the COPD Assessment Test (CAT), and future risk based on either severity of airflow limitation or exacerbation history [5]. The new assessment categorises patients into one of four groups: A: low risk, less symptoms; B: low risk, more symptoms; C: high risk, less symptoms; D: high risk, more symptoms.

Only a few studies have reported the utility of the new GOLD 2011 classification in COPD populations. In an analysis of COPD patients (n=6628) drawn from two large general population surveys, Lange et al found the new classification to have improved prognostic capacity to predict future exacerbation risk compared with the older version [6]. They also reported that patients with more symptoms (dyspnoea) but relatively well preserved lung function (subgroup B) had poorer survival than those with more severe obstruction but fewer symptoms (subgroup C) highlighting the existence of a number of distinctly different disease phenotypes with regard to future patient risk. Using a dataset of 1041 COPD patients (approximately one third from primary care, two thirds from specialist care), Nadeau and colleagues reported that a fifth of patients considered ‘low risk’ using the new GOLD classification, had actually suffered an exacerbation in the previous year [7]. As a history of a COPD exacerbation in the previous year is now recognised as a strong predictor of future exacerbation risk [8], this highlights potential limitations of the new classification system. Additional considerations that need addressing include identifying the optimal cut-points on the mMRC and CAT tools for symptomatic sub-grouping of patients. Using data from European primary care COPD patients (n=1810), Jones et al suggested a mMRC cut point of ≥ 1 rather than ≥ 2 had the closest equivalence to a CAT cut-point >10 [9].
Therefore, although the new GOLD classification may need future refinement it is clear that multidimensional assessment of COPD patients has important clinical implications. This is particularly the case in primary care where the majority of COPD patients are managed clinically. To determine the true distribution of COPD in primary care according to the newly recommended GOLD classification requires analysis of large and representative primary care databases. However, they may be limited by incomplete data on important variables including lung function, accurate recording of exacerbations, comorbidities and details of contacts with external agencies (hospital admissions, out-patient visits or out-of-hours consultations).

In attempt to overcome some of these limitations we have assembled a large and representative database of COPD patients attending primary care throughout the UK. Here, we utilise these data to report the distribution of COPD severity across the UK using the new GOLD 2011 assessment framework and its comparison with the previous GOLD 2007 classification.

METHODS

Study design and study population
This was an observational, multi-centre, retrospective, cohort study which aimed to evaluate prevalence, incidence, severity, co-morbidities and burden of disease in patients with COPD. Patients were identified by the presence of a Read Code diagnosis (the standard UK diagnosis classification system) for COPD [10]. Patients with a diagnosis of COPD for at least 1 year were eligible for inclusion; where available, data for the past 3 years were collected.

The database used was generated by the National Service for Health Improvement (NSHI) and sponsored by GlaxoSmithKline. General practices across the UK were invited to participate to ensure a broad representation of practices and COPD patients. NSHI nurse specialists carried out electronic searches of electronic patient records (MIQUEST search) [11] on identified COPD patients; supplemented by a hand search of all patient records held by each practice (see online supplement for further details of sampling procedure and database generation).

Consent was obtained from each general practice. Ethics approval was not required for this study as there was no patient contact and no patient-identifiable material was recorded; all data collected on the database were anonymised (ClinicalTrials.gov Identifier: NCT012633340).

Study outcomes
**Demographic characteristics**

Data collection items included: sex, age, height, weight, body mass index (BMI), duration of COPD, smoking history, concurrent therapies for COPD, and co-morbidities. Co-morbidities of common interest were: cardiovascular disease, cerebrovascular disease, cancer, diabetes mellitus, depression and osteoporosis; defined as an ever-recorded diagnosis.

**FEV₁**

The most recently recorded spirometry readings available to the practice were documented. The severity of airflow limitation categories was defined according to GOLD 2007 criteria: I (mild): FEV₁ ≥ 80% predicted; II (moderate): FEV₁ 50% to 79% predicted; III (severe): FEV₁ 30% to 49% predicted; IV (very severe): FEV₁ <30% predicted.¹

**COPD Exacerbations**

Exacerbations were identified by researcher scrutiny of physicians’ records (primary, secondary or emergency care) and this evidence was cross-checked by identification of one or more of the following:

- READ code in the clinical records
  or
- Prescription of oral corticosteroids in primary care for an intended duration of at least three days, where no other reason for the prescription other than the presence of COPD was identified
  or
- Attendance at an out of hours or emergency centre or A&E department with a primary diagnosis of COPD resulting in a prescription of oral corticosteroid for an intended duration of at least three days
  or
- A hospital admission with a primary diagnosis of COPD.

**Symptom assessment**

Where available, the most recently recorded score for mMRC was documented. The mMRC dyspnoea scale describes five grades of breathlessness ranging from 0 (least severe breathlessness) to 4 (most severe breathlessness) [12]. It was noted if a CAT score was available, although its magnitude was not recorded. The CAT is an 8-item health status questionnaire with total score ranging from 0 (best) to 40 (worst) [13].

**Statistical analysis**
Data were entered into SPSS (version 21) for analyses. Summary statistics included the mean and standard deviation (SD) or median and interquartile range (IQR) for continuous data and number (percentage) for categorical data. Demographic and clinical factors were compared across GOLD 2007/2011 groups using the chi-square test for categorical data and analysis of variance (ANOVA) for continuous data. A p-value of ≤0.05 was used to denote statistical significance throughout.

RESULTS

General practice sample characteristics
In total, 80 practices across the UK contributed to this study and represented most geographical areas. Thirty two (40%) of the practices were training practices; the majority of practices had 6 general practitioners (GPs) or fewer (90%) and 24 (30%) practices had ≤2 GPs. More details of the sample characteristics can be found in the online supplement.

Study population
The total patient population of all practices was 540793; data on 9219 patients with a Read Code diagnosis of COPD were collected between November 2010 and January 2012 (Figure 1), giving an overall prevalence of COPD of 1.7%. Across the 80 practices, COPD prevalence varied from 0.3% to 4.3% with a median (IQR) practice prevalence of 1.8% (1.2%-2.3%).

The mean (SD) age of the group was 69.5 (11.1) years; approximately half were male; most were either former (53%) or current (37%) smokers and mean (SD) BMI for the group was 27.2 (6.3) kg/m² (Table 1). According to GOLD spirometry-based severity criteria, patients were most commonly classified as having moderate COPD (52%) with very few patients being classified as having very severe COPD (5%). The mean annual rate of exacerbations in the previous 12 months was 0.89; just under a half of the patients had experienced at least one exacerbation in the previous year and approximately one fifth had had ≥2 exacerbations.

A score for mMRC was recorded for 7119 (77%) patients; the most common score was ‘1’ (breathless when hurrying on level or up a slight hill) (37%) and ‘2’ (walk slower than people of same age on the level due to breathlessness or stop for breath when walking on level at own pace) (27%). A total of 278 (3%) patients in 10 practices had a CAT score recorded. Due to the small number available, no further analysis was undertaken.

Co-morbidities
Approximately 79% (n=7267) of patients had one or more comorbidities.
The most commonly reported co-morbidities, defined as an ever-recorded diagnosis, were cardiovascular disease (46%) followed by asthma (23%) and cerebrovascular disease (17%). Depression, cancer and diabetes were reported by 14-15% of the population.

**Medication**

The proportion of patients receiving any COPD medication was 87.4%; the most commonly prescribed maintenance therapies were combined treatment with an inhaled corticosteroid (ICS) plus long-acting beta-agonist (LABA) (56%) and the long-acting anticholinergic tiotropium (46%).

**Demographic and clinical characteristics by the 2007 and 2011 GOLD categories**

Patient characteristics according to GOLD 2011 and 2007 classifications are shown in Table 2.

By GOLD 2007 classification, patients in the sub-groups with greatest airflow imitation (III and IV) were older, more likely to be male and have lower BMI compared with patients with GOLD stage I and II. Patients in GOLD stage II and III had higher levels of cardiovascular co-morbidities compared with stages I and IV. The proportions of patients with lung cancer increased with increasing severity of airflow limitation.

When classified by GOLD 2011, patients in the higher symptom sub-groups (B and D) were older and had higher levels of comorbidities compared with the low symptom sub-groups (A and C). Patients in group B (low risk with preserved lung function but more symptoms) were more likely to be female, have the highest BMI and the greatest proportion with cardiovascular co-morbidities. The proportions of patients with \( \geq 1 \) COPD exacerbation by sub-groups A to D were: A: 65.0%; B: 71.5%; C: 85.6%; D: 88.5%; and those with \( \geq 1 \) hospitalisation were: A: 4.0%; B: 8.3%; C: 13.2%; D: 26.4%. The proportions of patients who were prescribed an inhaled steroid preparation (either a monotherapy or as combination therapy) were: A: 55.8%; B: 70.9%; C: 76.5%; D: 89.6%.

**Distribution of patients according to new GOLD categories**

The distribution of patients into the new, GOLD 2011 categories, using patients with both a valid FEV\(_1\) percent predicted and a valid mMRC score (n=6283) is presented in Figure 2a. The largest proportion of patients (36%) were classified into sub-group A (low risk, less symptoms,); approximately a fifth of patients were classified into sub-groups B (low risk, more symptoms, 19.1%) and C (high risk, less symptoms, 19.6%) and approximately a quarter (25.3%) were classified into sub-group D (high risk, more symptoms).
Therefore, this new GOLD classification categorised 46% of the population as ‘high risk’ whereas only 30.5% of the population would be considered high risk (FEV\textsubscript{1} < 50%; GOLD stage III and IV) according to previous GOLD 2007 system. The distribution of patients in primary care using the new GOLD 2011 classification suggests a shift towards two distinct sub-groups; A (low risk, less symptoms) and D (high risk, more symptoms). Such a pattern of distribution is not so evident using previous GOLD 2007 classification (Figure 3). The GOLD 2013 recommends that one or more hospitalisations for COPD exacerbations should be considered high risk. The results of applying GOLD 2013 to our population is shown in Fig 2b.

Comparison of risk stratification by exacerbation frequency and by airflow limitation

For the 7480 patients with a valid FEV\textsubscript{1} percent predicted, the proportion of patients categorised as high risk differed according to risk assessment used (Figure 4). Of the 5198 patients who had low risk airflow limitation, 1029 (20%) had had ≥ 2 exacerbations in the previous 12 months. Of the patients with high risk on airflow limitation (n=2282), 1607 (70%) had a low exacerbation risk profile (0-1). Similarly for patients with a high exacerbation risk (n=1704), 1029 (60%) had FEV\textsubscript{1} ≥ 50% of predicted.

DISCUSSION

Our study is the first to describe the distribution of COPD patients in primary care throughout the UK, according to the new GOLD 2011 classification. Our key findings are that a greater proportion of COPD patients were identified as at high risk of adverse health outcome than would otherwise have been determined by the previous GOLD classification; the sub-groups with higher symptom levels (B and D) were associated with higher levels of co-morbidities compared with the low symptom level sub-groups A and C; there was a large heterogeneity in patient risk whether determined by FEV\textsubscript{1} % predicted or exacerbation history.

In our study we have shown that, compared with the GOLD 2007 classification, GOLD 2011 categorises more patients in both the mildest (group A) and most severe groups (group D). Similarly, Lange \textit{et al} reported a higher proportion of patients in the most severe group when comparing the GOLD 2011 (4.5%) and GOLD 2007 (0.7%) classifications [6]. In contrast, a higher proportion of subjects (77%) were categorised into group A compared with our own findings (36%) reflecting a major difference in the study populations; Lange’s study was a general population study comprising a large proportion of treatment naive patients.
contrasting our own population of already-diagnosed, largely treated, COPD patients. The
distribution of our patients into groups A to D was comparable to the findings of Nadeau et
al, although their population was skewed towards specialist care rather than primary care [7].
Using the GOLD 2011 classification on the Health-Related Quality of Life in COPD in
Europe Study (HEED) database Jones et al reported greater proportions of patients in the
higher risk sub-groups C and D compared to ours, likely reflecting distinct clinical
differences in study population [9]; the HEED database captured patients presenting to their
primary care physician for a scheduled visit, with 13% presenting with a COPD exacerbation
[14].

Our dataset provides further evidence of the difficulty in classifying disease risk in COPD;
one fifth of patients had experienced ≥ 2 exacerbations in the previous year despite having
‘low risk’ airflow limitation (FEV₁ >50% predicted), and the majority of patients (70%) with
‘high risk’ airflow limitation (FEV₁< 50% predicted), had 0 or 1 exacerbations in the previous
year. Although some studies suggest that there is an increased risk of exacerbations with
increased levels of airflow limitation [5,15], these data endorse the evidence that lung
function alone does not predict the likelihood of having an exacerbation. Our findings agree
with the ECLIPSE study which showed a heterogeneous response in the rate of exacerbations
across all GOLD stages [3]. Lange et al also reported heterogeneity in GOLD groups C and D
with respect to risk of future exacerbations, and perhaps surprisingly showed that patients in
group B, characterised by more severe dyspnoea, had worse survival outcomes than patients
in group C with greater airflow limitation [6].

Lange and colleagues speculated that the poorer survival outcomes may have been related to
the higher incidence of cardiovascular comorbidities in group B compared with group C [6].
The characteristics of patients in the GOLD 2011 sub-groups showed marked similarities in
our study and the Lange study. In both, the patients in sub-groups B and D were older and
had higher levels of cardiovascular co-morbidities. In our study, patients in sub-groups B and
D had higher levels of cerebrovascular disease, depression, cancer and diabetes compared
with patients in sub-groups A and C. We showed no clear pattern between GOLD 2007
categories and level of co-morbidities which was consistent with findings in the ECLIPSE
study [3]. Both our study and the Lange study suggest that sub-group B patients warrant close
attention; their symptoms possibly being driven by their associated co-morbidities, in
particular cardiovascular disease.
We report a prevalence of COPD of 1.7% which compares well with the reported mean prevalence of 1.7% for the UK in 2012, from a UK Quality and Outcomes Framework NHS database [16]. Our study did not evaluate specific age groups and may explain the lower reported prevalence. The differences may also be related to methods of reporting i.e. prevalences based on diagnosed and treated COPD compared with that based on spirometry measurements alone. COPD prevalence figures do, of course, vary considerably depending on the age range of the population included. For example, UK figures report a prevalence of 4.1% in adults age $\geq 30$ [17] and 13.3% in adults aged $> 35$ [18]. A review of prevalence data across Europe reported prevalences from 2.1% to 26.1% depending on country, age group and methods of assessment used [19].

For this study we sought to ensure the dataset generated was representative of primary care throughout the UK. Specifically we recruited general practices across a broad geographical distribution and ensured both the variation in practice size and the proportion with a primary care training designation was reflective of general practices in the UK. A particular strength of this database is its size and due to the comprehensive searches conducted, there were very few missing data items for demographic characteristics and relatively high numbers of records for clinical outcomes; providing a reliable source of data for exacerbations, FEV$_1$, mMRC and co-morbidities. Although the quality of FEV$_1$ recordings could not determined, the data represent real-world data capture. We acknowledge that findings from this study cannot be extrapolated to other European countries due to differences in healthcare systems including provision of primary care. Another limitation is the limited CAT data, collected in only three percent of patients across a few practices. However, these data are representative of primary care practice and reflect the fact that this health status measure is not yet routinely collected in primary care. In addition, the new GOLD classification allows the assessment of symptoms based on mMRC or CAT scores, and the optimal symptom questionnaire and associated cut-points are still the subject of some debate [9].

The new GOLD strategy combined assessment brings a welcome and marked change in the approach to managing COPD patients, which considers both the disease impact (current symptoms and activity limitation) and the future risk of disease progression (especially exacerbations). This may have important implications for patients in primary care in terms of treatment and their level of primary care contact. For clinicians, resources can be targeted to
patients with the greatest needs in terms of pharmacological therapy, treatment of comorbidities and frequency of monitoring required. For patients this should result in improved and targeted care which should maximise the treatment of symptoms and minimise the risk of exacerbations.

In conclusion, this study successfully used a comprehensive database, representative of COPD patients across the UK, and showed that classifying patients using the new GOLD 2011 criteria identified larger proportions of patients in the mildest and more severe groups compared with the GOLD 2007 classification. We also identified discordance between the level of airflow limitation and exacerbation risk highlighting the importance of recognising distinct COPD phenotypes. The GOLD 2011 classification should help primary care physicians target patients for pharmacological therapy and identify those in need of more frequent monitoring to reduce future long-term risks.

Acknowledgements
This study was undertaken by the National Service for Health Improvement (NSHI) using an unrestricted educational grant, funded by GlaxoSmithKline. The authors would like to thank the participating GP practices for their contribution to this study. The authors would also like to acknowledge editorial support in the form of assistance in the development of the manuscript drafts, including assembling tables and figures, collating author comments and copyediting, which was provided by Kate Hollingworth of Continuous Improvement Ltd. This support was funded by NSHI.
References


http://www.connectingforhealth.nhs.uk/systemsandservices/data/uktc/readcodes
http://www.connectingforhealth.nhs.uk/systemsandservices/data/miquest
Medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with 
Cataluna JJ, van der Molen T, Adamek L, Banik N.. Health-related quality of life in patients 
15. Hoogendoorn M, Feenstra TL, Hoogenveen RT, Al M, Rutten-van Mölken M. 
Association between lung function and exacerbation frequency in patients with COPD. Int J 
Chron Obstruct Pulmon Dis 2010;5 435–444.
http://www.gpcontract.co.uk/browse/UK/12
17. Frank TL, Hazell ML, Linehan MF, Morris JA, Frank PI. The estimated prevalence of 
chronic obstructive pulmonary disease in a general practice population. Prim Care Resp J 
dependence of chronic obstructive pulmonary disease in a nationally representative 
19. Atsou K, Chouaid C, Hejblum G. Variability of the chronic obstructive pulmonary 
Table 1: Demographic and baseline characteristics of COPD cohort

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total COPD cohort (n=9219)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years, mean (SD)</td>
<td>69.5 (11.1)</td>
</tr>
<tr>
<td>Males, n (%)</td>
<td>4693 (50.9)</td>
</tr>
<tr>
<td>Current smokers, n (%)</td>
<td>3399 (37.1)</td>
</tr>
<tr>
<td>Body mass index, kg/m², mean (SD)</td>
<td>27.2 (6.3)</td>
</tr>
<tr>
<td>FEV₁, % of predicted, mean (SD)</td>
<td>60.6 (19.9)</td>
</tr>
<tr>
<td>Severity of airflow limitation by percent predicted FEV₁ classification, n (%)</td>
<td></td>
</tr>
<tr>
<td>Mild (≥ 80%)</td>
<td>1307 (17.5)</td>
</tr>
<tr>
<td>Moderate (≥ 50 % and &lt; 80%)</td>
<td>3891 (52.0)</td>
</tr>
<tr>
<td>Severe (≥ 30% and &lt; 50%)</td>
<td>1891 (25.3)</td>
</tr>
<tr>
<td>Very severe (&lt; 30%)</td>
<td>391 (5.2)</td>
</tr>
<tr>
<td>Number of COPD exacerbations in last 12 months, n (%)</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>5140 (55.8)</td>
</tr>
<tr>
<td>1</td>
<td>2017 (21.9)</td>
</tr>
<tr>
<td>≥ 2</td>
<td>2062 (22.4)</td>
</tr>
<tr>
<td>Mean rate of exacerbations in last 12 months, per patient/per year</td>
<td>0.89</td>
</tr>
<tr>
<td>Modified Medical Research Council (mMRC) dyspnoea score, n (%)¹</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>7119</td>
</tr>
<tr>
<td>0</td>
<td>1184 (16.6)</td>
</tr>
<tr>
<td>1</td>
<td>2633 (37.0)</td>
</tr>
<tr>
<td>2</td>
<td>1952 (27.4)</td>
</tr>
<tr>
<td>3</td>
<td>1066 (15.0)</td>
</tr>
<tr>
<td>4</td>
<td>284 (4.0)</td>
</tr>
<tr>
<td>Co-morbidities, n (%)</td>
<td>7267 (78.8)</td>
</tr>
<tr>
<td>Any co-morbidity</td>
<td>4225 (46.1)</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>2141 (23.4)</td>
</tr>
<tr>
<td>Asthma</td>
<td>1504 (16.6)</td>
</tr>
<tr>
<td>Cerebrovascular disease</td>
<td>1348 (15.2)</td>
</tr>
<tr>
<td>Depression</td>
<td>1314 (14.7)</td>
</tr>
<tr>
<td>Lung cancer</td>
<td>308 (3.5)</td>
</tr>
<tr>
<td>Depression</td>
<td>1244 (13.6)</td>
</tr>
<tr>
<td>Cancer</td>
<td>978 (11.0)</td>
</tr>
<tr>
<td>COPD medications, n (%)</td>
<td>8053 (87.4)</td>
</tr>
<tr>
<td>Any COPD medication</td>
<td>6896 (74.8)</td>
</tr>
<tr>
<td>Short-acting beta-agonists</td>
<td>834 (9.0)</td>
</tr>
<tr>
<td>Short-acting anticholinergics</td>
<td>4273 (46.3)</td>
</tr>
<tr>
<td>Tiotropium</td>
<td>1228 (13.3)</td>
</tr>
<tr>
<td>Inhaled corticosteroids (ICS)</td>
<td>476 (5.2)</td>
</tr>
<tr>
<td>Long-acting beta agonists (LABA)</td>
<td>5146 (55.8)</td>
</tr>
</tbody>
</table>

¹MRC Grades 1-5 were captured; these are equivalent (and identical wording to) mMRC 0-4: 0=only breathless with strenuous exercise; 1=breathless when hurrying on level or up a slight hill; 2=walk slower than people of same age on the level due to breathlessness or stop for breath when walking on level at own pace; 3=stop for breath after walking 100 yards or a few minutes on the level; 4=too breathless to leave house or breathless when dressing
Table 2: Demographic and clinical characteristics split by 2011 and 2007 GOLD categories for patients with both a valid FEV1 and a valid mMRC score (n=6283)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total cohort</th>
<th>GOLD 2011</th>
<th>GOLD 2007</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>n (% of total population)</td>
<td>6283 (100)</td>
<td>2265 (36.0)</td>
<td>1198 (19.1)</td>
</tr>
<tr>
<td>Age, years, mean (SD)</td>
<td>69.2 (10.6)</td>
<td>67.7 (10.8)</td>
<td>71.6 (10.7)</td>
</tr>
<tr>
<td>Males, n (%)</td>
<td>3261 (52.0)</td>
<td>1202 (53.2)</td>
<td>560 (46.9)</td>
</tr>
<tr>
<td>Current smokers</td>
<td>2287 (36.5)</td>
<td>842 (37.3)</td>
<td>422 (35.3)</td>
</tr>
<tr>
<td>Body mass index, kg/m&lt;sup&gt;2&lt;/sup&gt;, mean (SD)</td>
<td>27.3 (6.2)</td>
<td>27.5 (5.8)</td>
<td>28.8 (7.5)</td>
</tr>
<tr>
<td>FEV1, % of predicted, mean (SD)</td>
<td>60.4 (19.8)</td>
<td>71.9 (14.6)</td>
<td>68.9 (14.0)</td>
</tr>
<tr>
<td>Co-morbidities, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any co-morbidity</td>
<td>4966 (79.0)</td>
<td>1702 (75.1)</td>
<td>1003 (83.7)</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>2984 (47.6)</td>
<td>950 (42.0)</td>
<td>621 (52.0)</td>
</tr>
<tr>
<td>Asthma</td>
<td>1366 (21.8)</td>
<td>461 (20.4)</td>
<td>246 (20.6)</td>
</tr>
<tr>
<td>Cerebrovascular disease</td>
<td>1005 (16.2)</td>
<td>338 (15.2)</td>
<td>234 (19.9)</td>
</tr>
<tr>
<td>Depression</td>
<td>872 (14.4)</td>
<td>282 (12.9)</td>
<td>179 (15.4)</td>
</tr>
<tr>
<td>Cancer</td>
<td>879 (14.4)</td>
<td>288 (13.1)</td>
<td>185 (15.8)</td>
</tr>
<tr>
<td>Lung cancer</td>
<td>210 (3.5)</td>
<td>72 (3.3)</td>
<td>46 (3.9)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>821 (13.1)</td>
<td>260 (11.5)</td>
<td>199 (16.6)</td>
</tr>
<tr>
<td>Osteoporosis</td>
<td>648 (10.7)</td>
<td>196 (8.9)</td>
<td>106 (9.0)</td>
</tr>
</tbody>
</table>

GOLD 2011: A: low risk, less symptoms; B: low risk, more symptoms; C: high risk, less symptoms; D: high risk, more symptoms; GOLD 2007: I FEV1≥80; II FEV1 50-79; III FEV1 30-49; IV FEV1≤29. <sup>1</sup>From chi-squared test or ANOVA across GOLD categories; <sup>2</sup>Not done (ND) as categories defined by differences in FEV1 % predicted
Figures legend

Figure 1: Study flow according to patient data

- Total patient population in 80 general practices
  N=540793

  Read Code diagnosis of COPD
  N=9219

  Diagnosis of COPD and spirometry data available
  N=7480

  - Diagnosis of COPD and a valid FEV₁ and a valid mMRC score
    N=6283
  
  - Diagnosis of COPD and a valid FEV₁ and a valid CAT score
    N=221
Figure 2a: Distribution of patients using the 2011 GOLD categories

NB: includes patients with both a valid FEV₁ and a valid mMRC score (n=6283)
Figure 2b: Distribution of patients using the 2013 GOLD categories

NB: includes patients with both a valid FEV₁ and a valid mMRC score (n=6283)
Figure 3: Comparison of the Distribution of patients using GOLD 2007 (I to IV) and GOLD 2011 classification (A to D)

NB: includes patients with both a valid FEV$_1$ and a valid mMRC score (n=6283)
Figure 4: Overlap of risk categories (airflow limitation and exacerbation history)

NB: includes patients with a valid FEV$_1$ (n=7480)