BMJ Open

Risk factors for acute exacerbations in primary care COPD patients: a retrospective observational cohort study

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2014-006171
Article Type:	Research
Date Submitted by the Author:	21-Jul-2014
Complete List of Authors:	Müllerová, Hana; GlaxoSmithKline, Worldwide Epidemiology Shukla, Amit; GlaxoSmithKline, Worldwide Epidemiology Hawkins, Adam; GlaxoSmithKline, Global Respiratory Franchise Quint, Jennifer; London School of Hygiene and Tropical Medicine, Epidemiology
Primary Subject Heading:	Respiratory medicine
Secondary Subject Heading:	Epidemiology
Keywords:	chronic obstructive pulmonary disease, database, electronic medical records, exacerbations, PRIMARY CARE

Risk factors for acute exacerbations in primary care COPD patients: a retrospective observational cohort study

Hana Müllerová,[1] Amit Shukla,[1] Adam Hawkins,[2] Jennifer Quint[3]

[1]Worldwide Epidemiology, GlaxoSmithKline, Uxbridge, UK

[2]Global Respiratory Franchise, GlaxoSmithKline, Uxbridge, UK

[3]London School of Hygiene and Tropical Medicine, London, UK

Correspondence to

Hana Müllerová; Worldwide Epidemiology, GlaxoSmithKline R&D, Building 9, Iron Bridge Road, Stockley Park West, Uxbridge, Middlesex, UB11 1BT, UK

Tel: +44 208 990 2647. Fax: +44 208 990 3505

hana.x.muellerova@gsk.com

Keywords

Chronic obstructive pulmonary disease, database, electronic medical records, exacerbations, primary care

Word count

(excluding title page, abstract, references, figures and tables): 3,127 words

ABSTRACT

Objectives: To evaluate risk factors associated with exacerbation frequency in primary care. Information on exacerbations of chronic obstructive pulmonary disease (COPD) has mainly been generated by secondary care-based clinical cohorts.

Design: Retrospective observational cohort study.

Setting: Electronic medical records database (England and Wales).

Participants: 58 589 COPD patients aged ≥40 years with COPD diagnosis recorded between 1 April 2009 and 30 September 2012, and with at least 365 days of follow-up before and after the COPD diagnosis, were identified in the Clinical Practice Research Datalink. Mean age 69 years; 47% female; mean forced expiratory volume in 1 second (FEV_1) 60% predicted.

Outcome measures: Data on moderate or severe exacerbation episodes defined by diagnosis and/or medication codes 12 months following cohort entry were retrieved, together with demographic and clinical characteristics. Associations between patient characteristics and odds of having none vs one, none vs frequent (≥2) and one vs frequent exacerbations over 12 months follow-up were evaluated using multivariate logistic regression models.

Results: During follow-up, 23% of patients had evidence of frequent moderate-to-severe COPD exacerbations (24% one; 53% none). Independent predictors of increased odds of having exacerbations during the follow-up, either frequent episodes or one episode, included prior exacerbations, increasing dyspnoea score, increasing grade of airflow limitation, females and prior or current history of several comorbidities (e.g. asthma, depression, anxiety, heart failure and cancer).

Conclusions: Primary care-managed COPD patients at the highest risk of exacerbations can be identified by exploring past medical history for presence of prior exacerbations, greater COPD disease severity and co-occurrence of other medical conditions.

Funding: GlaxoSmithKline (WEUSKOP5903)

ARTICLE SUMMARY

Strengths and limitations of the study

- We report the results of a retrospective cohort study of 58,589 patients identified in the UK Clinical Practice Research Datalink (CPRD) with spirometry-confirmed COPD diagnosis during a 42-month period from 1 April 2009, with the primary objective of evaluating risk factors for acute exacerbations of COPD.
- Strengths of the study include: the large size of the cohort; the cohort is representative of the UK primary care population; the utilisation of the CPRD, a comprehensive and well-established electronic medical records database which incorporates spirometric scores and dyspnoea grades recorded as part of the Quality Outcomes Framework.
- The main limitation of this study is that the event recording of acute exacerbations in the CPRD is not entirely standardized , necessitating the use of an algorithm containing prescription records for oral corticosteroid and antibiotics as well as primary care visits for exacerbations, and A&E and hospital admissions for COPD to maximise the detection of exacerbations.
- Another possible study limitation is that the true population of patients with milder airflow obstruction and those who did not experience exacerbations may be underrepresented in the dataset due to such patients not seeking healthcare or undergoing spirometry.

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is characterised by the progressive deterioration of lung function. In many patients, the course of the disease is punctuated by exacerbations, an acute worsening of symptoms which, in severe cases, can necessitate hospitalisation and even result in death.[1] Exacerbations are the major cost driver, directly and indirectly, in COPD.[2] COPD exacerbations are known to cluster in time and patients who have experienced one episode are at an increased risk of a recurrent exacerbation.[3] Furthermore, exacerbations accelerate the deterioration of lung function,[4] which in turn increases the likelihood and severity of further exacerbations.

Various definitions are used to mark the start, end, relapse and recurrence of exacerbations.[5, 6] Frequent exacerbations, usually defined as two or more episodes per year, are the focus of the majority of clinical epidemiology studies, as they are considered a marker of greater disease burden and reduced survival.[7] However, an increase in frequency of any moderate-to-severe exacerbation events has been associated with greater disease severity, and the best predictor of future events is a history of exacerbation(s).[7]

Previous research into the natural history of COPD exacerbations has mainly focused on prospective cohorts of patients who have usually been recruited from secondary and/or tertiary care. Limited data are available on the burden of COPD exacerbations in patients who are managed in primary care. The use of electronic medical record (EMR) databases allow for an identification of very large-scale patient cohorts in a setting much more representative of clinical practice than prospective studies carried out in selective cohorts. EMR-based cohorts represent all patients diagnosed and treated in primary care, with information reflecting real-world healthcare provider decisions and patient behaviours without any interventions introduced. Furthermore, the retrospective cohort approach avoids the potential for behavioural biases associated with selection and active participation in research, the so-called ‘Hawthorne effect’.[8]

This retrospective, observational cohort study comprises patients who were identified from primary care using the UK Clinical Practice Research Datalink (CPRD-GOLD).

We aimed to evaluate records of episodes of moderate-to-severe exacerbations of COPD and to determine the factors associated with COPD exacerbation frequency in patients managed through primary care.

METHODS

Study design

We used a retrospective cohort study design and identified patients in the CPRD-GOLD who had a record of COPD diagnosis, defined as ≥1 record of COPD-specific READ codes from 1 April 2009 until 30 September 2012. We then took the first record of COPD diagnosis during this predefined period and assigned it as a cohort entry date (cohort baseline). We then further limited the patient cohort to only those with a minimum of 365 days of available history in the database prior to and subsequent to this first record (**figure 1**).

CPRD-GOLD database

The CPRD-GOLD, formerly known as the General Practice Research Database, is one of the largest computerised databases of linked anonymised primary care medical records in the world.[9] The CPRD-GOLD reflects the complete EMR for all NHS primary healthcare collected from approximately 8% of the population of England and Wales.[10] All medical signs, symptoms, investigations and diagnoses deemed important for the continued care of the patients are coded within the EMR.[10, 11]

The CPRD research group, a part of the UK Medicines and Healthcare Products Regulatory Agency, continuously monitors data quality and removes practices from the database if they fail to maintain the required standards of data entry.

Patient population

The cohort consisted of patients with a record of COPD diagnosis between 1 April 2009 and 30 September 2012 who sought care for their condition; see Study design. COPD diagnosis was required to be confirmed by spirometry recording consistent with obstructive disease, defined as a forced expiratory volume in 1 second (FEV_1)/forced vital capacity (FVC) ratio of <0.7 recorded any time before and up to 3 months from the date of diagnosis of COPD record that qualified the patient to enter this cohort.

Exacerbations of COPD

An algorithm was assembled to collect data on recorded events of moderate-to-severe COPD exacerbations from the 12 months preceding and the 12 months subsequent to cohort entry date. Moderate exacerbations were defined as a record of a diagnosis of exacerbation, acute bronchitis, or the management of COPD with specific antibiotics and

oral corticosteroids recorded within 5 days of each other. Admissions to hospital or visits to accident and emergency associated with COPD as recorded by GPs were considered as severe exacerbations. An exacerbation episode was defined based on events recorded over 2 weeks from the initial exacerbation record. The start of an episode was considered to be the date of the first qualifying event. Subsequently, a 14-day rolling window was applied to identify a period of at least 2 exacerbation-free weeks in order to ensure that a relapse was not categorised as a separate exacerbation episode.[12] The frequency of individual patients' exacerbations during the 12 months prior to the start of observation and also during the 12-month follow-up period was split into three categories: none, one, or two or more (frequent) episodes of moderate-to-severe exacerbation of COPD. The incidence of exacerbation episodes was also expressed as a rate per person per year, a sum of the episodes per patient divided by 365 days of follow-up.

Other key study variables

The following variables were also retrieved from the database using the latest record prior to cohort entry date: age, gender, smoking status, body mass index (BMI; or imputed BMI using the latest height and weight measurements where BMI data were not recorded), Medical Research Council (MRC) dyspnoea grade and FEV_1 percent predicted assessment with the closest date to cohort entry.[13]

Mean BMI was summarised as both a continuous and categorical variable using the World Health Organisation classification of underweight (<18.5), normal (18.5 to 24.9), overweight (25.0 to 29.9) and obese (≥30).[14]

The Global Initiative for Chronic Obstructive Lung Disease (GOLD) grade classification [15] was used to determine the severity of airflow limitation. Due to the nature of the data available from CPRD-GOLD, it was not possible to determine whether the spirometry data had only been recorded following the administration of a bronchodilator. However, given that the NHS Quality Outcomes Framework requires all COPD patients to have their COPD diagnosis confirmed with post-bronchodilator spirometry,[16] it is expected that the majority were recorded following bronchodilator administration. The categories of airflow severity limitation were defined using the FEV_1 cut-off points ≥80% predicted for mild, ≥50% to <80% predicted for moderate, ≥30% to <50% predicted for severe and >30% predicted for very severe airflow limitation.

MRC dyspnoea grade is a unidimensional measure of breathlessness related to activity.[17] The MRC dyspnoea scale is equivalent to the modified MRC (mMRC) scale, which is an established method in the published literature. The key difference with the

mMRC dyspnoea scale is a shift in the scoring range: whereas MRC ranged from 1 to 5, mMRC employs a range from 0 to 4. Therefore, mMRC Grade 0 is equal to Grade 1 on the MRC scale.[18]

Records of diagnoses of selected comorbidities, including depression, anxiety and asthma, occurring prior to the start of the study were retrieved from the database. Data on the frequency of prescription treatment for COPD in the 12 months prior to the start of this cohort study were also collected. Patients who had one or more records for medication within a therapeutic class, with the exception of oral corticosteroids, were considered to use that medication. For oral corticosteroids, four or more prescriptions within the 12-month period were considered to indicate regular use rather than acute use for COPD exacerbation(s). Lastly, the number of recorded contacts with a GP for any reason in the 12 months before the cohort entry was ascertained and standardised per 365.25 days.

Statistical analyses

A multiple, multilevel, logistic regression was used to assess risk factors associated with frequency of exacerbations over the 12-month follow-up period: frequent vs none, frequent vs one, and one vs none. These analyses were carried out using the PROC LOGISTICS procedures in SAS version 9.2. All covariates listed in descriptive tables, with the exception of respiratory medications, were entered into the models. Missing values for BMI and percent predicted FEV_1 values (less than 2% of the cohort) and MRC dyspnoea score (12% of the cohort) were not imputed.

The relationships between the number of exacerbation episodes during the 12-month follow-up period and the percentage predicted FEV_1 value and MRC dyspnoea grade were assessed using Spearman's correlation coefficient.

The study protocol (WEUSKOP5903) was reviewed and approved as Protocol 12_118 by the CPRD Scientific and Ethics committee.

RESULTS

Patients

The cohort consisted of 58,589 patients with COPD (**figure 1**). Overall, 47% of patients were female, mean age was 69 years and mean FEV_1 was 60% predicted.

Incidence of exacerbation

During the 12-month follow-up period, 53% of patients (n=31 049) had no recorded episodes of moderate-to-severe exacerbation of COPD, 24% of patients (n=14 189) had one recorded exacerbation episode and frequent episodes were recorded in 23% of patients (n=13 351). The rate of moderate-to-severe exacerbations during the 12-month follow-up period was 0.89 per person per year. A maximum of eight episodes per patient was observed.

In comparison to patients with no recorded exacerbation episodes, patients with frequent exacerbations were more likely to be female (51% vs 44%), had poorer lung function (mean FEV_1 percent predicted: 54% vs 62%) and more dyspnoea (MRC grade ≥3: 58% vs 39%). When compared with patients with no recorded exacerbation episodes during the 12-month follow-up, patients with frequent exacerbations were more regularly diagnosed with comorbidities including myocardial infarction, heart failure and depression, and had been managed with maintenance treatment more often in the past 12 months (≥1 prescribed inhaled corticosteroid/long-acting beta agonist 68% vs 39%, ≥1 prescribed long-acting muscarinic antagonist 53% vs 28%) (**table 1**).

Table 1 Characteristics of COPD patients by frequency of recorded exacerbation episodes during 12-month follow-up

	Moderate-to-severe exacerbation frequency		
	None	**One**	**Two or more**
Patient characteristic	**n=31 049**	**n=14 189**	**n=13 351**
Age, mean (SD)	69.3 (10.5)	69.8 (10.3)	69.4 (10.2)
Smoking status			
Current	10 531 (34)	4667 (33)	4173 (31)
Former	15 988 (51)	7474 (53)	7345 (55)
Never	3312 (11)	1492 (11)	1301 (10)
Other	1218 (4)	556 (4)	532 (4)
Female	13 619 (44)	6831 (48)	6832 (51)
BMI underweight (<18.5)	1327 (4)	641 (5)	751 (6)
BMI normal (18.5–24.9)	10 533 (34)	4810 (34)	4668 (35)
BMI overweight (25.0–29.9)	10 633 (34)	4778 (34)	4274 (32)
BMI obese (≥30.0)	8037 (26)	3760 (26)	3472 (26)
Rate of moderate-to-severe exacerbations in prior 12 months per person per year, mean	0.37	0.74	1.67
Rate of exacerbations with hospital admission in prior 12 months per person per year, mean	0.07	0.12	0.2
Comorbidities any time in history			
Acute myocardial infarction	2530 (8)	1386 (10)	1350 (10)
Congestive heart disease	1857 (6)	1041 (7)	1165 (9)
Depression	3976 (13)	2116 (15)	2393 (18)
≥1 prescription within 12 months before cohort entry			
ICS	5789 (19)	2883 (20)	2501 (19)
LABA	2379 (8)	1278 (9)	1295 (10)

ICS / LABA	12 150 (39)	7328 (52)	9133 (68)
LAMA	8618 (28)	5260 (37)	7049 (53)
MRC Dyspnoea scale score			
MRC 1	6251 (23)	2059 (16)	1165 (10)
MRC 2	11 552 (42)	4765 (38)	3546 (31)
MRC 3	6478 (23)	3487 (28)	3286 (29)
MRC 4	2959 (11)	1876 (15)	2605 (23)
MRC 5	524 (2)	398 (3)	690 (6)
Airflow limitation grade			
mild	5475 (18)	1971 (14)	1398 (10)
moderate	16 815 (54)	7324 (52)	5693 (43)
severe	7013 (23)	3855 (27)	4629 (35)
very severe	1152 (4)	814 (6)	1411 (11)

All data are n (%) unless otherwise stated. Percentages expressed as column %.

BMI, body mass index; COPD, chronic obstructive pulmonary disease; ICS, inhaled corticosteroid; LABA, long-acting beta$_2$ agonist; LAMA, long-acting muscarinic antagonist; MRC, Medical Research Council; SD, standard deviation

Of the patients with no recorded exacerbation episodes during the 12-month follow-up period, 72% had no record of a moderate-to-severe exacerbation in the prior 12 months; 46% of patients who experienced frequent moderate-to-severe exacerbations during the follow-up period had also experienced two or more exacerbations in the prior 12 months (**figure 2**).

A weak, although statistically significant, relationship was observed between the frequency of COPD exacerbations and the stage of airflow limitation (Spearman's r=0.16, p<0.001). Thirty-four percent and 45% of patients with no recorded exacerbations also had poor lung function (moderate and severe airflow limitation, respectively), in contrast to 16% and 19% of patients with frequent exacerbation episodes who had mild or moderate airflow limitation, respectively (**figure 3a**).

Factors associated with exacerbation frequency

Comparison of patients with no exacerbation, one exacerbation and frequent exacerbations during the 12-month follow-up showed that several factors were associated with an increased likelihood of prospective frequent exacerbations; the strongest association was with prior exacerbations. Experiencing one moderate exacerbation, compared with none, during the 12 months preceding the start of the study was associated with one moderate-to-severe exacerbation occurring (odds ratio [OR] 1.9) or frequent (≥2) moderate-to-severe exacerbations occurring (OR 3.3) during the 12-month follow-up period (**table 2**). The OR for frequent (≥2) moderate-to-severe exacerbations increased to 13.6 if patients had experienced two or more moderate episodes, compared with none in the 12 months preceding the start of the study.

Table 2 Factors associated with moderate-to-severe exacerbations during the 12-month follow-up period [three-category multi-level model]

	Moderate-to-severe exacerbations during 12-month follow-up					
	≥2 vs none		**One vs none**		**≥2 vs one**	
Factor	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
COPD exacerbations, 12 months prior to observation period start						
No moderate event	Reference					
1 moderate episode	3.31	3.12 to 3.51	1.89	1.79 to 1.99	1.76	1.65 to 1.87
≥2 moderate episodes	13.64	12.67 to 14.68	3.11	2.88 to 3.37	4.41	4.11 to 4.74
No hospitalised episode	Reference					
≥1 hospitalised episode	2.13	1.95 to 2.31	1.49	1.38 to 1.61	1.44	1.32 to 1.57
Airflow limitation level nearest to observation period start						
Mild (FEV_1 ≥80% predicted)	Reference					
Moderate (≥50% FEV_1 <80%	1.23	1.13 to 1.33	1.19	1.12 to 1.27	1.04	0.95 to 1.14

predicted)						
Severe (≥30% FEV_1 <50% predicted)	1.81	1.66 to 1.98	1.36	1.27 to 1.47	1.37	1.24 to 1.50
Very severe (FEV_1 >30% predicted)	2.37	2.08 to 2.70	1.59	1.41 to 1.79	1.56	1.37 to 1.78
MRC dyspnoea scale nearest to observation period start						
1	Reference					
2	1.39	1.28 to 1.51	1.16	1.09 to 1.23	1.16	1.06 to 1.27
3	1.86	1.70 to 2.03	1.41	1.31 to 1.51	1.32	1.20 to 1.44
4	2.74	2.49 to 3.03	1.55	1.43 to 1.69	1.74	1.57 to 1.93
5	3.02	2.58 to 3.55	1.62	1.39 to 1.88	1.84	1.57 to 2.16
Comorbidities, history of medical diagnosis before observation period start (absence is referent category)						
Heart failure	1.2	1.08 to 1.32	1.1	1.00 to 1.20	1.11	1.00 to 1.23
Myocardial infarction	1.13	1.03 to 1.24	1.15	1.06 to 1.24	0.96	0.87 to 1.05
Rheumatological disease	1.12	1.01 to 1.24	1.1	1.01 to 1.21	1.04	0.94 to 1.16
Renal disease	0.9	0.84 to 0.97	0.96	0.91 to 1.02	0.93	0.86 to 1.00
Anxiety	1.16	1.08 to 1.25	1.14	1.08 to 1.22	1.02	0.95 to 1.09
Depression	1.25	1.16 to 1.35	1.1	1.02 to 1.17	1.12	1.04 to 1.21
Asthma	1.51	1.43 to 1.60	1.24	1.19 to 1.30	1.23	1.16 to 1.30
Cancer	1.28	1.19 to 1.38	1.1	1.03 to 1.17	1.14	1.05 to 1.23
Number of contacts with GP, 12 months prior to observation period start						
Low: 0–5	Reference					
Medium: 6–10	1.02	0.95 to 1.10	1.07	1.00 to 1.13	0.99	0.91 to 1.07
High: ≥10	1.24	1.16 to 1.32	1.18	1.12 to 1.24	1.08	1.01 to 1.16
Sex						
Female	1.19	1.13 to 1.26	1.12	1.07 to 1.17	1.07	1.01 to 1.14

Models further adjusted for age, smoking and BMI.

BMI, body mass index; CI, confidence interval; COPD, chronic obstructive pulmonary disease; FEV_1, forced expiratory volume in 1 second; MRC: Medical Research Council

The risk of frequent exacerbations compared with none, during the 12-month observation period, increased with greater severity of airway limitation; OR of 1.2 for moderate limitation (≥50% FEV_1 <80% predicted) and OR of 2.4 for very severe limitation (FEV_1 >30% predicted) (**table 2**). Increasing dyspnoea grade was also associated with an increased OR for frequent exacerbations compared with no exacerbations, from 1.4 for MRC2 to 3.0 for MRC5 when compared with MRC1 (**table 2**; **figure 3b**). The presence of specific comorbidities was associated with a significant increase in the risk of patients experiencing frequent exacerbations compared with none during the observational period (**table 2**). Among the specific comorbidities assessed, asthma, cancer and depression had the strongest observed associations with exacerbation frequency, with corresponding ORs for frequent vs no exacerbations of 1.5, 1.3 and 1.3, respectively (**table 2**). Similar risk factors for one vs no exacerbation, albeit of lower magnitude, were observed as for the risk of frequent vs no exacerbations.

DISCUSSION

We set out to explore COPD exacerbation frequency and associated factors, using a retrospective cohort approach, in a large primary care cohort of 58 589 patients in England and Wales. We found marked differences between patients managed in primary care who had frequent exacerbations (23% of patients) or one exacerbation (24%) as compared with patients with no recorded COPD exacerbations (53%) during the 12-month follow-up period.

The characteristic most strongly associated with moderate-to-severe exacerbation frequency was a history of exacerbations. A gradient relationship between moderate exacerbations in the prior 12 months and prospective moderate-to-severe exacerbations was apparent: patients who had no exacerbations in the prior 12 months had the lowest risk of future events; patients who had one exacerbation were at a greater risk of future episodes than those who had none, but were at a lower risk than those who had two or more prior exacerbations; patients with two or more prior exacerbations were at the highest risk of future exacerbations. Prior episodes of severe exacerbations also independently increased the odds of future exacerbations. This result is in agreement with previously reported findings.[7] When moderate and severe prior exacerbations are considered separately, our findings provide additional supportive evidence of an independent and 'dose-like' relationship of prior moderate events and future exacerbation risk.

Exacerbation frequency also increased with increasing grade of airflow limitation and, similarly, with increasing dyspnoea score. A relationship between exacerbation frequency and symptom severity is well established on the basis of findings from prospective studies;[19, 20] however, the two factors are semi-independent and patients with moderate airflow limitation can experience recurrent exacerbations.[7] Again, both airflow limitation and dyspnoea were associated with a higher risk for any exacerbation event, not only with frequent events. Likewise, the slightly increased probability of frequent exacerbations observed to be associated with females is consistent with previous reports.[19] No relationships between age or smoking status and exacerbation frequency were observed.

Finally, we observed a relationship between a prior diagnosis of selected comorbidities in patient history and a risk of any future moderate-to-severe exacerbations. Consistent independent associations were observed between a history of asthma, depression, anxiety, heart failure and cancer and either one or two or more exacerbation episodes. It has previously been reported that COPD exacerbations are associated with cardiovascular events, including myocardial infarction and heart failure,[21] and the presence of these co-morbidities increases the likelihood of mortality and costs associated with exacerbations.[2] Likewise, the importance of comorbid asthma as a risk factor for more frequent exacerbation

and accelerated disease progression has led to the description of an 'asthma-COPD overlap syndrome'.[22, 23] It has been estimated that overlap syndrome may be associated with up to a three-fold increase in the frequency and severity of exacerbations.[22] Overlap syndrome is typically seen in younger patients with a smoking burden that is below the average for COPD patients, potentially providing a partial explanation of the absence of observed relationships between age, smoking status and exacerbation frequency in this study.

Analyses of this large, primary care-based cohort supported the established hypothesis of the existence of a 'frequent exacerbator' phenotype, widely reported in the COPD literature as an explanation for observations of a subgroup of patients who experience highly recurrent exacerbations.[7, 24] The pathophysiology underlying the frequent exacerbator phenotype is not yet fully understood, and identification of the interrelated factors is potentially of value in aiding the identification of patients who are likely to be at an increased risk of further exacerbations. Early intervention is associated with faster recovery from exacerbation and improved health-related quality of life.[25]

The factors that we have observed to be associated with any future moderate-to-severe exacerbations are suggestive of a shared inflammatory biological mechanism that gives rise to further exacerbation events and also leads to worsening of COPD symptoms. Markers for airway and systemic inflammation have been shown to be associated with exacerbation frequency,[1] may precipitate lung function decline,[19] and may also partly explain the observed relationships between exacerbation frequency and comorbidities such as cardiovascular disease [21, 26] and asthma.[27] Recent studies have identified a number of potential susceptibility factors for exacerbation of COPD, including increased levels of inflammatory markers such as fibrinogen.[28, 29] Furthermore, dynamic lung hyperinflation correlates strongly with the patient-reported symptom of dyspnoea and is a consequence of the acute expiratory flow limitation that occurs during exacerbation.[30, 31] In combination with the heightened inflammation, this deterioration in airway function may increase the likelihood and severity of future exacerbation events, as reflected in the current GOLD guidelines.[15]

We found that the characteristics of patients who experienced at least one exacerbation event during the follow-up period differed from those who did not have any exacerbations. This observation replicated earlier findings from a clinical cohort of 2138 patients enrolled in the Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints (ECLIPSE) study.[7] However, it also should be noted that patients who do not experience exacerbations and have mild airflow limitation are less likely to seek healthcare and are, therefore, less likely to be diagnosed with COPD.[32]

In this study, an inclusive definition of exacerbations was used, encompassing prescription records, as well as exacerbation event coding and GP records of hospital admissions and accident and emergency visits. Despite this, more than half of the cohort experienced no exacerbations during the 12-month observation period. Although a spirometry recording was required for eligibility, COPD patients in the present study were identified based on a record of their diagnosis from a primary care database rather than actively recruited when visiting a physician. Therefore, we cannot ensure that all patients met the definition of COPD as per the GOLD guidelines.[15]

A strength of this study is the exploration of exacerbation data from a large sample of patients diagnosed with COPD obtained from a well-established, comprehensive primary care EMR-based data resource, the CPRD-GOLD. This resource provides representative coverage of the patient population in England and Wales, and has incorporated spirometry and dyspnoea score data as a part of routine COPD patient disease management since 2004. The CPRD-GOLD has previously been used in COPD research to determine disease prevalence in the UK [33] and to investigate comorbidities in COPD and asthma.[34, 35]

In conclusion, we demonstrated that, in a primary population of patients diagnosed with COPD, prior history of exacerbations (as well as increasing severity of airflow limitation and dyspnoea) and presence of comorbidities predict risk of future moderate-to-severe exacerbation events. We also showed that not only patients with prior frequent exacerbations, but also those with only one prior moderate exacerbation, are at increased risk of future exacerbation events. As exacerbations can accelerate the progression of the disease, our findings underline the importance of identifying COPD patients who are at risk of exacerbation and intervening as early as possible. Such identification is feasible for a primary care practitioner based on a review of clinical history routinely recorded in the UK electronic health record.

Collaborators Editorial support in the form of development of the manuscript first draft in consultation with the authors, editorial suggestions to draft versions of this paper, assembling tables and figures, collating author comments, copyediting, fact checking, referencing and graphic services was provided by Ian Grieve, PhD, at Gardiner-Caldwell Communications, an Ashfield company (Macclesfield, UK) and was funded by GlaxoSmithKline.

Competing interests HM, AS and AH are employees of and hold stock in GlaxoSmithKline. JQ received funding from an MRC Population Health Scientist Fellowship, grant number G0902135, and declares no competing interests.

Contributors HM wrote the study protocol and analysis plan, oversaw the data analysis and drafted the manuscript. JQ and AH contributed to the conception and design of the study and the interpretation of the data. AS contributed to the analysis and the interpretation of the data.

Funding This work was supported by GlaxoSmithKline (study number WEUSKOP5903).

Ethics approval The study protocol (WEUSKOP5903) was reviewed and approved as Protocol 12_118 by the CPRD Scientific and Ethics committee.

Data sharing statement No additional data available: access to the data used in this analysis, CPRD-GOLD database, is governed by the Ethics and Scientific committee: www.cprd.com/isac/. The authors cannot allow access to the CPRD-GOLD database because of contractual and ethics regulations imposed by the CPRD and its Ethics committee.

The analysis was conducted according to the study protocol. Additional statistical analyses included in the final report are available from the corresponding author.

REFERENCES

1. Donaldson GC, Seemungal TA, Patel IS, *et al*. Airway and systemic inflammation and decline in lung function in patients with COPD. *Chest* 2005;128(4):1995–2004.
2. Perera PN, Armstrong EP, Sherrill DL, *et al*. Acute exacerbations of COPD in the United States: inpatient burden and predictors of costs and mortality. *COPD* 2012;9(2):131–41.
3. Hurst JR, Donaldson GR, Quint JK, *et al*. Temporal Clustering of Exacerbations in Chronic Obstructive Pulmonary Disease. *Am J Respir Crit Care Med* 2009;179:369–74.
4. Vestbo J, Edwards LD, Scanlon PD, *et al*. Changes in forced expiratory volume in 1 second over time in COPD. *New Engl J Med* 2011;365(13):1184–92.
5. Aaron SD, Donaldson GC, Whitmore GA, *et al*. Time course and pattern of COPD exacerbation onset. *Thorax* 2012;67(3):238–43.
6. Leidy NK, Murray LT. Patient-reported outcome (PRO) measures for clinical trials of COPD: the EXACT and E-RS. *COPD* 2013;10(3):393–8.
7. Hurst JR, Vestbo J, Anzueto A, *et al*. Susceptibility to Exacerbation in Chronic Obstructive Pulmonary Disease. *New Engl J Med* 2010;363:1128–38.
8. McCambridge J, Witton J, Elbourne DR. Systematic review of the Hawthorne effect: New concepts are needed to study research participation effects. *J Clin Epidemiol* 2014;67:267–77.
9. Tate AR, Beloff N, Al-Radwan B, *et al*. Exploiting the potential of large databases of electronic health records for research using rapid search algorithms and an intuitive query interface. *J Am Med Inform Assoc* 2014;21:292–8.
10. Williams JEA, Green RH, Warrington V, *et al*. Development of the i-BODE: validation of the incremental shuttle walking test within the BODE index. *Respir Med* 2012;106(3):390–6.
11. Wood L, Martinez C. The general practice research database: role in pharmacovigilance. *Drug Safety* 2004;27:871–81.
12. Leidy NK, Wilcox TK, Jones PW, *et al*. Standardizing measurement of chronic obstructive pulmonary disease exacerbations. Reliability and validity of a patient-reported diary. *American J Respir Crit Care Med* 2011;183(3):323–9.

13. Bhaskaran K, Forbes HJ, Douglas I, *et al*. Representativeness and optimal use of body mass index (BMI) in the UK Clinical Practice Research Datalink (CPRD). *BMJ Open* 2013;3:e003389.

14. World Health Organisation. Obesity: preventing and managing the global epidemic. Report of a WHO Consultation. WHO Technical Report Series 894. Geneva: World Health Organisation, 2000. Available at: http://whqlibdoc.who.int/trs/WHO_TRS_894.pdf [last accessed July 2014].

15. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global Strategy for the Diagnosis, Management and Prevention of Chronic Obstructive Pulmonary Disease. Revised 2014. Available at: http://www.goldcopd.org/uploads/users/files/GOLD_Report_2014_Jun11.pdf [last accessed July 2014].

16. NICE. Quality and Outcomes Framework for COPD, NM63 Advisory Committee Recommendations. 2012. Available at: http://www.nice.org.uk/aboutnice/qof/indicators_detail.jsp?summary=13808 [last accessed July 2014].

17. Fletcher CM, Elmes PC, Fairbairn AS, *et al*. The significance of respiratory symptoms and the diagnosis of chronic bronchitis in a working population. *BMJ* 1959;2:257–66.

18. Williams T, van Staa T, Puri S, *et al*. Recent advances in the utility and use of the General Practice Research Database as an example of a UK Primary Care Data resource. *Ther Adv Drug Safety* 2012;3:89–99.

19. Burge S, Wedzicha JA. COPD exacerbations: definitions and classifications. *Eur Respir J* 2003;21(Suppl. 41):46s–53s.

20. Donaldson GC, Wedzicha JA. COPD exacerbations 1: epidemiology. *Thorax* 2006;61:164–8.

21. Donaldson GC, Hurst JR, Smith CJ, *et al*. Increased risk of myocardial infarction and stroke following exacerbation of COPD. *Chest* 2010;137(5):1091–7.

22. Hardin M, Silverman EK, Barr RG, *et al*. The clinical features of the overlap between COPD and asthma. *Respir Res* 2011;12:127.

23. De Marco R, Pesce G, Marcon A, *et al*. The Coexistence of Asthma and Chronic Obstructive Pulmonary Disease (COPD): Prevalence and Risk Factors in Young, Middle-aged and Elderly People from the General Population. *PLoS ONE* 2013;8(5):e62985.

24. Wedzicha JA, Brill SE, Allinson JP, *et al*. Mechanisms and impact of the frequent exacerbator phenotype in chronic obstructive pulmonary disease. *BMC Med* 2013;11:181.

25. Wilkinson TMA, Donaldson GC, Hurst JR, *et al*. Early Therapy Improves Outcomes of Exacerbations of Chronic Obstructive Pulmonary Disease. *Am J Respir Crit Care Med* 2004;169(12):1298–303.

26. Halpin DMG, Decramer M, Celli B, *et al*. Risk of nonlower respiratory serious adverse events following exacerbations in the 4-year UPLIFT trial. *Lung* 2011;189:261–8.

27. Zeki AA, Schivo M, Chan A, *et al*. The Asthma-COPD Overlap Syndrome: A Common Clinical Problem in the Elderly. *J Allergy* 2011:861926.

28. Thomsen M, Ingebrigtsen TS, Marott JL, *et al*. Inflammatory biomarkers and exacerbations in chronic obstructive pulmonary disease. *JAMA* 2013;309(22):2353.

29. Valvi D, Mannino DM, Müllerova H, *et al*. Fibrinogen, chronic obstructive pulmonary disease (COPD) and outcomes in two United States cohorts. *Int J Chron Obstruct Pulmon Dis* 2012;7:173–82.

30. O'Donnell DE, Laveneziana P. Dyspnea and activity limitation in COPD: mechanical factors. *COPD* 2007;4(3):225–36.

31. Stevenson NJ, Walker PP, Costello RW, *et al*. Lung mechanics and dyspnea during exacerbations of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2005;172:1510–6.

32. Bednarek M, Maciejewski J, Wozniak M, *et al*. Prevalence, severity and underdiagnosis of COPD in the primary care setting. *Thorax* 2008;63:402–7.

33. Soriano JB, Maier WC, Egger P, *et al*. Recent trends in physician diagnosed COPD in women and men in the UK. *Thorax* 2000;55:789–94.

34. Soriano JB, Visick GT, Muellerova H, *et al*. Patterns of co-morbidities in newly diagnosed COPD and asthma in primary care. *Chest* 2005;128(4):2099–107.

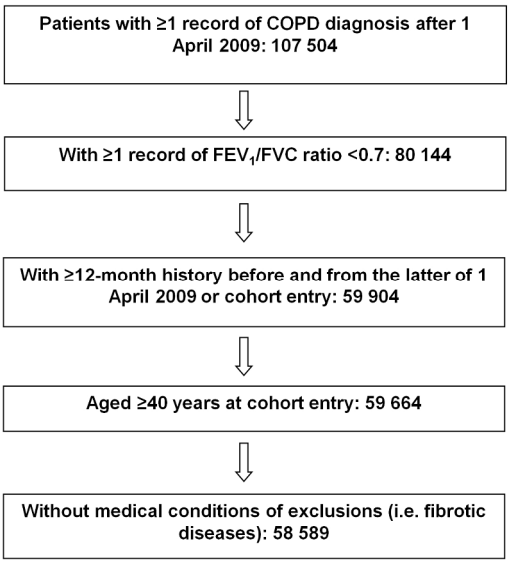
35. García Rodríguez LA, Wallander M-A, Martín-Merino E, *et al*. Heart failure, myocardial infarction, lung cancer and death in COPD patients: A UK primary care study. *Respir Med* 2010;104(11):1691–9.

FIGURE LEGENDS

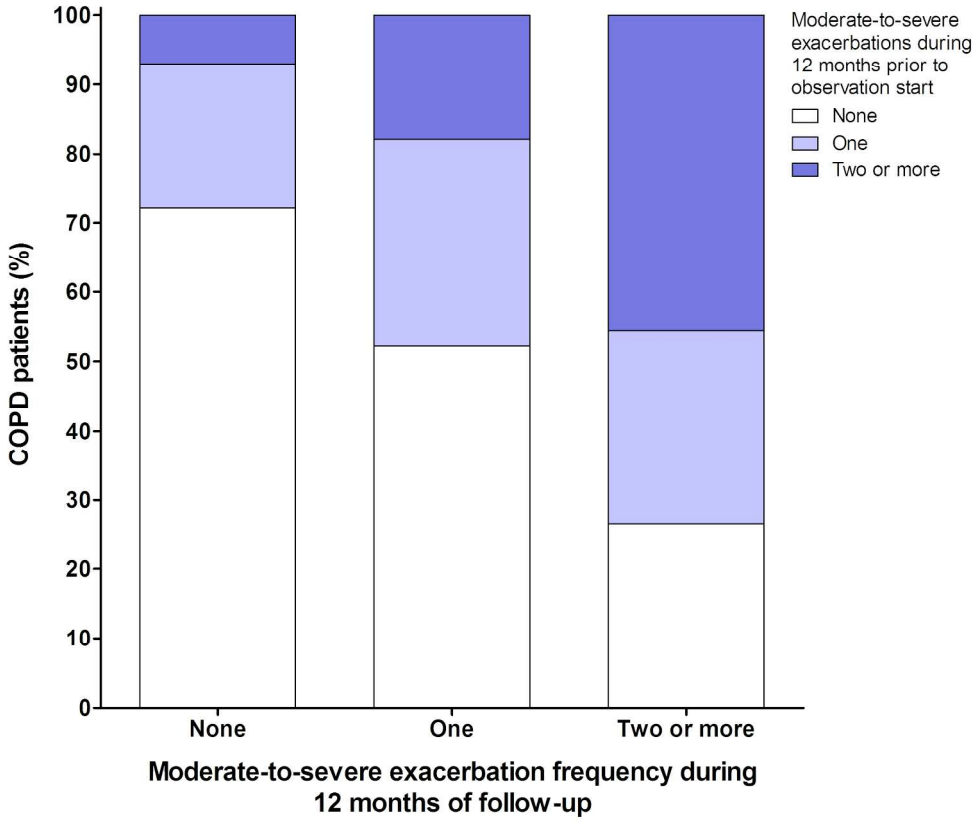
Figure 1 Patient recruitment flow chart.

Figure 2 Frequency of moderate-to-severe COPD exacerbations in the 12 months prior to cohort entry in relation to the frequency of exacerbation episodes during the 12-month follow-up period.

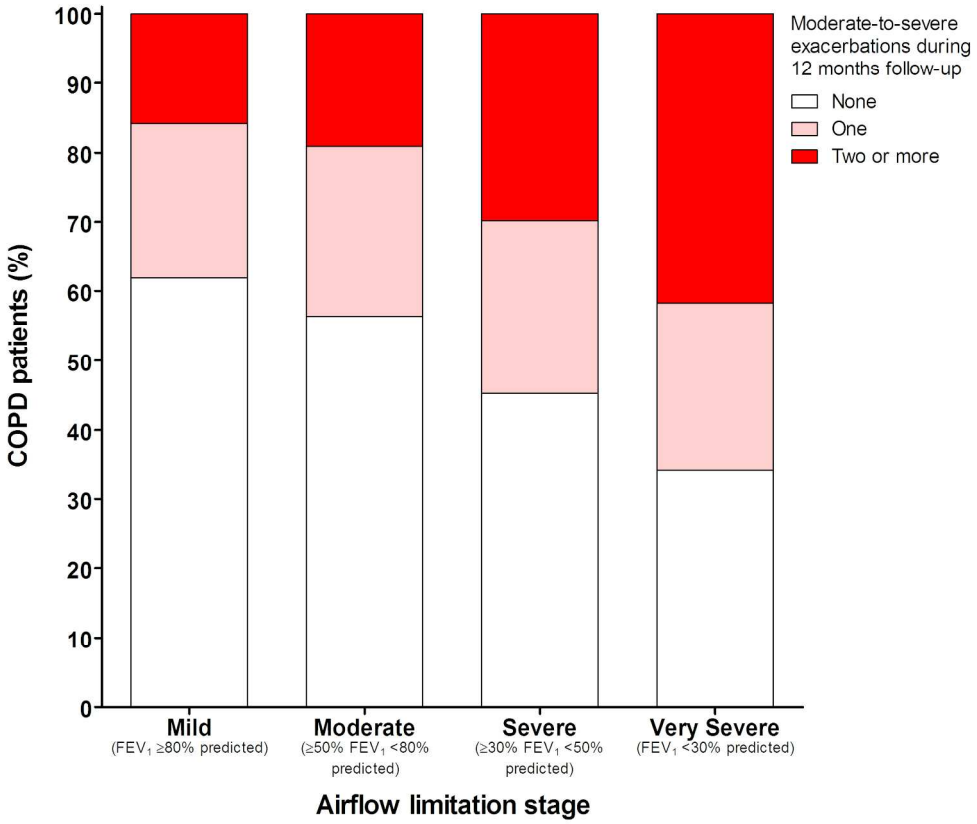
Figure 3 COPD exacerbation frequency during the 12-month follow-up period by a) airflow limitation stage, b) MRC dyspnoea grade.



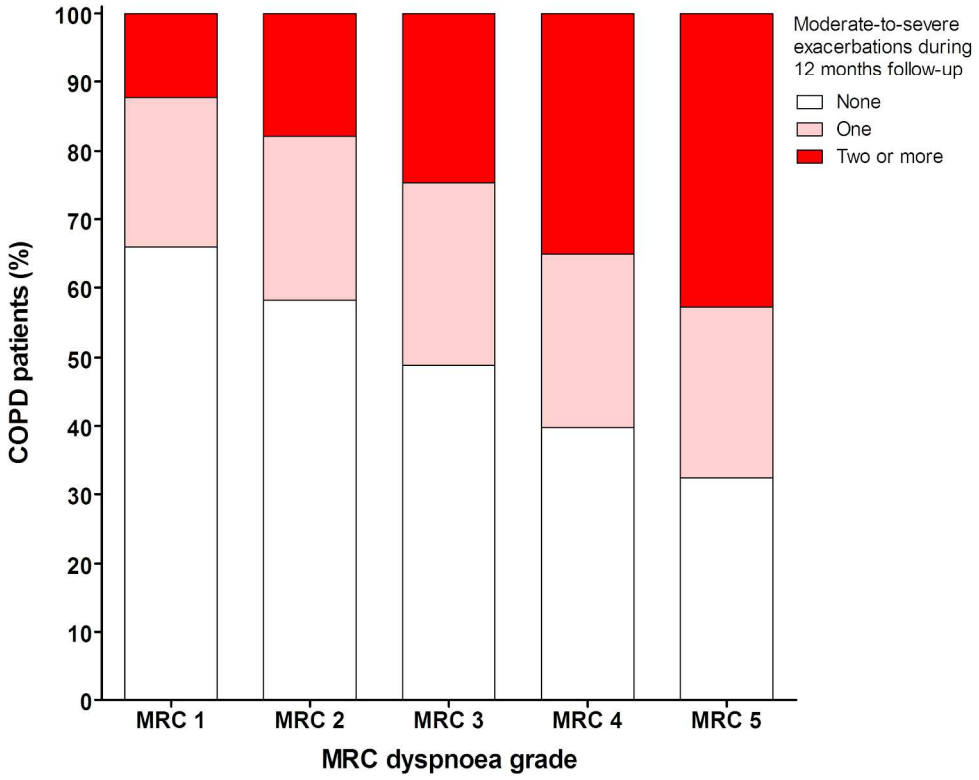
Patient recruitment flow chart.
190x254mm (300 x 300 DPI)



Frequency of moderate-to-severe COPD exacerbations in the 12 months prior to cohort entry in relation to the frequency of exacerbation episodes during the 12-month follow-up period.
213x185mm (300 x 300 DPI)



COPD exacerbation frequency during the 12 month follow-up period by airflow limitation stage.
214x184mm (300 x 300 DPI)



COPD exacerbation frequency during the 12 month follow-up period by MRC dyspnoea grade.
216x176mm (300 x 300 DPI)

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *cohort studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	4
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	5, Fig. 1
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5–7
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	5–7
Bias	9	Describe any efforts to address potential sources of bias	15–16
Study size	10	Explain how the study size was arrived at	5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	7
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	7
		(*b*) Describe any methods used to examine subgroups and interactions	7
		(*c*) Explain how missing data were addressed	7
		(*d*) If applicable, explain how loss to follow-up was addressed	N/A, all patients follow-up for a fixed period of 12 months

		(*e*) Describe any sensitivity analyses	N/A
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Fig. 1
		(b) Give reasons for non-participation at each stage	N/A
		(c) Consider use of a flow diagram	Fig. 1
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	8, Table 1
		(b) Indicate number of participants with missing data for each variable of interest	7
		(c) Summarise follow-up time (eg, average and total amount)	8
Outcome data	15*	Report numbers of outcome events or summary measures over time	8
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Table 2
		(*b*) Report category boundaries when continuous variables were categorized	Table 2
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	8
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	11, 13
Discussion			
Key results	18	Summarise key results with reference to study objectives	14
Limitations			
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	14–16
Generalisability	21	Discuss the generalisability (external validity) of the study results	14–16
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	17

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Risk factors for acute exacerbations in primary care COPD patients: a retrospective observational cohort study

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2014-006171.R1
Article Type:	Research
Date Submitted by the Author:	17-Oct-2014
Complete List of Authors:	Müllerová, Hana; GlaxoSmithKline, Worldwide Epidemiology Shukla, Amit; GlaxoSmithKline, Worldwide Epidemiology Hawkins, Adam; GlaxoSmithKline, Global Respiratory Franchise Quint, Jennifer; London School of Hygiene and Tropical Medicine, Epidemiology
Primary Subject Heading:	Respiratory medicine
Secondary Subject Heading:	Epidemiology
Keywords:	chronic obstructive pulmonary disease, database, electronic medical records, exacerbations, PRIMARY CARE

Risk factors for acute exacerbations in primary care COPD patients: a retrospective observational cohort study

Hana Müllerová,[1] Amit Shukla,[1] Adam Hawkins,[2] Jennifer Quint[3]

[1]Worldwide Epidemiology, GlaxoSmithKline, Uxbridge, UK

[2]Global Respiratory Franchise, GlaxoSmithKline, Uxbridge, UK

[3]London School of Hygiene and Tropical Medicine, London, UK

Correspondence to

Hana Müllerová; Worldwide Epidemiology, GlaxoSmithKline R&D, Building 9, Iron Bridge Road, Stockley Park West, Uxbridge, Middlesex, UB11 1BT, UK

Tel: +44 208 990 2647. Fax: +44 208 990 3505

hana.x.muellerova@gsk.com

Keywords

Chronic obstructive pulmonary disease, database, electronic medical records, exacerbations, primary care

Word count

(excluding title page, abstract, references, figures and tables): 3,127 words

ABSTRACT

Objectives: To evaluate risk factors associated with exacerbation frequency in primary care. Information on exacerbations of chronic obstructive pulmonary disease (COPD) has mainly been generated by secondary care-based clinical cohorts.

Design: Retrospective observational cohort study.

Setting: Electronic medical records database (England and Wales).

Participants: 58 589 COPD patients aged ≥40 years with COPD diagnosis recorded between 1 April 2009 and 30 September 2012, and with at least 365 days of follow-up before and after the COPD diagnosis, were identified in the Clinical Practice Research Datalink. Mean age 69 years; 47% female; mean forced expiratory volume in 1 second (FEV_1) 60% predicted.

Outcome measures: Data on moderate or severe exacerbation episodes defined by diagnosis and/or medication codes 12 months following cohort entry were retrieved, together with demographic and clinical characteristics. Associations between patient characteristics and odds of having none vs one, none vs frequent (≥2) and one vs frequent exacerbations over 12 months follow-up were evaluated using multivariate logistic regression models.

Results: During follow-up, 23% of patients had evidence of frequent moderate-to-severe COPD exacerbations (24% one; 53% none). Independent predictors of increased odds of having exacerbations during the follow-up, either frequent episodes or one episode, included prior exacerbations, increasing dyspnoea score, increasing grade of airflow limitation, females and prior or current history of several comorbidities (e.g. asthma, depression, anxiety, heart failure and cancer).

Conclusions: Primary care-managed COPD patients at the highest risk of exacerbations can be identified by exploring past medical history for presence of prior exacerbations, greater COPD disease severity and co-occurrence of other medical conditions.

Funding: GlaxoSmithKline (WEUSKOP5903)

ARTICLE SUMMARY

Strengths and limitations of the study

- We report the results of a retrospective cohort study of 58,589 patients identified in the UK Clinical Practice Research Datalink (CPRD) with spirometry-confirmed COPD diagnosis during a 42-month period from 1 April 2009, with the primary objective of evaluating risk factors for acute exacerbations of COPD.
- Strengths of the study include: the large size of the cohort; the cohort is representative of the UK primary care population; the utilisation of the CPRD, a comprehensive and well-established electronic medical records database which incorporates spirometric scores and dyspnoea grades recorded as part of the Quality Outcomes Framework.
- The main limitation of this study is that the event recording of acute exacerbations in the CPRD is not entirely standardized , necessitating the use of an algorithm containing prescription records for oral corticosteroid and antibiotics as well as primary care visits for exacerbations, and A&E and hospital admissions for COPD to maximise the detection of exacerbations.
- Another possible study limitation is that the true population of patients with milder airflow obstruction and those who did not experience exacerbations may be underrepresented in the dataset due to such patients not seeking healthcare or undergoing spirometry.

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is characterised by the progressive deterioration of lung function. In many patients, the course of the disease is punctuated by exacerbations, an acute worsening of symptoms which, in severe cases, can necessitate hospitalisation and even result in death.[1] Exacerbations are the major cost driver, directly and indirectly, in COPD.[2] COPD exacerbations are known to cluster in time and patients who have experienced one episode are at an increased risk of a recurrent exacerbation.[3] Furthermore, exacerbations accelerate the deterioration of lung function,[4] which in turn increases the likelihood and severity of further exacerbations.

Various definitions are used to mark the start, end, relapse and recurrence of exacerbations.[5, 6] Frequent exacerbations, usually defined as two or more episodes per year, are the focus of the majority of clinical epidemiology studies, as they are considered a marker of greater disease burden and reduced survival.[7] However, an increase in frequency of any moderate-to-severe exacerbation events has been associated with greater disease severity, and the best predictor of future events is a history of exacerbation(s).[7]

Previous research into the natural history of COPD exacerbations has mainly focused on prospective cohorts of patients who have usually been recruited from secondary and/or tertiary care. Limited data are available on the burden of COPD exacerbations in patients who are managed in primary care. The use of electronic medical record (EMR) databases allow for an identification of very large-scale patient cohorts in a setting much more representative of clinical practice than prospective studies carried out in selective cohorts. EMR-based cohorts represent all patients diagnosed and treated in primary care, with information reflecting real-world healthcare provider decisions and patient behaviours without any interventions introduced. Furthermore, the retrospective cohort approach avoids the potential for behavioural biases associated with selection and active participation in research, the so-called ‘Hawthorne effect’.[8]

This retrospective, observational cohort study comprises patients who were identified from primary care using the UK Clinical Practice Research Datalink (CPRD-GOLD).

We aimed to evaluate records of episodes of moderate-to-severe exacerbations of COPD and to determine the factors associated with COPD exacerbation frequency in patients managed through primary care.

METHODS

Study design

We used a retrospective cohort study design and identified patients in the CPRD-GOLD who had a record of COPD diagnosis, defined as ≥1 record of COPD-specific READ codes from 1 April 2009 until 30 September 2012. We then took the first record of COPD diagnosis during this predefined period and assigned it as a cohort entry date (cohort baseline). We then further limited the patient cohort to only those with a minimum of 365 days of available history in the database prior to and subsequent to this first record (**figure 1**).

CPRD-GOLD database

The CPRD-GOLD, formerly known as the General Practice Research Database, is one of the largest computerised databases of linked anonymised primary care medical records in the world.[9] The CPRD-GOLD reflects the complete EMR for all NHS primary healthcare collected from approximately 8% of the population of England and Wales.[10] All medical signs, symptoms, investigations and diagnoses deemed important for the continued care of the patients are coded within the EMR.[10, 11]

The CPRD research group, a part of the UK Medicines and Healthcare Products Regulatory Agency, continuously monitors data quality and removes practices from the database if they fail to maintain the required standards of data entry.

Patient population

The cohort consisted of patients with a record of COPD diagnosis between 1 April 2009 and 30 September 2012 who sought care for their condition; see Study design. COPD diagnosis was required to be accompanied by spirometry recording consistent with obstructive disease, defined as a forced expiratory volume in 1 second (FEV_1)/forced vital capacity (FVC) ratio of <0.7 recorded any time before and up to 3 months from the date of diagnosis of COPD record that qualified the patient to enter this cohort.

Exacerbations of COPD

An algorithm was assembled to collect data on recorded events of moderate-to-severe COPD exacerbations from the 12 months preceding and the 12 months subsequent to cohort entry date. Moderate exacerbations were defined as a record of a diagnosis of exacerbation, acute bronchitis, or the management of COPD with specific antibiotics and

oral corticosteroids recorded on the same day or up to 5 days from each other. Admissions to hospital or visits to accident and emergency associated with COPD as recorded by GPs were considered as severe exacerbations. An exacerbation episode was defined based on events recorded over 2 weeks from the initial exacerbation record. The start of an episode was considered to be the date of the first qualifying event. Subsequently, a 14-day rolling window was applied to identify a period of at least 2 exacerbation-free weeks in order to ensure that a relapse was not categorised as a separate exacerbation episode.[12] The frequency of individual patients' exacerbations during the 12 months prior to the start of observation and also during the 12-month follow-up period was split into three categories: none, one, or two or more (frequent) episodes of moderate-to-severe exacerbation of COPD. The incidence of exacerbation episodes was also expressed as a rate per person per year, a sum of the episodes per patient divided by 365 days of follow-up. For a minority of patients, who would had been first diagnosed with COPD during the year prior to observation period start, we treated their exacerbation history similarly to those with established COPD.

Other key study variables

The following variables were also retrieved from the database using the latest record prior to cohort entry date: age, gender, smoking status, body mass index (BMI; or imputed BMI using the latest height and weight measurements where BMI data were not recorded), Medical Research Council (MRC) dyspnoea grade and FEV_1 percent predicted assessment with the closest date to cohort entry.[13]

Mean BMI was summarised as both a continuous and categorical variable using the World Health Organisation classification of underweight (<18.5), normal (18.5 to 24.9), overweight (25.0 to 29.9) and obese (≥30).[14]

The Global Initiative for Chronic Obstructive Lung Disease (GOLD) grade classification [15] was used to determine the severity of airflow limitation. Due to the nature of the data available from CPRD-GOLD, it was not possible to determine whether the spirometry data had only been recorded following the administration of a bronchodilator. However, given that the NHS Quality Outcomes Framework requires all COPD patients to have their COPD diagnosis confirmed with post-bronchodilator spirometry,[16] it is expected that the majority were recorded following bronchodilator administration. The categories of airflow severity limitation were defined using the FEV_1 cut-off points ≥80% predicted for mild, ≥50% to <80% predicted for moderate, ≥30% to <50% predicted for severe and >30% predicted for very severe airflow limitation.

MRC dyspnoea grade is a unidimensional measure of breathlessness related to activity.[17] The MRC dyspnoea scale is equivalent to the modified MRC (mMRC) scale, which is an established method in the published literature. The key difference with the mMRC dyspnoea scale is a shift in the scoring range: whereas MRC ranged from 1 to 5, mMRC employs a range from 0 to 4. Therefore, mMRC Grade 0 is equal to Grade 1 on the MRC scale.[18]

Records of diagnoses of selected comorbidities, including depression, anxiety and asthma, occurring prior to the start of the study were retrieved from the database. Data on the frequency of prescription treatment for COPD in the 12 months prior to the start of this cohort study were also collected. Patients who had one or more records for medication within a therapeutic class, with the exception of oral corticosteroids, were considered to use that medication. For oral corticosteroids, four or more prescriptions within the 12-month period were considered to indicate regular use rather than acute use for COPD exacerbation(s). Lastly, the number of recorded contacts with a GP for any reason in the 12 months before the cohort entry was ascertained and standardised per 365.25 days.

Statistical analyses

A multiple, multilevel, logistic regression was used to assess risk factors associated with frequency of exacerbations over the 12-month follow-up period: frequent vs none, frequent vs one, and one vs none. These analyses were carried out using the PROC LOGISTICS procedures in SAS version 9.2. All covariates listed in descriptive tables, with the exception of respiratory medications, were entered into the models. Missing values for BMI and percent predicted FEV_1 values (less than 2% of the cohort) and MRC dyspnoea score (12% of the cohort) were not imputed.

The relationships between the number of exacerbation episodes during the 12-month follow-up period and the percentage predicted FEV_1 value and MRC dyspnoea grade were assessed using Spearman's correlation coefficient.

The study protocol (WEUSKOP5903) was reviewed and approved as Protocol 12_118 by the CPRD Scientific and Ethics committee.

RESULTS

Patients

The cohort consisted of 58,589 patients with COPD (**figure 1**). Overall, 47% of patients were female, mean age was 69 years and mean FEV_1 was 60% predicted.

Incidence of exacerbation

During the 12-month follow-up period, 53% of patients (n=31 049) had no recorded episodes of moderate-to-severe exacerbation of COPD, 24% of patients (n=14 189) had one recorded exacerbation episode and frequent episodes were recorded in 23% of patients (n=13 351). The rate of moderate-to-severe exacerbations during the 12-month follow-up period was 0.89 per person per year. A maximum of eight episodes per patient was observed.

In comparison to patients with no recorded exacerbation episodes, patients with frequent exacerbations were more likely to be female (51% vs 44%), had poorer lung function (mean FEV_1 percent predicted: 54% vs 62%) and more dyspnoea (MRC grade ≥3: 58% vs 39%). When compared with patients with no recorded exacerbation episodes during the 12-month follow-up, patients with frequent exacerbations were more regularly diagnosed with comorbidities including myocardial infarction, heart failure and depression, and had been managed with maintenance treatment more often in the past 12 months (≥1 prescribed inhaled corticosteroid/long-acting beta agonist 68% vs 39%, ≥1 prescribed long-acting muscarinic antagonist 53% vs 28%) (**table 1**).

Table 1 Characteristics of COPD patients by frequency of recorded exacerbation episodes during 12-month follow-up

	Moderate-to-severe exacerbation frequency		
	None	**One**	**Two or more**
Patient characteristic	**n=31 049**	**n=14 189**	**n=13 351**
Age, mean (SD)	69.3 (10.5)	69.8 (10.3)	69.4 (10.2)
Smoking status			
Current	10 531 (34)	4667 (33)	4173 (31)
Former	15 988 (51)	7474 (53)	7345 (55)
Never	3312 (11)	1492 (11)	1301 (10)
Other	1218 (4)	556 (4)	532 (4)
Female	13 619 (44)	6831 (48)	6832 (51)
BMI underweight (<18.5)	1327 (4)	641 (5)	751 (6)
BMI normal (18.5–24.9)	10 533 (34)	4810 (34)	4668 (35)
BMI overweight (25.0–29.9)	10 633 (34)	4778 (34)	4274 (32)
BMI obese (≥30.0)	8037 (26)	3760 (26)	3472 (26)
Rate of moderate-to-severe exacerbations in prior 12 months per person per year, mean	0.37	0.74	1.67
Rate of exacerbations with hospital admission in prior 12 months per person per year, mean	0.07	0.12	0.2
Comorbidities any time in history			
Acute myocardial infarction	2530 (8)	1386 (10)	1350 (10)
Congestive heart disease	1857 (6)	1041 (7)	1165 (9)
Depression	3976 (13)	2116 (15)	2393 (18)
≥1 prescription within 12 months before cohort entry			
ICS	5789 (19)	2883 (20)	2501 (19)
LABA	2379 (8)	1278 (9)	1295 (10)

ICS / LABA	12 150 (39)	7328 (52)	9133 (68)
LAMA	8618 (28)	5260 (37)	7049 (53)
MRC Dyspnoea scale score			
MRC 1	6251 (23)	2059 (16)	1165 (10)
MRC 2	11 552 (42)	4765 (38)	3546 (31)
MRC 3	6478 (23)	3487 (28)	3286 (29)
MRC 4	2959 (11)	1876 (15)	2605 (23)
MRC 5	524 (2)	398 (3)	690 (6)
Airflow limitation grade			
mild	5475 (18)	1971 (14)	1398 (10)
moderate	16 815 (54)	7324 (52)	5693 (43)
severe	7013 (23)	3855 (27)	4629 (35)
very severe	1152 (4)	814 (6)	1411 (11)

All data are n (%) unless otherwise stated. Percentages expressed as column %.

BMI, body mass index; COPD, chronic obstructive pulmonary disease; ICS, inhaled corticosteroid; LABA, long-acting beta$_2$ agonist; LAMA, long-acting muscarinic antagonist; MRC, Medical Research Council; SD, standard deviation

Of the patients with no recorded exacerbation episodes during the 12-month follow-up period, 72% had no record of a moderate-to-severe exacerbation in the prior 12 months; 46% of patients who experienced frequent moderate-to-severe exacerbations during the follow-up period had also experienced two or more exacerbations in the prior 12 months (**figure 2**).

A weak, although statistically significant, relationship was observed between the frequency of COPD exacerbations and the stage of airflow limitation (Spearman's r=0.16, p<0.001). Thirty-four percent and 45% of patients with no recorded exacerbations also had poor lung function (moderate and severe airflow limitation, respectively), in contrast to 16% and 19% of patients with frequent exacerbation episodes who had mild or moderate airflow limitation, respectively (**figure 3a**).

Factors associated with exacerbation frequency

Comparison of patients with no exacerbation, one exacerbation and frequent exacerbations during the 12-month follow-up showed that several factors were associated with an increased likelihood of prospective frequent exacerbations; the strongest association was with prior exacerbations. Experiencing one moderate exacerbation, compared with none, during the 12 months preceding the start of the study was associated with one moderate-to-severe exacerbation occurring (odds ratio [OR] 1.9) or frequent (≥2) moderate-to-severe exacerbations occurring (OR 3.3) during the 12-month follow-up period (**table 2**). The OR for frequent (≥2) moderate-to-severe exacerbations increased to 13.6 if patients had experienced two or more moderate episodes, compared with none in the 12 months preceding the start of the study.

Table 2 Factors associated with moderate-to-severe exacerbations during the 12-month follow-up period [three-category multi-level model]

	Moderate-to-severe exacerbations during 12-month follow-up					
	≥2 vs none		**One vs none**		**≥2 vs one**	
Factor	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
COPD exacerbations, 12 months prior to observation period start						
No moderate event	Reference					
1 moderate episode	3.31	3.12 to 3.51	1.89	1.79 to 1.99	1.76	1.65 to 1.87
≥2 moderate episodes	13.64	12.67 to 14.68	3.11	2.88 to 3.37	4.41	4.11 to 4.74
No hospitalised episode	Reference					
≥1 hospitalised episode	2.13	1.95 to 2.31	1.49	1.38 to 1.61	1.44	1.32 to 1.57
Airflow limitation level nearest to observation period start						
Mild (FEV_1 ≥80% predicted)	Reference					
Moderate (≥50% FEV_1 <80%	1.23	1.13 to 1.33	1.19	1.12 to 1.27	1.04	0.95 to 1.14

predicted)						
Severe (≥30% FEV_1 <50% predicted)	1.81	1.66 to 1.98	1.36	1.27 to 1.47	1.37	1.24 to 1.50
Very severe (FEV_1 >30% predicted)	2.37	2.08 to 2.70	1.59	1.41 to 1.79	1.56	1.37 to 1.78
MRC dyspnoea scale nearest to observation period start						
1	Reference					
2	1.39	1.28 to 1.51	1.16	1.09 to 1.23	1.16	1.06 to 1.27
3	1.86	1.70 to 2.03	1.41	1.31 to 1.51	1.32	1.20 to 1.44
4	2.74	2.49 to 3.03	1.55	1.43 to 1.69	1.74	1.57 to 1.93
5	3.02	2.58 to 3.55	1.62	1.39 to 1.88	1.84	1.57 to 2.16
Comorbidities, history of medical diagnosis before observation period start (absence is referent category)						
Heart failure	1.2	1.08 to 1.32	1.1	1.00 to 1.20	1.11	1.00 to 1.23
Myocardial infarction	1.13	1.03 to 1.24	1.15	1.06 to 1.24	0.96	0.87 to 1.05
Rheumatological disease	1.12	1.01 to 1.24	1.1	1.01 to 1.21	1.04	0.94 to 1.16
Renal disease	0.9	0.84 to 0.97	0.96	0.91 to 1.02	0.93	0.86 to 1.00
Anxiety	1.16	1.08 to 1.25	1.14	1.08 to 1.22	1.02	0.95 to 1.09
Depression	1.25	1.16 to 1.35	1.1	1.02 to 1.17	1.12	1.04 to 1.21
Asthma	1.51	1.43 to 1.60	1.24	1.19 to 1.30	1.23	1.16 to 1.30
Cancer	1.28	1.19 to 1.38	1.1	1.03 to 1.17	1.14	1.05 to 1.23
Number of contacts with GP, 12 months prior to observation period start						
Low: 0–5	Reference					
Medium: 6–10	1.02	0.95 to 1.10	1.07	1.00 to 1.13	0.99	0.91 to 1.07
High: ≥10	1.24	1.16 to 1.32	1.18	1.12 to 1.24	1.08	1.01 to 1.16
Sex						
Female	1.19	1.13 to 1.26	1.12	1.07 to 1.17	1.07	1.01 to 1.14

Models further adjusted for age, smoking and BMI.

BMI, body mass index; CI, confidence interval; COPD, chronic obstructive pulmonary disease; FEV_1, forced expiratory volume in 1 second; MRC: Medical Research Council

The risk of frequent exacerbations compared with none, during the 12-month observation period, increased with greater severity of airway limitation; OR of 1.2 for moderate limitation (≥50% FEV_1 <80% predicted) and OR of 2.4 for very severe limitation (FEV_1 >30% predicted) (**table 2**). Increasing dyspnoea grade was also associated with an increased OR for frequent exacerbations compared with no exacerbations, from 1.4 for MRC2 to 3.0 for MRC5 when compared with MRC1 (**table 2**; **figure 3b**). The presence of specific comorbidities was associated with a significant increase in the risk of patients experiencing frequent exacerbations compared with none during the observational period (**table 2**). Among the specific comorbidities assessed, asthma, cancer and depression had the strongest observed associations with exacerbation frequency, with corresponding ORs for frequent vs no exacerbations of 1.5, 1.3 and 1.3, respectively (**table 2**). Similar risk factors for one vs no exacerbation, albeit of lower magnitude, were observed as for the risk of frequent vs no exacerbations.

DISCUSSION

We set out to explore COPD exacerbation frequency and associated factors, using a retrospective cohort approach, in a large primary care cohort of 58 589 patients in England and Wales. We found marked differences between patients managed in primary care who had frequent exacerbations (23% of patients) or one exacerbation (24%) as compared with patients with no recorded COPD exacerbations (53%) during the 12-month follow-up period.

The characteristic most strongly associated with moderate-to-severe exacerbation frequency was a history of exacerbations. A gradient relationship between moderate exacerbations in the prior 12 months and prospective moderate-to-severe exacerbations was apparent: patients who had no exacerbations in the prior 12 months had the lowest risk of future events; patients who had one exacerbation were at a greater risk of future episodes than those who had none, but were at a lower risk than those who had two or more prior exacerbations; patients with two or more prior exacerbations were at the highest risk of future exacerbations. Prior episodes of severe exacerbations also independently increased the odds of future exacerbations. This result is in agreement with previously reported findings.[7] When moderate and severe prior exacerbations are considered separately, our findings provide additional supportive evidence of an independent and 'dose-like' relationship of prior moderate events and future exacerbation risk.

Exacerbation frequency also increased with increasing grade of airflow limitation and, similarly, with increasing dyspnoea score. A relationship between exacerbation frequency and symptom severity is well established on the basis of findings from prospective studies;[19, 20] however, the two factors are semi-independent and patients with moderate airflow limitation can experience recurrent exacerbations.[7] Again, both airflow limitation and dyspnoea were associated with a higher risk for any exacerbation event, not only with frequent events. Likewise, the slightly increased probability of frequent exacerbations observed to be associated with females is consistent with previous reports.[19] No relationships between age or smoking status and exacerbation frequency were observed.

Finally, we observed a relationship between a prior diagnosis of selected comorbidities in patient history and a risk of any future moderate-to-severe exacerbations. Consistent independent associations were observed between a history of asthma, depression, anxiety, heart failure and cancer and either one or two or more exacerbation episodes. It has previously been reported that COPD exacerbations are associated with cardiovascular events, including myocardial infarction and heart failure,[21] and the presence of these co-morbidities increases the likelihood of mortality and costs associated with exacerbations.[2] Likewise, the importance of comorbid asthma as a risk factor for more frequent exacerbation

and accelerated disease progression has led to the description of an 'asthma-COPD overlap syndrome'.[22, 23] It has been estimated that overlap syndrome may be associated with up to a three-fold increase in the frequency and severity of exacerbations.[22] Overlap syndrome is typically seen in younger patients with a smoking burden that is below the average for COPD patients, potentially providing a partial explanation of the absence of observed relationships between age, smoking status and exacerbation frequency in this study.

Analyses of this large, primary care-based cohort supported the established hypothesis of the existence of a 'frequent exacerbator' phenotype, widely reported in the COPD literature as an explanation for observations of a subgroup of patients who experience highly recurrent exacerbations.[7, 24] The pathophysiology underlying the frequent exacerbator phenotype is not yet fully understood, and identification of the interrelated factors is potentially of value in aiding the identification of patients who are likely to be at an increased risk of further exacerbations. Early intervention is associated with faster recovery from exacerbation and improved health-related quality of life.[25]

The factors that we have observed to be associated with any future moderate-to-severe exacerbations are suggestive of a shared inflammatory biological mechanism that gives rise to further exacerbation events and also leads to worsening of COPD symptoms. Markers for airway and systemic inflammation have been shown to be associated with exacerbation frequency,[1] may precipitate lung function decline,[19] and may also partly explain the observed relationships between exacerbation frequency and comorbidities such as cardiovascular disease [21, 26] and asthma.[27].

We found that the characteristics of patients who experienced at least one exacerbation event during the follow-up period differed from those who did not have any exacerbations. This observation replicated earlier findings from a clinical cohort of 2138 patients enrolled in the Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints (ECLIPSE) study.[7] However, it also should be noted that patients who do not experience exacerbations and have mild airflow limitation are less likely to seek healthcare and are, therefore, less likely to be diagnosed with COPD.[28]

In this study, an inclusive definition of exacerbations was used, encompassing prescription records, as well as exacerbation event coding and GP records of hospital admissions and accident and emergency visits. Despite this, more than half of the cohort experienced no exacerbations during the 12-month observation period. It is possible that some exacerbations were un-reported or self-managed at home with rescue packs of antibiotics and/or oral corticosteroids. Although a spirometry recording was required for

eligibility, COPD patients in the present study were identified based on a record of their diagnosis from a primary care database rather than actively recruited when visiting a physician. Therefore, we cannot ensure that all patients met the definition of COPD as per the GOLD guidelines.[15] Further, we required all patients to be present in the database for at least a year post observation period entry. This could have biased our results by removing the most severe patients from the analysis.

A strength of this study is the exploration of exacerbation data from a large sample of patients diagnosed with COPD obtained from a well-established, comprehensive primary care EMR-based data resource, the CPRD-GOLD. This resource provides representative coverage of the patient population in England and Wales, and has incorporated spirometry and dyspnoea score data as a part of routine COPD patient disease management since 2004. The CPRD-GOLD has previously been used in COPD research to determine disease prevalence in the UK [29] and to investigate comorbidities in COPD and asthma.[30,31]

In conclusion, we demonstrated that, in a primary population of patients diagnosed with COPD, prior history of exacerbations (as well as increasing severity of airflow limitation and dyspnoea) and presence of comorbidities predict risk of future moderate-to-severe exacerbation events. We also showed that not only patients with prior frequent exacerbations, but also those with only one prior moderate exacerbation, are at increased risk of future exacerbation events. As exacerbations can accelerate the progression of the disease, our findings underline the importance of identifying COPD patients who are at risk of exacerbation and intervening as early as possible. Such identification is feasible for a primary care practitioner based on a review of clinical history routinely recorded in the UK electronic health record.

Collaborators Editorial support in the form of development of the manuscript first draft in consultation with the authors, editorial suggestions to draft versions of this paper, assembling tables and figures, collating author comments, copyediting, fact checking, referencing and graphic services was provided by Ian Grieve, PhD, at Gardiner-Caldwell Communications, an Ashfield company (Macclesfield, UK) and was funded by GlaxoSmithKline.

Competing interests HM, AS and AH are employees of and hold stock in GlaxoSmithKline. JQ received funding from an MRC Population Health Scientist Fellowship, grant number G0902135, and declares no competing interests.

Contributors HM wrote the study protocol and analysis plan, oversaw the data analysis and drafted the manuscript. JQ and AH contributed to the conception and design of the study and the interpretation of the data. AS contributed to the analysis and the interpretation of the data.

Funding This study was funded by GlaxoSmithKline R&D. The funder provided support in form of salaries for authors HM, AS, and AH, but did not have any additional role in the study design, data collection and analysis or preparation of the manuscript except that the study protocol was peer-reviewed by the GSK R&D Epidemiology Protocol Review Forum and the report was also peer-reviewed by the medical and epidemiology supervisors. The decision to submit for publication is required by GSK policy as part of being transparent regarding research activities. Editorial support in the form of development of the manuscript first draft in consultation with the authors, editorial suggestions to draft versions of this paper, assembling tables and figures, collating author comments, copyediting, fact checking, referencing and graphic services was provided by Ian Grieve, PhD, at Gardiner-Caldwell Communications, an Ashfield company (Macclesfield, UK) and was funded by GlaxoSmithKline.

Ethics approval The study protocol (WEUSKOP5903) was reviewed and approved as Protocol 12_118 by the CPRD Scientific and Ethics committee.

Data sharing statement No additional data available: access to the data used in this analysis, CPRD-GOLD database, is governed by the Ethics and Scientific committee: www.cprd.com/isac/. The authors cannot allow access to the CPRD-GOLD database because of contractual and ethics regulations imposed by the CPRD and its Ethics committee.

The analysis was conducted according to the study protocol. Additional statistical analyses included in the final report are available from the corresponding author.

REFERENCES

1. Donaldson GC, Seemungal TA, Patel IS, *et al*. Airway and systemic inflammation and decline in lung function in patients with COPD. *Chest* 2005;128(4):1995–2004.
2. Perera PN, Armstrong EP, Sherrill DL, *et al*. Acute exacerbations of COPD in the United States: inpatient burden and predictors of costs and mortality. *COPD* 2012;9(2):131–41.
3. Hurst JR, Donaldson GR, Quint JK, *et al*. Temporal Clustering of Exacerbations in Chronic Obstructive Pulmonary Disease. *Am J Respir Crit Care Med* 2009;179:369–74.
4. Vestbo J, Edwards LD, Scanlon PD, *et al*. Changes in forced expiratory volume in 1 second over time in COPD. *New Engl J Med* 2011;365(13):1184–92.
5. Aaron SD, Donaldson GC, Whitmore GA, *et al*. Time course and pattern of COPD exacerbation onset. *Thorax* 2012;67(3):238–43.
6. Leidy NK, Murray LT. Patient-reported outcome (PRO) measures for clinical trials of COPD: the EXACT and E-RS. *COPD* 2013;10(3):393–8.
7. Hurst JR, Vestbo J, Anzueto A, *et al*. Susceptibility to Exacerbation in Chronic Obstructive Pulmonary Disease. *New Engl J Med* 2010;363:1128–38.
8. McCambridge J, Witton J, Elbourne DR. Systematic review of the Hawthorne effect: New concepts are needed to study research participation effects. *J Clin Epidemiol* 2014;67:267–77.
9. Tate AR, Beloff N, Al-Radwan B, *et al*. Exploiting the potential of large databases of electronic health records for research using rapid search algorithms and an intuitive query interface. *J Am Med Inform Assoc* 2014;21:292–8.
10. Williams JEA, Green RH, Warrington V, *et al*. Development of the i-BODE: validation of the incremental shuttle walking test within the BODE index. *Respir Med* 2012;106(3):390–6.
11. Wood L, Martinez C. The general practice research database: role in pharmacovigilance. *Drug Safety* 2004;27:871–81.
12. Leidy NK, Wilcox TK, Jones PW, *et al*. Standardizing measurement of chronic obstructive pulmonary disease exacerbations. Reliability and validity of a patient-reported diary. *American J Respir Crit Care Med* 2011;183(3):323–9.

13. Bhaskaran K, Forbes HJ, Douglas I, *et al*. Representativeness and optimal use of body mass index (BMI) in the UK Clinical Practice Research Datalink (CPRD). *BMJ Open* 2013;3:e003389.

14. World Health Organisation. Obesity: preventing and managing the global epidemic. Report of a WHO Consultation. WHO Technical Report Series 894. Geneva: World Health Organisation, 2000. Available at: http://whqlibdoc.who.int/trs/WHO_TRS_894.pdf [last accessed July 2014].

15. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global Strategy for the Diagnosis, Management and Prevention of Chronic Obstructive Pulmonary Disease. Revised 2014. Available at: http://www.goldcopd.org/uploads/users/files/GOLD_Report_2014_Jun11.pdf [last accessed July 2014].

16. NICE. Quality and Outcomes Framework for COPD, NM63 Advisory Committee Recommendations. 2012. Available at: http://www.nice.org.uk/aboutnice/qof/indicators_detail.jsp?summary=13808 [last accessed July 2014].

17. Fletcher CM, Elmes PC, Fairbairn AS, *et al*. The significance of respiratory symptoms and the diagnosis of chronic bronchitis in a working population. *BMJ* 1959;2:257–66.

18. Williams T, van Staa T, Puri S, *et al*. Recent advances in the utility and use of the General Practice Research Database as an example of a UK Primary Care Data resource. *Ther Adv Drug Safety* 2012;3:89–99.

19. Burge S, Wedzicha JA. COPD exacerbations: definitions and classifications. *Eur Respir J* 2003;21(Suppl. 41):46s–53s.

20. Donaldson GC, Wedzicha JA. COPD exacerbations 1: epidemiology. *Thorax* 2006;61:164–8.

21. Donaldson GC, Hurst JR, Smith CJ, *et al*. Increased risk of myocardial infarction and stroke following exacerbation of COPD. *Chest* 2010;137(5):1091–7.

22. Hardin M, Silverman EK, Barr RG, *et al*. The clinical features of the overlap between COPD and asthma. *Respir Res* 2011;12:127.

23. De Marco R, Pesce G, Marcon A, *et al*. The Coexistence of Asthma and Chronic Obstructive Pulmonary Disease (COPD): Prevalence and Risk Factors in Young, Middle-aged and Elderly People from the General Population. *PLoS ONE* 2013;8(5):e62985.

24. Wedzicha JA, Brill SE, Allinson JP, *et al*. Mechanisms and impact of the frequent exacerbator phenotype in chronic obstructive pulmonary disease. *BMC Med* 2013;11:181.

25. Wilkinson TMA, Donaldson GC, Hurst JR, *et al*. Early Therapy Improves Outcomes of Exacerbations of Chronic Obstructive Pulmonary Disease. *Am J Respir Crit Care Med* 2004;169(12):1298–303.

26. Halpin DMG, Decramer M, Celli B, *et al*. Risk of nonlower respiratory serious adverse events following exacerbations in the 4-year UPLIFT trial. *Lung* 2011;189:261–8.

27. Zeki AA, Schivo M, Chan A, *et al*. The Asthma-COPD Overlap Syndrome: A Common Clinical Problem in the Elderly. *J Allergy* 2011:861926.

28. Bednarek M, Maciejewski J, Wozniak M, *et al*. Prevalence, severity and underdiagnosis of COPD in the primary care setting. *Thorax* 2008;63:402–7.

29. Soriano JB, Maier WC, Egger P, *et al*. Recent trends in physician diagnosed COPD in women and men in the UK. *Thorax* 2000;55:789–94.

30. Soriano JB, Visick GT, Muellerova H, *et al*. Patterns of co-morbidities in newly diagnosed COPD and asthma in primary care. *Chest* 2005;128(4):2099–107.

31. García Rodríguez LA, Wallander M-A, Martín-Merino E, *et al*. Heart failure, myocardial infarction, lung cancer and death in COPD patients: A UK primary care study. *Respir Med* 2010;104(11):1691–9.

FIGURE LEGENDS

Figure 1 Patient recruitment flow chart.

Figure 2 Frequency of moderate-to-severe COPD exacerbations in the 12 months prior to cohort entry in relation to the frequency of exacerbation episodes during the 12-month follow-up period.

Figure 3 COPD exacerbation frequency during the 12-month follow-up period by a) airflow limitation stage, b) MRC dyspnoea grade.

Risk factors for acute exacerbations in primary care COPD patients: a retrospective observational cohort study

Hana Müllerová,[1] Amit Shukla,[1] Adam Hawkins,[2] Jennifer Quint[3]

[1]Worldwide Epidemiology, GlaxoSmithKline, Uxbridge, UK

[2]Global Respiratory Franchise, GlaxoSmithKline, Uxbridge, UK

[3]London School of Hygiene and Tropical Medicine, London, UK

Correspondence to

Hana Müllerová; Worldwide Epidemiology, GlaxoSmithKline R&D, Building 9, Iron Bridge Road, Stockley Park West, Uxbridge, Middlesex, UB11 1BT, UK

Tel: +44 208 990 2647. Fax: +44 208 990 3505

hana.x.muellerova@gsk.com

Keywords

Chronic obstructive pulmonary disease, database, electronic medical records, exacerbations, primary care

Word count

(excluding title page, abstract, references, figures and tables): 3,127 words

ABSTRACT

Objectives: To evaluate risk factors associated with exacerbation frequency in primary care. Information on exacerbations of chronic obstructive pulmonary disease (COPD) has mainly been generated by secondary care-based clinical cohorts.

Design: Retrospective observational cohort study.

Setting: Electronic medical records database (England and Wales).

Participants: 58 589 COPD patients aged ≥40 years with COPD diagnosis recorded between 1 April 2009 and 30 September 2012, and with at least 365 days of follow-up before and after the COPD diagnosis, were identified in the Clinical Practice Research Datalink. Mean age 69 years; 47% female; mean forced expiratory volume in 1 second (FEV_1) 60% predicted.

Outcome measures: Data on moderate or severe exacerbation episodes defined by diagnosis and/or medication codes 12 months following cohort entry were retrieved, together with demographic and clinical characteristics. Associations between patient characteristics and odds of having none vs one, none vs frequent (≥2) and one vs frequent exacerbations over 12 months follow-up were evaluated using multivariate logistic regression models.

Results: During follow-up, 23% of patients had evidence of frequent moderate-to-severe COPD exacerbations (24% one; 53% none). Independent predictors of increased odds of having exacerbations during the follow-up, either frequent episodes or one episode, included prior exacerbations, increasing dyspnoea score, increasing grade of airflow limitation, females and prior or current history of several comorbidities (e.g. asthma, depression, anxiety, heart failure and cancer).

Conclusions: Primary care-managed COPD patients at the highest risk of exacerbations can be identified by exploring past medical history for presence of prior exacerbations, greater COPD disease severity and co-occurrence of other medical conditions.

Funding: GlaxoSmithKline (WEUSKOP5903)

ARTICLE SUMMARY

Strengths and limitations of the study

- We report the results of a retrospective cohort study of 58,589 patients identified in the UK Clinical Practice Research Datalink (CPRD) with spirometry-confirmed COPD diagnosis during a 42-month period from 1 April 2009, with the primary objective of evaluating risk factors for acute exacerbations of COPD.
- Strengths of the study include: the large size of the cohort; the cohort is representative of the UK primary care population; the utilisation of the CPRD, a comprehensive and well-established electronic medical records database which incorporates spirometric scores and dyspnoea grades recorded as part of the Quality Outcomes Framework.
- The main limitation of this study is that the event recording of acute exacerbations in the CPRD is not entirely standardized , necessitating the use of an algorithm containing prescription records for oral corticosteroid and antibiotics as well as primary care visits for exacerbations, and A&E and hospital admissions for COPD to maximise the detection of exacerbations.
- Another possible study limitation is that the true population of patients with milder airflow obstruction and those who did not experience exacerbations may be underrepresented in the dataset due to such patients not seeking healthcare or undergoing spirometry.

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is characterised by the progressive deterioration of lung function. In many patients, the course of the disease is punctuated by exacerbations, an acute worsening of symptoms which, in severe cases, can necessitate hospitalisation and even result in death.[1] Exacerbations are the major cost driver, directly and indirectly, in COPD.[2] COPD exacerbations are known to cluster in time and patients who have experienced one episode are at an increased risk of a recurrent exacerbation.[3] Furthermore, exacerbations accelerate the deterioration of lung function,[4] which in turn increases the likelihood and severity of further exacerbations.

Various definitions are used to mark the start, end, relapse and recurrence of exacerbations.[5, 6] Frequent exacerbations, usually defined as two or more episodes per year, are the focus of the majority of clinical epidemiology studies, as they are considered a marker of greater disease burden and reduced survival.[7] However, an increase in frequency of any moderate-to-severe exacerbation events has been associated with greater disease severity, and the best predictor of future events is a history of exacerbation(s).[7]

Previous research into the natural history of COPD exacerbations has mainly focused on prospective cohorts of patients who have usually been recruited from secondary and/or tertiary care. Limited data are available on the burden of COPD exacerbations in patients who are managed in primary care. The use of electronic medical record (EMR) databases allow for an identification of very large-scale patient cohorts in a setting much more representative of clinical practice than prospective studies carried out in selective cohorts. EMR-based cohorts represent all patients diagnosed and treated in primary care, with information reflecting real-world healthcare provider decisions and patient behaviours without any interventions introduced. Furthermore, the retrospective cohort approach avoids the potential for behavioural biases associated with selection and active participation in research, the so-called ‘Hawthorne effect’.[8]

This retrospective, observational cohort study comprises patients who were identified from primary care using the UK Clinical Practice Research Datalink (CPRD-GOLD).

We aimed to evaluate records of episodes of moderate-to-severe exacerbations of COPD and to determine the factors associated with COPD exacerbation frequency in patients managed through primary care.

METHODS

Study design

We used a retrospective cohort study design and identified patients in the CPRD-GOLD who had a record of COPD diagnosis, defined as ≥1 record of COPD-specific READ codes from 1 April 2009 until 30 September 2012. We then took the first record of COPD diagnosis during this predefined period and assigned it as a cohort entry date (cohort baseline). We then further limited the patient cohort to only those with a minimum of 365 days of available history in the database prior to and subsequent to this first record (**figure 1**).

CPRD-GOLD database

The CPRD-GOLD, formerly known as the General Practice Research Database, is one of the largest computerised databases of linked anonymised primary care medical records in the world.[9] The CPRD-GOLD reflects the complete EMR for all NHS primary healthcare collected from approximately 8% of the population of England and Wales.[10] All medical signs, symptoms, investigations and diagnoses deemed important for the continued care of the patients are coded within the EMR.[10, 11]

The CPRD research group, a part of the UK Medicines and Healthcare Products Regulatory Agency, continuously monitors data quality and removes practices from the database if they fail to maintain the required standards of data entry.

Patient population

The cohort consisted of patients with a record of COPD diagnosis between 1 April 2009 and 30 September 2012 who sought care for their condition; see Study design. COPD diagnosis was required to be accompanied ~~confirmed~~ by spirometry recording consistent with obstructive disease, defined as a forced expiratory volume in 1 second (FEV_1)/forced vital capacity (FVC) ratio of <0.7 recorded any time before and up to 3 months from the date of diagnosis of COPD record that qualified the patient to enter this cohort.

Exacerbations of COPD

An algorithm was assembled to collect data on recorded events of moderate-to-severe COPD exacerbations from the 12 months preceding and the 12 months subsequent to cohort entry date. Moderate exacerbations were defined as a record of a diagnosis of exacerbation, acute bronchitis, or the management of COPD with specific antibiotics and

oral corticosteroids recorded on the same day or up to~~within~~ 5 days from~~of~~ each other. Admissions to hospital or visits to accident and emergency associated with COPD as recorded by GPs were considered as severe exacerbations. An exacerbation episode was defined based on events recorded over 2 weeks from the initial exacerbation record. The start of an episode was considered to be the date of the first qualifying event. Subsequently, a 14-day rolling window was applied to identify a period of at least 2 exacerbation-free weeks in order to ensure that a relapse was not categorised as a separate exacerbation episode.[12] The frequency of individual patients' exacerbations during the 12 months prior to the start of observation and also during the 12-month follow-up period was split into three categories: none, one, or two or more (frequent) episodes of moderate-to-severe exacerbation of COPD. The incidence of exacerbation episodes was also expressed as a rate per person per year, a sum of the episodes per patient divided by 365 days of follow-up. For a minority of patients, who would had been first diagnosed with COPD during the year prior to observation period start, we treated their exacerbation history similarly to those with established COPD.

Other key study variables

The following variables were also retrieved from the database using the latest record prior to cohort entry date: age, gender, smoking status, body mass index (BMI; or imputed BMI using the latest height and weight measurements where BMI data were not recorded), Medical Research Council (MRC) dyspnoea grade and FEV_1 percent predicted assessment with the closest date to cohort entry.[13]

Mean BMI was summarised as both a continuous and categorical variable using the World Health Organisation classification of underweight (<18.5), normal (18.5 to 24.9), overweight (25.0 to 29.9) and obese (≥30).[14]

The Global Initiative for Chronic Obstructive Lung Disease (GOLD) grade classification [15] was used to determine the severity of airflow limitation. Due to the nature of the data available from CPRD-GOLD, it was not possible to determine whether the spirometry data had only been recorded following the administration of a bronchodilator. However, given that the NHS Quality Outcomes Framework requires all COPD patients to have their COPD diagnosis confirmed with post-bronchodilator spirometry,[16] it is expected that the majority were recorded following bronchodilator administration. The categories of airflow severity limitation were defined using the FEV_1 cut-off points ≥80% predicted for mild, ≥50% to <80% predicted for moderate, ≥30% to <50% predicted for severe and >30% predicted for very severe airflow limitation.

MRC dyspnoea grade is a unidimensional measure of breathlessness related to activity.[17] The MRC dyspnoea scale is equivalent to the modified MRC (mMRC) scale, which is an established method in the published literature. The key difference with the mMRC dyspnoea scale is a shift in the scoring range: whereas MRC ranged from 1 to 5, mMRC employs a range from 0 to 4. Therefore, mMRC Grade 0 is equal to Grade 1 on the MRC scale.[18]

Records of diagnoses of selected comorbidities, including depression, anxiety and asthma, occurring prior to the start of the study were retrieved from the database. Data on the frequency of prescription treatment for COPD in the 12 months prior to the start of this cohort study were also collected. Patients who had one or more records for medication within a therapeutic class, with the exception of oral corticosteroids, were considered to use that medication. For oral corticosteroids, four or more prescriptions within the 12-month period were considered to indicate regular use rather than acute use for COPD exacerbation(s). Lastly, the number of recorded contacts with a GP for any reason in the 12 months before the cohort entry was ascertained and standardised per 365.25 days.

Statistical analyses

A multiple, multilevel, logistic regression was used to assess risk factors associated with frequency of exacerbations over the 12-month follow-up period: frequent vs none, frequent vs one, and one vs none. These analyses were carried out using the PROC LOGISTICS procedures in SAS version 9.2. All covariates listed in descriptive tables, with the exception of respiratory medications, were entered into the models. Missing values for BMI and percent predicted FEV_1 values (less than 2% of the cohort) and MRC dyspnoea score (12% of the cohort) were not imputed.

The relationships between the number of exacerbation episodes during the 12-month follow-up period and the percentage predicted FEV_1 value and MRC dyspnoea grade were assessed using Spearman's correlation coefficient.

The study protocol (WEUSKOP5903) was reviewed and approved as Protocol 12_118 by the CPRD Scientific and Ethics committee.

RESULTS

Patients

The cohort consisted of 58,589 patients with COPD (**figure 1**). Overall, 47% of patients were female, mean age was 69 years and mean FEV_1 was 60% predicted.

Incidence of exacerbation

During the 12-month follow-up period, 53% of patients (n=31 049) had no recorded episodes of moderate-to-severe exacerbation of COPD, 24% of patients (n=14 189) had one recorded exacerbation episode and frequent episodes were recorded in 23% of patients (n=13 351). The rate of moderate-to-severe exacerbations during the 12-month follow-up period was 0.89 per person per year. A maximum of eight episodes per patient was observed.

In comparison to patients with no recorded exacerbation episodes, patients with frequent exacerbations were more likely to be female (51% vs 44%), had poorer lung function (mean FEV_1 percent predicted: 54% vs 62%) and more dyspnoea (MRC grade ≥3: 58% vs 39%). When compared with patients with no recorded exacerbation episodes during the 12-month follow-up, patients with frequent exacerbations were more regularly diagnosed with comorbidities including myocardial infarction, heart failure and depression, and had been managed with maintenance treatment more often in the past 12 months (≥1 prescribed inhaled corticosteroid/long-acting beta agonist 68% vs 39%, ≥1 prescribed long-acting muscarinic antagonist 53% vs 28%) (**table 1**).

Table 1 Characteristics of COPD patients by frequency of recorded exacerbation episodes during 12-month follow-up

	Moderate-to-severe exacerbation frequency		
	None	**One**	**Two or more**
Patient characteristic	**n=31 049**	**n=14 189**	**n=13 351**
Age, mean (SD)	69.3 (10.5)	69.8 (10.3)	69.4 (10.2)
Smoking status			
Current	10 531 (34)	4667 (33)	4173 (31)
Former	15 988 (51)	7474 (53)	7345 (55)
Never	3312 (11)	1492 (11)	1301 (10)
Other	1218 (4)	556 (4)	532 (4)
Female	13 619 (44)	6831 (48)	6832 (51)
BMI underweight (<18.5)	1327 (4)	641 (5)	751 (6)
BMI normal (18.5–24.9)	10 533 (34)	4810 (34)	4668 (35)
BMI overweight (25.0–29.9)	10 633 (34)	4778 (34)	4274 (32)
BMI obese (≥30.0)	8037 (26)	3760 (26)	3472 (26)
Rate of moderate-to-severe exacerbations in prior 12 months per person per year, mean	0.37	0.74	1.67
Rate of exacerbations with hospital admission in prior 12 months per person per year, mean	0.07	0.12	0.2
Comorbidities any time in history			
Acute myocardial infarction	2530 (8)	1386 (10)	1350 (10)
Congestive heart disease	1857 (6)	1041 (7)	1165 (9)
Depression	3976 (13)	2116 (15)	2393 (18)
≥1 prescription within 12 months before cohort entry			
ICS	5789 (19)	2883 (20)	2501 (19)
LABA	2379 (8)	1278 (9)	1295 (10)

ICS / LABA	12 150 (39)	7328 (52)	9133 (68)
LAMA	8618 (28)	5260 (37)	7049 (53)
MRC Dyspnoea scale score			
MRC 1	6251 (23)	2059 (16)	1165 (10)
MRC 2	11 552 (42)	4765 (38)	3546 (31)
MRC 3	6478 (23)	3487 (28)	3286 (29)
MRC 4	2959 (11)	1876 (15)	2605 (23)
MRC 5	524 (2)	398 (3)	690 (6)
Airflow limitation grade			
mild	5475 (18)	1971 (14)	1398 (10)
moderate	16 815 (54)	7324 (52)	5693 (43)
severe	7013 (23)	3855 (27)	4629 (35)
very severe	1152 (4)	814 (6)	1411 (11)

All data are n (%) unless otherwise stated. Percentages expressed as column %.

BMI, body mass index; COPD, chronic obstructive pulmonary disease; ICS, inhaled corticosteroid; LABA, long-acting beta$_2$ agonist; LAMA, long-acting muscarinic antagonist; MRC, Medical Research Council; SD, standard deviation

Of the patients with no recorded exacerbation episodes during the 12-month follow-up period, 72% had no record of a moderate-to-severe exacerbation in the prior 12 months; 46% of patients who experienced frequent moderate-to-severe exacerbations during the follow-up period had also experienced two or more exacerbations in the prior 12 months (**figure 2**).

A weak, although statistically significant, relationship was observed between the frequency of COPD exacerbations and the stage of airflow limitation (Spearman's r=0.16, p<0.001). Thirty-four percent and 45% of patients with no recorded exacerbations also had poor lung function (moderate and severe airflow limitation, respectively), in contrast to 16% and 19% of patients with frequent exacerbation episodes who had mild or moderate airflow limitation, respectively (**figure 3a**).

Factors associated with exacerbation frequency

Comparison of patients with no exacerbation, one exacerbation and frequent exacerbations during the 12-month follow-up showed that several factors were associated with an increased likelihood of prospective frequent exacerbations; the strongest association was with prior exacerbations. Experiencing one moderate exacerbation, compared with none, during the 12 months preceding the start of the study was associated with one moderate-to-severe exacerbation occurring (odds ratio [OR] 1.9) or frequent (≥2) moderate-to-severe exacerbations occurring (OR 3.3) during the 12-month follow-up period (**table 2**). The OR for frequent (≥2) moderate-to-severe exacerbations increased to 13.6 if patients had experienced two or more moderate episodes, compared with none in the 12 months preceding the start of the study.

Table 2 Factors associated with moderate-to-severe exacerbations during the 12-month follow-up period [three-category multi-level model]

	Moderate-to-severe exacerbations during 12-month follow-up					
	≥2 vs none		**One vs none**		**≥2 vs one**	
Factor	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
COPD exacerbations, 12 months prior to observation period start						
No moderate event	Reference					
1 moderate episode	3.31	3.12 to 3.51	1.89	1.79 to 1.99	1.76	1.65 to 1.87
≥2 moderate episodes	13.64	12.67 to 14.68	3.11	2.88 to 3.37	4.41	4.11 to 4.74
No hospitalised episode	Reference					
≥1 hospitalised episode	2.13	1.95 to 2.31	1.49	1.38 to 1.61	1.44	1.32 to 1.57
Airflow limitation level nearest to observation period start						
Mild (FEV_1 ≥80% predicted)	Reference					
Moderate (≥50% FEV_1 <80%	1.23	1.13 to 1.33	1.19	1.12 to 1.27	1.04	0.95 to 1.14

predicted)						
Severe (≥30% FEV_1 <50% predicted)	1.81	1.66 to 1.98	1.36	1.27 to 1.47	1.37	1.24 to 1.50
Very severe (FEV_1 >30% predicted)	2.37	2.08 to 2.70	1.59	1.41 to 1.79	1.56	1.37 to 1.78
MRC dyspnoea scale nearest to observation period start						
1	Reference					
2	1.39	1.28 to 1.51	1.16	1.09 to 1.23	1.16	1.06 to 1.27
3	1.86	1.70 to 2.03	1.41	1.31 to 1.51	1.32	1.20 to 1.44
4	2.74	2.49 to 3.03	1.55	1.43 to 1.69	1.74	1.57 to 1.93
5	3.02	2.58 to 3.55	1.62	1.39 to 1.88	1.84	1.57 to 2.16
Comorbidities, history of medical diagnosis before observation period start (absence is referent category)						
Heart failure	1.2	1.08 to 1.32	1.1	1.00 to 1.20	1.11	1.00 to 1.23
Myocardial infarction	1.13	1.03 to 1.24	1.15	1.06 to 1.24	0.96	0.87 to 1.05
Rheumatological disease	1.12	1.01 to 1.24	1.1	1.01 to 1.21	1.04	0.94 to 1.16
Renal disease	0.9	0.84 to 0.97	0.96	0.91 to 1.02	0.93	0.86 to 1.00
Anxiety	1.16	1.08 to 1.25	1.14	1.08 to 1.22	1.02	0.95 to 1.09
Depression	1.25	1.16 to 1.35	1.1	1.02 to 1.17	1.12	1.04 to 1.21
Asthma	1.51	1.43 to 1.60	1.24	1.19 to 1.30	1.23	1.16 to 1.30
Cancer	1.28	1.19 to 1.38	1.1	1.03 to 1.17	1.14	1.05 to 1.23
Number of contacts with GP, 12 months prior to observation period start						
Low: 0–5	Reference					
Medium: 6–10	1.02	0.95 to 1.10	1.07	1.00 to 1.13	0.99	0.91 to 1.07
High: ≥10	1.24	1.16 to 1.32	1.18	1.12 to 1.24	1.08	1.01 to 1.16
Sex						
Female	1.19	1.13 to 1.26	1.12	1.07 to 1.17	1.07	1.01 to 1.14

Models further adjusted for age, smoking and BMI.

BMI, body mass index; CI, confidence interval; COPD, chronic obstructive pulmonary disease; FEV_1, forced expiratory volume in 1 second; MRC: Medical Research Council

The risk of frequent exacerbations compared with none, during the 12-month observation period, increased with greater severity of airway limitation; OR of 1.2 for moderate limitation (≥50% FEV_1 <80% predicted) and OR of 2.4 for very severe limitation (FEV_1 >30% predicted) (**table 2**). Increasing dyspnoea grade was also associated with an increased OR for frequent exacerbations compared with no exacerbations, from 1.4 for MRC2 to 3.0 for MRC5 when compared with MRC1 (**table 2**; **figure 3b**). The presence of specific comorbidities was associated with a significant increase in the risk of patients experiencing frequent exacerbations compared with none during the observational period (**table 2**). Among the specific comorbidities assessed, asthma, cancer and depression had the strongest observed associations with exacerbation frequency, with corresponding ORs for frequent vs no exacerbations of 1.5, 1.3 and 1.3, respectively (**table 2**). Similar risk factors for one vs no exacerbation, albeit of lower magnitude, were observed as for the risk of frequent vs no exacerbations.

DISCUSSION

We set out to explore COPD exacerbation frequency and associated factors, using a retrospective cohort approach, in a large primary care cohort of 58 589 patients in England and Wales. We found marked differences between patients managed in primary care who had frequent exacerbations (23% of patients) or one exacerbation (24%) as compared with patients with no recorded COPD exacerbations (53%) during the 12-month follow-up period.

The characteristic most strongly associated with moderate-to-severe exacerbation frequency was a history of exacerbations. A gradient relationship between moderate exacerbations in the prior 12 months and prospective moderate-to-severe exacerbations was apparent: patients who had no exacerbations in the prior 12 months had the lowest risk of future events; patients who had one exacerbation were at a greater risk of future episodes than those who had none, but were at a lower risk than those who had two or more prior exacerbations; patients with two or more prior exacerbations were at the highest risk of future exacerbations. Prior episodes of severe exacerbations also independently increased the odds of future exacerbations. This result is in agreement with previously reported findings.[7] When moderate and severe prior exacerbations are considered separately, our findings provide additional supportive evidence of an independent and 'dose-like' relationship of prior moderate events and future exacerbation risk.

Exacerbation frequency also increased with increasing grade of airflow limitation and, similarly, with increasing dyspnoea score. A relationship between exacerbation frequency and symptom severity is well established on the basis of findings from prospective studies;[19, 20] however, the two factors are semi-independent and patients with moderate airflow limitation can experience recurrent exacerbations.[7] Again, both airflow limitation and dyspnoea were associated with a higher risk for any exacerbation event, not only with frequent events. Likewise, the slightly increased probability of frequent exacerbations observed to be associated with females is consistent with previous reports.[19] No relationships between age or smoking status and exacerbation frequency were observed.

Finally, we observed a relationship between a prior diagnosis of selected comorbidities in patient history and a risk of any future moderate-to-severe exacerbations. Consistent independent associations were observed between a history of asthma, depression, anxiety, heart failure and cancer and either one or two or more exacerbation episodes. It has previously been reported that COPD exacerbations are associated with cardiovascular events, including myocardial infarction and heart failure,[21] and the presence of these co-morbidities increases the likelihood of mortality and costs associated with exacerbations.[2] Likewise, the importance of comorbid asthma as a risk factor for more frequent exacerbation

and accelerated disease progression has led to the description of an 'asthma-COPD overlap syndrome'.[22, 23] It has been estimated that overlap syndrome may be associated with up to a three-fold increase in the frequency and severity of exacerbations.[22] Overlap syndrome is typically seen in younger patients with a smoking burden that is below the average for COPD patients, potentially providing a partial explanation of the absence of observed relationships between age, smoking status and exacerbation frequency in this study.

Analyses of this large, primary care-based cohort supported the established hypothesis of the existence of a 'frequent exacerbator' phenotype, widely reported in the COPD literature as an explanation for observations of a subgroup of patients who experience highly recurrent exacerbations.[7, 24] The pathophysiology underlying the frequent exacerbator phenotype is not yet fully understood, and identification of the interrelated factors is potentially of value in aiding the identification of patients who are likely to be at an increased risk of further exacerbations. Early intervention is associated with faster recovery from exacerbation and improved health-related quality of life.[25]

The factors that we have observed to be associated with any future moderate-to-severe exacerbations are suggestive of a shared inflammatory biological mechanism that gives rise to further exacerbation events and also leads to worsening of COPD symptoms. Markers for airway and systemic inflammation have been shown to be associated with exacerbation frequency,[1] may precipitate lung function decline,[19] and may also partly explain the observed relationships between exacerbation frequency and comorbidities such as cardiovascular disease [21, 26] and asthma.[27]. ~~Recent studies have identified a number of potential susceptibility factors for exacerbation of COPD, including increased levels of inflammatory markers such as fibrinogen.[28, 29] Furthermore, dynamic lung hyperinflation correlates strongly with the patient-reported symptom of dyspnoea and is a consequence of the acute expiratory flow limitation that occurs during exacerbation.[30, 31] In combination with the heightened inflammation, this deterioration in airway function may increase the likelihood and severity of future exacerbation events, as reflected in the current GOLD guidelines.[15]~~

We found that the characteristics of patients who experienced at least one exacerbation event during the follow-up period differed from those who did not have any exacerbations. This observation replicated earlier findings from a clinical cohort of 2138 patients enrolled in the Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints (ECLIPSE) study.[7] However, it also should be noted that patients who do not experience exacerbations and have mild airflow limitation are less likely to seek healthcare and are, therefore, less likely to be diagnosed with COPD.[28~~32~~]

In this study, an inclusive definition of exacerbations was used, encompassing prescription records, as well as exacerbation event coding and GP records of hospital admissions and accident and emergency visits. Despite this, more than half of the cohort experienced no exacerbations during the 12-month observation period. It is possible that some exacerbations were un-reported or self-managed at home with rescue packs of antibiotics and/or oral corticosteroids. Although a spirometry recording was required for eligibility, COPD patients in the present study were identified based on a record of their diagnosis from a primary care database rather than actively recruited when visiting a physician. Therefore, we cannot ensure that all patients met the definition of COPD as per the GOLD guidelines.[15] Further, we required all patients to be present in the database for at least a year post observation period entry. This could have biased our results by removing the most severe patients from the analysis.

A strength of this study is the exploration of exacerbation data from a large sample of patients diagnosed with COPD obtained from a well-established, comprehensive primary care EMR-based data resource, the CPRD-GOLD. This resource provides representative coverage of the patient population in England and Wales, and has incorporated spirometry and dyspnoea score data as a part of routine COPD patient disease management since 2004. The CPRD-GOLD has previously been used in COPD research to determine disease prevalence in the UK [29~~33~~] and to investigate comorbidities in COPD and asthma.[30,31~~34, 35~~]

In conclusion, we demonstrated that, in a primary population of patients diagnosed with COPD, prior history of exacerbations (as well as increasing severity of airflow limitation and dyspnoea) and presence of comorbidities predict risk of future moderate-to-severe exacerbation events. We also showed that not only patients with prior frequent exacerbations, but also those with only one prior moderate exacerbation, are at increased risk of future exacerbation events. As exacerbations can accelerate the progression of the disease, our findings underline the importance of identifying COPD patients who are at risk of exacerbation and intervening as early as possible. Such identification is feasible for a primary care practitioner based on a review of clinical history routinely recorded in the UK electronic health record.

Collaborators Editorial support in the form of development of the manuscript first draft in consultation with the authors, editorial suggestions to draft versions of this paper, assembling tables and figures, collating author comments, copyediting, fact checking, referencing and graphic services was provided by Ian Grieve, PhD, at Gardiner-Caldwell Communications, an Ashfield company (Macclesfield, UK) and was funded by GlaxoSmithKline.

Competing interests HM, AS and AH are employees of and hold stock in GlaxoSmithKline. JQ received funding from an MRC Population Health Scientist Fellowship, grant number G0902135, and declares no competing interests.

Contributors HM wrote the study protocol and analysis plan, oversaw the data analysis and drafted the manuscript. JQ and AH contributed to the conception and design of the study and the interpretation of the data. AS contributed to the analysis and the interpretation of the data.

Funding This work was supported by GlaxoSmithKline (study number WEUSKOP5903).

Ethics approval The study protocol (WEUSKOP5903) was reviewed and approved as Protocol 12_118 by the CPRD Scientific and Ethics committee.

Data sharing statement No additional data available: access to the data used in this analysis, CPRD-GOLD database, is governed by the Ethics and Scientific committee: www.cprd.com/isac/. The authors cannot allow access to the CPRD-GOLD database because of contractual and ethics regulations imposed by the CPRD and its Ethics committee.

The analysis was conducted according to the study protocol. Additional statistical analyses included in the final report are available from the corresponding author.

REFERENCES

1. Donaldson GC, Seemungal TA, Patel IS, *et al*. Airway and systemic inflammation and decline in lung function in patients with COPD. *Chest* 2005;128(4):1995–2004.
2. Perera PN, Armstrong EP, Sherrill DL, *et al*. Acute exacerbations of COPD in the United States: inpatient burden and predictors of costs and mortality. *COPD* 2012;9(2):131–41.
3. Hurst JR, Donaldson GR, Quint JK, *et al*. Temporal Clustering of Exacerbations in Chronic Obstructive Pulmonary Disease. *Am J Respir Crit Care Med* 2009;179:369–74.
4. Vestbo J, Edwards LD, Scanlon PD, *et al*. Changes in forced expiratory volume in 1 second over time in COPD. *New Engl J Med* 2011;365(13):1184–92.
5. Aaron SD, Donaldson GC, Whitmore GA, *et al*. Time course and pattern of COPD exacerbation onset. *Thorax* 2012;67(3):238–43.
6. Leidy NK, Murray LT. Patient-reported outcome (PRO) measures for clinical trials of COPD: the EXACT and E-RS. *COPD* 2013;10(3):393–8.
7. Hurst JR, Vestbo J, Anzueto A, *et al*. Susceptibility to Exacerbation in Chronic Obstructive Pulmonary Disease. *New Engl J Med* 2010;363:1128–38.
8. McCambridge J, Witton J, Elbourne DR. Systematic review of the Hawthorne effect: New concepts are needed to study research participation effects. *J Clin Epidemiol* 2014;67:267–77.
9. Tate AR, Beloff N, Al-Radwan B, *et al*. Exploiting the potential of large databases of electronic health records for research using rapid search algorithms and an intuitive query interface. *J Am Med Inform Assoc* 2014;21:292–8.
10. Williams JEA, Green RH, Warrington V, *et al*. Development of the i-BODE: validation of the incremental shuttle walking test within the BODE index. *Respir Med* 2012;106(3):390–6.
11. Wood L, Martinez C. The general practice research database: role in pharmacovigilance. *Drug Safety* 2004;27:871–81.
12. Leidy NK, Wilcox TK, Jones PW, *et al*. Standardizing measurement of chronic obstructive pulmonary disease exacerbations. Reliability and validity of a patient-reported diary. *American J Respir Crit Care Med* 2011;183(3):323–9.

13. Bhaskaran K, Forbes HJ, Douglas I, *et al*. Representativeness and optimal use of body mass index (BMI) in the UK Clinical Practice Research Datalink (CPRD). *BMJ Open* 2013;3:e003389.

14. World Health Organisation. Obesity: preventing and managing the global epidemic. Report of a WHO Consultation. WHO Technical Report Series 894. Geneva: World Health Organisation, 2000. Available at: http://whqlibdoc.who.int/trs/WHO_TRS_894.pdf [last accessed July 2014].

15. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global Strategy for the Diagnosis, Management and Prevention of Chronic Obstructive Pulmonary Disease. Revised 2014. Available at: http://www.goldcopd.org/uploads/users/files/GOLD_Report_2014_Jun11.pdf [last accessed July 2014].

16. NICE. Quality and Outcomes Framework for COPD, NM63 Advisory Committee Recommendations. 2012. Available at: http://www.nice.org.uk/aboutnice/qof/indicators_detail.jsp?summary=13808 [last accessed July 2014].

17. Fletcher CM, Elmes PC, Fairbairn AS, *et al*. The significance of respiratory symptoms and the diagnosis of chronic bronchitis in a working population. *BMJ* 1959;2:257–66.

18. Williams T, van Staa T, Puri S, *et al*. Recent advances in the utility and use of the General Practice Research Database as an example of a UK Primary Care Data resource. *Ther Adv Drug Safety* 2012;3:89–99.

19. Burge S, Wedzicha JA. COPD exacerbations: definitions and classifications. *Eur Respir J* 2003;21(Suppl. 41):46s–53s.

20. Donaldson GC, Wedzicha JA. COPD exacerbations 1: epidemiology. *Thorax* 2006;61:164–8.

21. Donaldson GC, Hurst JR, Smith CJ, *et al*. Increased risk of myocardial infarction and stroke following exacerbation of COPD. *Chest* 2010;137(5):1091–7.

22. Hardin M, Silverman EK, Barr RG, *et al*. The clinical features of the overlap between COPD and asthma. *Respir Res* 2011;12:127.

23. De Marco R, Pesce G, Marcon A, *et al*. The Coexistence of Asthma and Chronic Obstructive Pulmonary Disease (COPD): Prevalence and Risk Factors in Young, Middle-aged and Elderly People from the General Population. *PLoS ONE* 2013;8(5):e62985.

24. Wedzicha JA, Brill SE, Allinson JP, *et al*. Mechanisms and impact of the frequent exacerbator phenotype in chronic obstructive pulmonary disease. *BMC Med* 2013;11:181.

25. Wilkinson TMA, Donaldson GC, Hurst JR, *et al*. Early Therapy Improves Outcomes of Exacerbations of Chronic Obstructive Pulmonary Disease. *Am J Respir Crit Care Med* 2004;169(12):1298–303.

26. Halpin DMG, Decramer M, Celli B, *et al*. Risk of nonlower respiratory serious adverse events following exacerbations in the 4-year UPLIFT trial. *Lung* 2011;189:261–8.

27. Zeki AA, Schivo M, Chan A, *et al*. The Asthma-COPD Overlap Syndrome: A Common Clinical Problem in the Elderly. *J Allergy* 2011:861926.

~~28. Thomsen M, Ingebrigtsen TS, Marott JL, *et al*. Inflammatory biomarkers and exacerbations in chronic obstructive pulmonary disease. *JAMA* 2013;309(22):2353.~~

~~29. Valvi D, Mannino DM, Müllerova H, *et al*. Fibrinogen, chronic obstructive pulmonary disease (COPD) and outcomes in two United States cohorts. *Int J Chron Obstruct Pulmon Dis* 2012;7:173–82.~~

~~30. O'Donnell DE, Laveneziana P. Dyspnea and activity limitation in COPD: mechanical factors. *COPD* 2007;4(3):225–36.~~

~~31. Stevenson NJ, Walker PP, Costello RW, *et al*. Lung mechanics and dyspnea during exacerbations of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2005;172:1510–6.~~

~~32.~~28. Bednarek M, Maciejewski J, Wozniak M, *et al*. Prevalence, severity and underdiagnosis of COPD in the primary care setting. *Thorax* 2008;63:402–7.

~~33.~~29. Soriano JB, Maier WC, Egger P, *et al*. Recent trends in physician diagnosed COPD in women and men in the UK. *Thorax* 2000;55:789–94.

~~34.~~30. Soriano JB, Visick GT, Muellerova H, *et al*. Patterns of co-morbidities in newly diagnosed COPD and asthma in primary care. *Chest* 2005;128(4):2099–107.

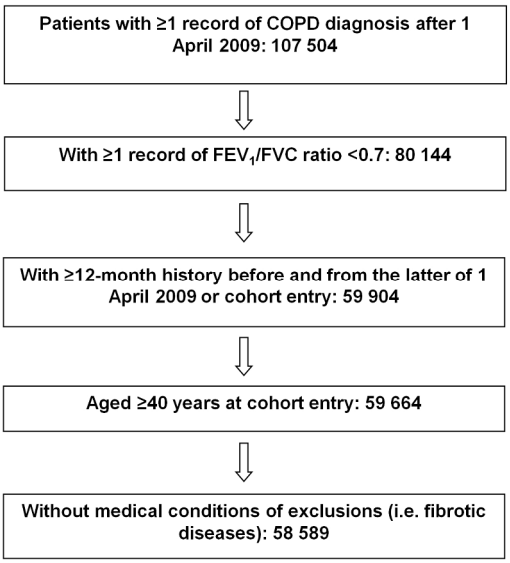
~~35.~~31. García Rodríguez LA, Wallander M-A, Martín-Merino E, *et al*. Heart failure, myocardial infarction, lung cancer and death in COPD patients: A UK primary care study. *Respir Med* 2010;104(11):1691–9.

FIGURE LEGENDS

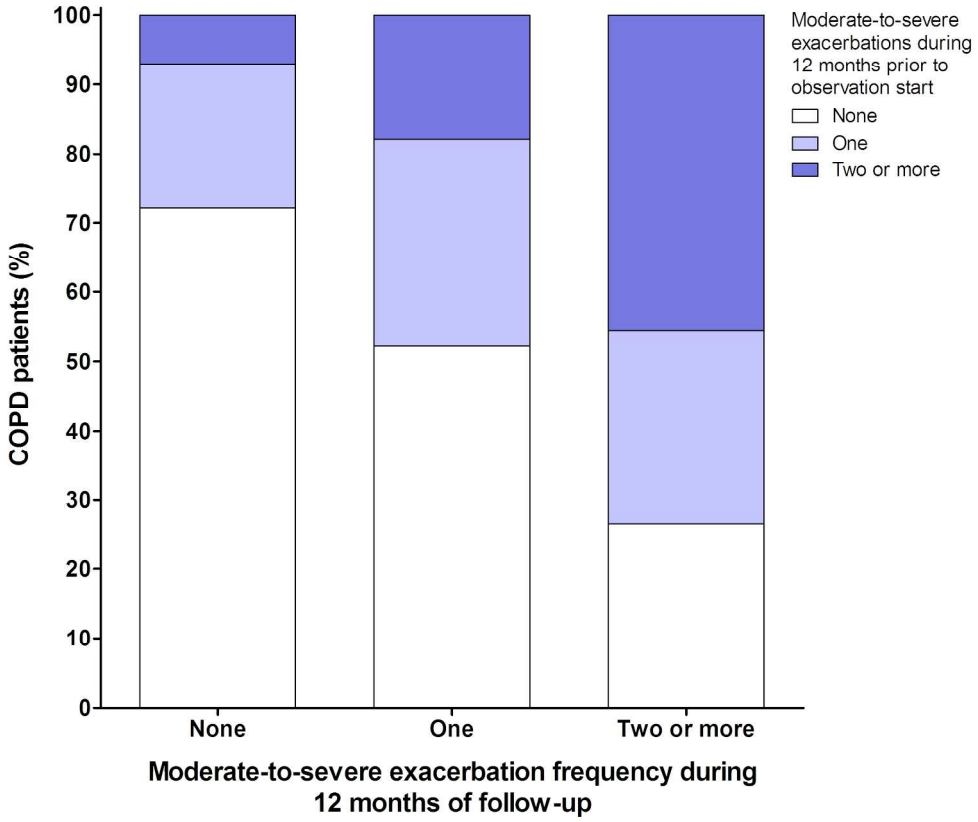
Figure 1 Patient recruitment flow chart.

Figure 2 Frequency of moderate-to-severe COPD exacerbations in the 12 months prior to cohort entry in relation to the frequency of exacerbation episodes during the 12-month follow-up period.

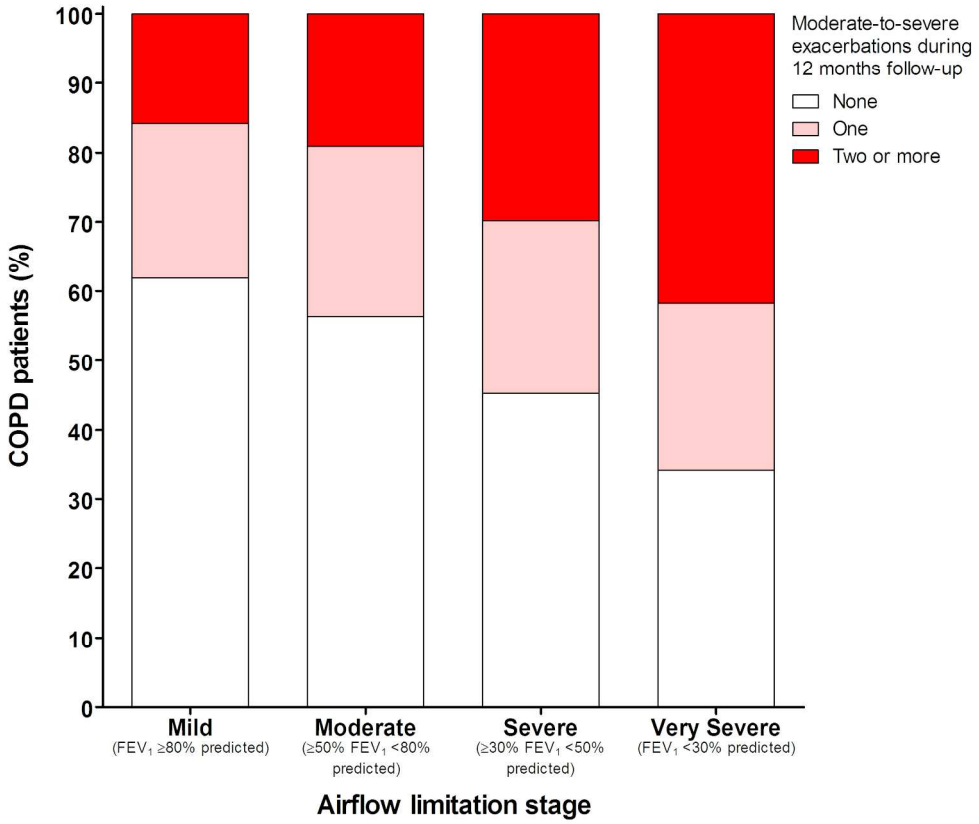
Figure 3 COPD exacerbation frequency during the 12-month follow-up period by a) airflow limitation stage, b) MRC dyspnoea grade.



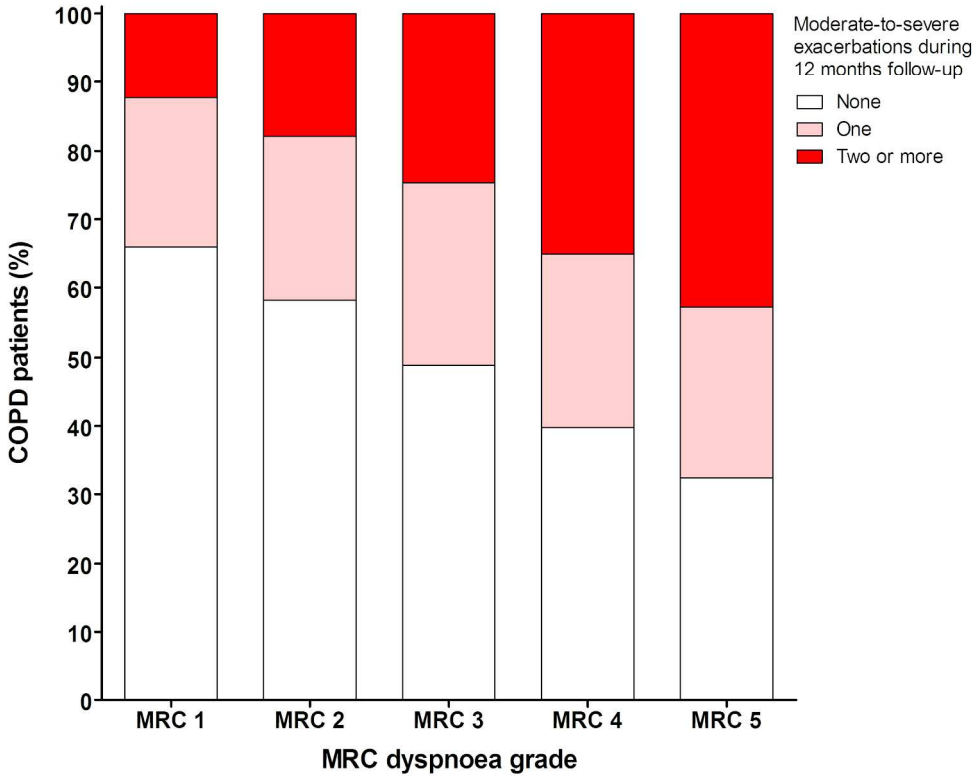
Patient recruitment flow chart.
190x254mm (300 x 300 DPI)



Frequency of moderate-to-severe COPD exacerbations in the 12 months prior to cohort entry in relation to the frequency of exacerbation episodes during the 12-month follow-up period.
213x185mm (300 x 300 DPI)



COPD exacerbation frequency during the 12 month follow-up period by airflow limitation stage.
214x184mm (300 x 300 DPI)



COPD exacerbation frequency during the 12 month follow-up period by MRC dyspnoea grade.
216x176mm (300 x 300 DPI)

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *cohort studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	4
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	5, Fig. 1
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5–7
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	5–7
Bias	9	Describe any efforts to address potential sources of bias	15–16
Study size	10	Explain how the study size was arrived at	5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	7
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	7
		(*b*) Describe any methods used to examine subgroups and interactions	7
		(*c*) Explain how missing data were addressed	7
		(*d*) If applicable, explain how loss to follow-up was addressed	N/A, all patients follow-up for a fixed period of 12 months

		(*e*) Describe any sensitivity analyses	N/A
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Fig. 1
		(b) Give reasons for non-participation at each stage	N/A
		(c) Consider use of a flow diagram	Fig. 1
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	8, Table 1
		(b) Indicate number of participants with missing data for each variable of interest	7
		(c) Summarise follow-up time (eg, average and total amount)	8
Outcome data	15*	Report numbers of outcome events or summary measures over time	8
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Table 2
		(*b*) Report category boundaries when continuous variables were categorized	Table 2
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	8
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	11, 13
Discussion			
Key results	18	Summarise key results with reference to study objectives	14
Limitations			
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	14–16
Generalisability	21	Discuss the generalisability (external validity) of the study results	14–16
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	17

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.