ORIGINAL RESEARCH

Development and Validation of a Method to Estimate COPD Severity in Multiple Datasets: A Retrospective Study

Jennifer K. Quint · Caroline O'Leary · Alessandra Venerus ·

Ulf Holmgren · Precil Varghese · Claudia Cabrera

ABSTRACT

Introduction: Outcomes in chronic obstructive pulmonary disease (COPD) such as symptoms, hospitalisations and mortality rise with increasing disease severity. However, the heterogeneity of electronic medical records presents a significant challenge in measuring severity across geographies. We aimed to develop and validate a method to approximate COPD severity using the Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2011 classification scheme, which categorises patients based on forced expiratory volume in 1 s, hospitalisations and the modified Medical Research Council dyspnoea scale or COPD Assessment Test.

Methods: This analysis was part of a comprehensive retrospective study, including patients sourced from the IQVIA Medical Research Data [IMRD; incorporating data from The Health Improvement Network (THIN), a Cegedim database] and the Clinical Practice Research Datalink (CPRD) in the UK, the Disease Analyzer in Germany and the Longitudinal Patient Data in Italy, France and Australia. Patients in the CPRD with the complete set of information required to calculate GOLD 2011 groups were used to develop the method. Ordinal logistic models at COPD diagnosis and at index (first episode of triple therapy) were then used to validate the method to estimate COPD severity, and this was applied to the full study population to estimate GOLD 2011 categories.

Results: Overall, 4579 and 12,539 patients were included in the model at COPD diagnosis and at index, respectively. Models correctly classified 74.4% and 75.9% of patients into severe and non-severe categories at COPD diagnosis and at index, respectively. Age, gender, time between diagnosis and start of triple therapy, healthcare resource use, comorbid conditions and prescriptions were included as covariates.

Conclusion: This study developed and validated a method to approximate disease severity based on GOLD 2011 categories that can potentially be used in patients without all the key parameters needed for this calculation.

Keywords: Chronic obstructive pulmonary disease; Disease severity; Logistical model;

J. K. Quint
Imperial College London, London, UK

C. O'Leary · A. Venerus
IQVIA, London, UK

U. Holmgren · C. Cabrera (✉)
Real World Science and Digital, BioPharmaceuticals Medical, AstraZeneca, Gothenburg, Sweden
e-mail: claudia.s.cabrera@astrazeneca.com

P. Varghese
Biopharmaceuticals Medical, Respiratory and Immunology, AstraZeneca, Gaithersburg, MD, USA

Retrospective study; Treatment initiation; Triple therapy

Key Summary Points

Why carry out this study?

There is a need to distinguish between severe and non-severe COPD when investigating patient outcomes; however, the heterogeneity of electronic medical records (EMR) in terms of diagnostic coverage of key parameters makes it difficult to measure COPD severity across different populations.

We aimed to develop and validate a method, based on GOLD 2011 categories, to estimate severity of disease in patients with COPD who did not have all the key parameters needed for direct calculation.

What was learned from the study?

Our method correctly classified approximately three-quarters of patients into severe and non-severe categories both at COPD diagnosis and at first instance of triple therapy.

This method provides a framework for the integration of such models into electronic healthcare records so that COPD severity can be estimated in patients without all the key parameters needed for this calculation in a real-world setting, and it has the potential for use in future EMR retrospective studies.

DIGITAL FEATURES

This article is published with digital features, including a summary slide, to facilitate understanding of the article. To view digital features for this article go to https://doi.org/10.6084/m9.figshare.13176932.

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is characterised by airflow obstruction that is not fully reversible [1] and is the third leading cause of death worldwide [2]. In 2016, estimated global prevalence was 251 million [3, 4], with over 3 million deaths in 2015, corresponding to 5% of all deaths globally [5]. The disease burden experienced by patients with severe COPD can be high, including symptoms such as cough, dyspnoea, fatigue, weight loss, sleep disturbance and anorexia. Both hospitalisations and mortality risk are greater for patients with severe COPD compared to non-severe COPD [6–10]. Therefore, it is important to distinguish between severe and non-severe COPD when investigating patient outcomes. The Global Initiative for Chronic Obstructive Lung Disease (GOLD) classification scheme can be used as a proxy for disease severity. In 2007, this assessment rated the degree of airflow obstruction by post-bronchodilator spirometry results alone to categorise patients' disease severity. The 2011 update included the degree of symptoms measured through the COPD Assessment Test (CAT) or the modified Medical Research Council dyspnoea scale (mMRC) and exacerbation history. Previous work shows that the GOLD 2007 classification had the same predictive ability as the GOLD 2011 classification in a pooled analysis [11]. Subsequent GOLD updates were formulated in 2017–2020 that were similar to the GOLD 2011 classification, but separated pulmonary function from patient risk assessment groups, highlighting the importance of symptoms and exacerbation history in patients with COPD [10]. The GOLD 2011 classification was used in this study, as this was the current standard, which was most impactful throughout the study period, and yet this analysis still has relevance for more recent GOLD updates that separate assessment of airflow limitation severity from symptom burden and exacerbation risk. GOLD 2011–2020 grades patient risk from A (least severe) to D (most severe) based on a combined assessment [10], with A and B categories equating to GOLD 1 (mild) and 2 (moderate), based on the severity of airflow

limitation measured, using forced expiratory volume in 1 s (FEV_1), and C and D categories equating to GOLD 3 (severe) and 4 (very severe). In this study, A and B categories in the combined GOLD 2011 assessment were used as a proxy for non-severe COPD, and C and D categories for severe COPD.

Heterogeneity of electronic medical records (EMR) data quality, frequency of diagnostic capture and coverage for key parameters (FEV_1, hospitalisations, mMRC or CAT) present a significant challenge in measuring COPD severity across geographies. As cross-country comparisons become commonplace, a way of approximating disease severity, where it is not implicitly recorded, is needed. Therefore, this study aimed to develop and validate a method of categorising COPD disease status that could be used to estimate severity in patients without these data. Patient populations from the UK, Germany, Italy, France and Australia were included given the similarities in healthcare infrastructure in these countries, while allowing assessment of potential population diversity and its impact on treatment.

Treatment strategies for COPD include long-term inhaled pharmacologic therapies, including short- and long-acting β_2-agonists (SABA and LABA) and/or short- and long-acting muscarinic antagonists (SAMA and LAMA) with or without inhaled corticosteroids (ICS). Triple therapy combination with ICS, LABA and LAMA is recommended in patients who are inadequately controlled despite dual therapy [10, 12].

METHODS

Study Objective

This analysis aimed to develop and validate a method to approximate COPD severity using the GOLD 2011 classification scheme. Patients with complete information required to calculate GOLD 2011 groups were used to develop and validate this method, which was then applied to patients with incomplete information to calculate GOLD 2011 groups. This analysis was part of a large, multi-country, retrospective study to understand treatment pathways to triple therapy where COPD risk, based on GOLD groups A/B (less severe) and C/D (more severe), was an adjusting factor in analyses.

Study Population

Patients from the IQVIA Medical Research Data [(IMRD), incorporating data from The Health Improvement Network (THIN), a Cegedim database] [13, 14] and the Clinical Practice Research Datalink (CPRD) [15] in the UK, the Disease Analyzer (DA) [16] in Germany (GP and pneumologist panels) and the Longitudinal Patient Data (LPD) in Italy, France and Australia [17, 18] were included. All databases outside the UK included primary care only and, therefore, could not be used to calculate disease severity directly, as secondary care was not linked. In the UK, IMRD and CPRD were combined to obtain a larger sample, and a subset of these patients in the CPRD linked to Hospital Episodes Statistics (HES) with complete data (FEV_1, hospitalisations, mMRC or CAT) were used to develop and validate the severity categorisation method.

CPRD contains primary care medical records from 5.5 million patients in more than 670 UK practices, approximately 8% of the UK population. HES is derived from secondary care records in England. To derive information on secondary care episodes for the identification of exacerbations, CPRD was linked to HES. Linked CPRD-HES data were available to March 2016.

For the broader multi-country study, index date was defined as the first instance of triple therapy during the study period (1 January 2005 to 1 May 2016), defined as a prescription from each class of ICS, LABA and LAMA with at least 14 days of overlap, according to recorded duration or calculated based on quantity and dose. Patients were followed until the earliest of death, transfer out of practice or end of study.

Patients were included in the wider multi-country study if they initiated triple therapy during the study period and had at least 12 months of data prior to index. Patients required a diagnosis of COPD, defined as evidence of smoking (current or ex-smoker) at any point in their record (in the UK) or a

confirmatory diagnosis of COPD (in all other countries) and at least one COPD diagnosis code on or after their 40th birthday. Also, to be included in the COPD severity categorisation method development and validation, linked CPRD-HES data and a record of all variables needed to calculate GOLD 2011 categories at COPD diagnosis and/or at index were required. Patients with unknown gender were excluded from all analyses. Because of common misclassification of asthma and COPD, it was decided that patients with asthma would not be excluded, as this could result in excluding patients with COPD [19].

Model Implementation/Validation

Severity models were developed to be used in estimating GOLD 2011 categories (Table 1) for those without all the information needed to calculate this directly. Two models were developed, one at COPD diagnosis and the other at index date (the first instance of triple therapy) (Fig. 1).

The dependent variable in each model was the derived GOLD classification calculated directly from the data for those with complete information. The variables used to calculate GOLD 2011 categories at COPD diagnosis and at index were mMRC or CAT (mMRC was preferred where both were present), FEV_1 (recorded or calculated) and exacerbations per year. Patients needed at least one record for each of these variables in the 12 months prior to, and including, the date of COPD diagnosis/index date. If more records were present, the closest to the date of diagnosis/index date was used. The

percentage of predicted FEV_1 was as recorded, or calculated, as a function of age and height, if not available [20]. Patients' exacerbations were considered in the 12 months prior to COPD diagnosis/index date. A distinction was made between exacerbations recorded in primary care, estimated through an algorithm by Rothnie et al. [21, 22], and those resulting in hospitalisation (recorded in HES or in the CPRD as "hospitalisations due to COPD"), as GOLD treats exacerbation in secondary and primary care differently (Table 1).

The independent covariates were chosen based on clinical judgement and data availability across all databases; this meant the model was developed and validated in patients with complete information in order to calculate the GOLD group (dependent variable), which could be applied to those with incomplete mMRC, CAT, FEV_1 or exacerbation data. Covariates included age, gender, time between diagnosis and first triple therapy (index date), comorbid conditions, prescriptions and healthcare resource use (Fig. 1).

Once GOLD 2011 categories were calculated for linked CPRD-HES patients with complete information, the COPD severity identification method was developed on the same population using these categories as the outcome, allowing comparison between actual and estimated GOLD categories. Two ordinal logistic regression models were developed using patients with complete information at COPD diagnosis/index.

The study populations with complete information used to estimate GOLD categories were randomly split into development (80%) and validation datasets (20%). Models were

Table 1 Classification of COPD based on GOLD criteria 2011

GOLD category	Characteristics	Classification of airflow limitation	Exacerbations/year	mMRC	CAT
A	Low risk, fewer symptoms	GOLD 1–2	≤ 1 (not leading to admission)	0–1	< 10
B	Low risk, more symptoms	GOLD 1–2	≤ 1 (not leading to admission)	2	≥ 10
C	High risk, fewer symptoms	GOLD 3–4	2 or ≥ 1 leading to hospital admission	0–1	< 10
D	High risk, more symptoms	GOLD 3–4	2 or ≥ 1 leading to hospital admission	2	≥ 10

CAT COPD Assessment Test, *COPD* chronic obstructive pulmonary disease, *GOLD* Global Initiative for Chronic Obstructive Lung Disease, *mMRC* modified Medical Research Council dyspnoea scale
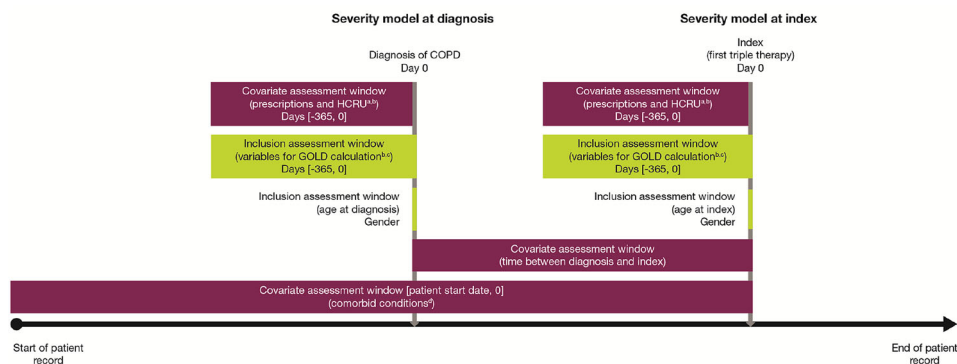
Fig. 1 Overview of variables included in the models. *CAT* COPD Assessment Test, *COPD* chronic obstructive pulmonary disease, *FEV₁* forced expiratory volume in 1 s, *GOLD* Global Initiative for Chronic Obstructive Lung Disease, *GP* general practitioner, *HCRU* healthcare resource utilisation, *ICS* inhaled corticosteroids, *LABA* long-acting β₂-agonist, *LAMA* long-acting muscarinic antagonist, *mMRC* modified Medical Research Council dyspnoea scale, *OCS* oral corticosteroids, *SABA* short-acting β₂-agonist, *SAMA* short-acting muscarinic antagonist. [a]Prescriptions is the number of prescriptions for LABA, LAMA, ICS, LABA + ICS, LABA + LAMA, SABA, SAMA, SABA + SAMA, OCS and other COPD drugs (oxygen, mucolytic products, roflumilast, theophylline and azithromycin), and healthcare resource use is the total number of visits (including GP visits, hospitalisations and annual reviews). [b]In cases where either

the period between the start of the patient's record and COPD diagnosis or the period between COPD diagnosis and index date was 6–12 months, all covariates collected prior to index date and post-COPD diagnosis were annualised to 12 months (e.g., two visits in 6 months became four visits in 12 months). Patients with < 6 months were excluded from the calculation at that time point. [c]mMRC or CAT (mMRC was preferred in cases where both were present), FEV₁ (recorded or calculated) and exacerbations per year. [d]Comorbid conditions: cardiovascular disease (ischaemic heart disease, angina, myocardial infarction, coronary artery bypass graft/percutaneous coronary intervention and/or hypertension), heart failure, atrial fibrillation, osteoporosis, depression and/or anxiety, diabetes and gastroesophageal reflux disease

estimated in the development datasets and validated in the validation datasets. Different models were tested, and the model with the best goodness of fit was chosen as the final model at COPD diagnosis/index. The goodness of fit measures included Kappa index (measure of agreement) and percentage agreement between each patient's calculated GOLD category and that estimated by the models, which were then applied to the study population with incomplete information in the UK, Germany, Italy, France and Australia to estimate GOLD categories.

An independent scientific advisory committee approved the use of CPRD data (16_298R), and an independent scientific review committee approved the use of IMRD data (16THIN097). No ethics approval was required for the DA or LPD databases, as German law allows the use of anonymous electronic medical records for research purposes under certain

conditions. Data was collected and processed in full compliance with General Data Protection Regulation (GDPR) and local privacy regulations requirements.

RESULTS

Predictive Models

Models for predicting two GOLD severity groups showed better goodness of fit compared to those predicting all four GOLD severity groups; therefore, the final models only estimated two severity groups: less risk (GOLD A or B) and more risk (GOLD C or D).

Based on positive predictive values (PPV) and negative predictive values (NPV), a probability level of 0.6 was chosen to classify patients as severe, both at COPD diagnosis and at index.

The probability threshold was data-driven and chosen to maximise the NPV while maintaining a sufficiently high PPV. Therefore, patients with an estimated probability ≥ 0.6, according to the model, were assigned to the severe group (C/D), while patients with an estimated probability < 0.6 were assigned to the non-severe group (A/B) (Table 2).

Severity Model at Diagnosis

At COPD diagnosis, 3660 and 919 patients were included in the development and validation datasets, respectively (Fig. 2). The model correctly predicted COPD severity for 74.4% of patients in the validation dataset, with a PPV of 82.3% and an NPV of 50.2% (Table 2), with similar figures observed in the development dataset. Overall accuracy was 74.4% and balanced accuracy 73.5%.

Model estimates show that factors associated with severity included gender, time between diagnosis and initiation of triple therapy, prescriptions [SABA, oral corticosteroids (OCS) and antibiotics] and comorbidities (cardiovascular disease, depression/anxiety, gastroesophageal reflux disease and asthma) (Table 3).

Severity Model at Index

At index, 10,032 and 2507 patients were included in the development and validations datasets, respectively (Fig. 2). The distribution of GOLD categories at index in the development and validation cohorts, respectively, was as follows: A, 12.6% and 12.6%; B, 14.9% and 15.2%;

C, 27.4% and 26.6%; D, 45.1% and 46.6%. The model correctly predicted COPD severity for 75.9% patients in the validation dataset, with a PPV of 84.8% and an NPV of 56.6% (Table 2), with similar figures observed in the development dataset. Overall accuracy was 76.4% and balanced accuracy 70.4%.

Factors associated with severity included gender, time between diagnosis and initiation of triple therapy, healthcare resource use, prescriptions (SABA, SAMA, SABA-SAMA fixed combinations, OCS, antibiotics and other drugs for COPD) and comorbidities (cardiovascular disease, heart failure, depression/anxiety, asthma and gastroesophageal reflux disease) (Table 3).

GOLD 2011 Categories

These methods were used to categorise COPD severity in the main study population; the results are shown in Table 4. In the UK, GOLD category was calculated for linked CPRD-HES patients with all variables needed for direct calculation and estimated using the methods to identify COPD severity for the remaining patients. Patients without records for model covariates were not eligible for GOLD estimation through these methods and, therefore, were not assigned a severity classification.

At COPD diagnosis, most patients were estimated to be in GOLD A or B categories, ranging from 50.6% in Australia [95% confidence interval (CI): 48.8–53.4%] to 94.1% in Germany (pneumologist-treated) (95% CI 93.1–95.0%); however, in the UK only 38.7% (95% CI

Table 2 Probability values for severity models at COPD diagnosis and index date

Probability level	PPV and NPV—Validation dataset				Correctly predicted—Validation dataset	
	At COPD diagnosis		At index (triple therapy)		At COPD diagnosis	At index (triple therapy)
	PPV	NPV	PPV	NPV	%	%
0.500	0.7822	0.5022	0.8153	0.5916	73.1	76.3
0.600	0.8232	0.5020	0.8479	0.5655	74.4	75.9
0.700	0.8886	0.4930	0.8902	0.5381	72.6	75.0

COPD chronic obstructive pulmonary disease, *NPV* negative predicted value, *PPV* positive predicted value
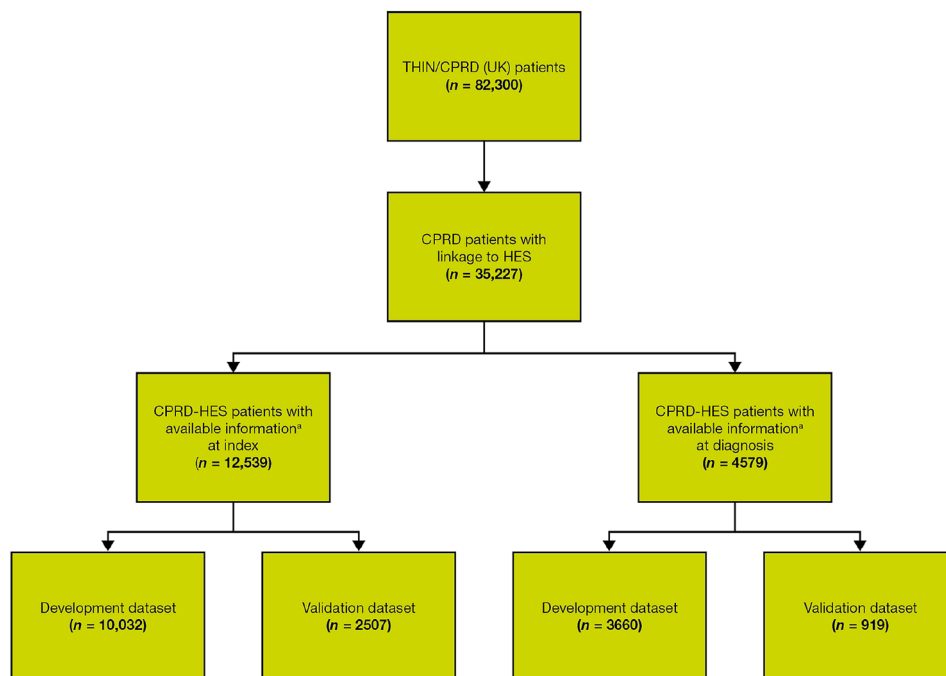
Fig. 2 Study population included in the model flowchart. *CAT* COPD Assessment Test, *CPRD* Clinical Practice Research Datalink, *FEV₁* forced expiratory volume in 1 s, *HES* Hospital Episodes Statistics, *mMRC* modified Medical Research Council dyspnoea scale, *THIN* The Health Improvement Network. [a]Available information includes mMRC or CAT, FEV$_1$ (recorded or calculated) and exacerbations per year

38.0–39.3%) of patients were estimated to be in GOLD A or B categories.

The proportion of patients estimated to be in GOLD C or D categories pre-triple therapy increased from diagnosis in all countries, with the greatest increase in Germany [GP-treated: 13.9% (95% CI 13.4–14.5%) to 32.8% (95% CI 32.1–33.4%); pneumologist-treated: 5.9% (95% CI 5.0–6.9%) to 11.6% (95% CI 10.8–12.4%)] and Italy [95% CI 38.4% (95% CI 37.4–39.5%) to 74.6% (95% CI 73.7–75.5%)], where the proportion of patients in GOLD categories C or D nearly doubled.

Fewer patients had missing estimated GOLD category at index, compared to those at COPD diagnosis. Patients with missing GOLD 2011 categories ranged from 21.8 to 67.8% at COPD diagnosis and from 10.8 to 37.6% at index. Most notably, the proportion of patients with missing estimated GOLD categories decreased almost sevenfold in Germany, and by more than half in the UK.

DISCUSSION

This study used linked primary (CPRD) and secondary (HES) UK EMR data to develop and validate a method to categorise COPD severity among patients in the UK, France, Germany, Italy and Australia, and demonstrate that COPD severity can be estimated among patients who do not have key clinical measures (FEV$_1$, hospitalisations, mMRC or CAT) in their EMR data. While GOLD 2011 categories were used in this study, GOLD 2020 categories use the same measurements for symptoms (CAT or mMRC), exacerbations per year (≤ 1 or ≥ 2) and the degree of airflow limitation (GOLD 1, 2, 3 or 4). The difference between GOLD 2011 and GOLD 2020 assessment criteria is that GOLD 2020 is only applied at the initial assessment, and subsequent treatment is based on major treatable traits and current medication. Therefore, our findings can be used to approximate COPD

Table 3 Covariates included in the COPD severity models and association with disease severity

Covariate	Odds ratio	At COPD diagnosis 95% CI		p value	Odds ratio	At index 95% CI		p value
		Lower	Upper			Lower	Upper	
Demographics								
Age (at COPD diagnosis/index)	1.008	1.000	1.016	0.0625	1.002	0.996	1.008	0.4941
Gender (female vs male)	0.707	0.596	0.839	< 0.0001	0.711	0.637	0.793	< 0.0001
Time between diagnosis and triple therapy	1.000	1.000	1.000	0.0011	1.000	1.000	1.000	0.0009
Healthcare resource use in the 12 months before COPD diagnosis/index[a]								
Number of visits	1.243	0.915	1.690	0.1640	2.753	2.135	3.550	< 0.0001
Number of prescriptions in the 12 months before COPD diagnosis/index[a]								
LABA	1.156	0.908	1.473	0.2392	1.123	0.998	1.262	0.0531
LAMA	1.034	0.780	1.369	0.8177	1.017	0.879	1.178	0.8185
ICS	1.035	0.847	1.265	0.7355	1.108	0.994	1.236	0.0643
LABA + ICS	1.079	0.870	1.339	0.4884	1.030	0.886	1.199	0.6993
LABA + LAMA	10.916	0.326	365.784	0.1822	0.868	0.459	1.642	0.6637
SABA	1.364	1.059	1.757	0.0163	1.479	1.315	1.663	< 0.0001
SAMA	1.070	0.845	1.355	0.5738	1.276	1.160	1.404	< 0.0001
SABA + SAMA	1.289	0.930	1.786	0.1277	1.575	1.372	1.807	< 0.0001
OCS	2.580	2.062	3.229	< 0.0001	1.924	1.719	2.154	< 0.0001
Antibiotics	6.832	5.825	8.014	< 0.0001	4.332	3.931	4.774	< 0.0001
Other COPD drugs	1.201	0.885	1.630	0.2405	1.226	1.090	1.378	0.0007
Total number of prescriptions	0.708	0.469	1.071	0.1018	0.417	0.285	0.612	< 0.0001
Comorbidities at any point in the patient history up to and including index (yes/no)								
Cardiovascular disease	0.822	0.687	0.985	0.0335	0.781	0.693	0.879	< 0.0001
Heart failure	1.040	0.752	1.438	0.8137	1.229	1.004	1.505	0.0460
Atrial fibrillation	0.981	0.705	1.365	0.9103	1.081	0.891	1.312	0.4304
Osteoporosis	0.985	0.718	1.351	0.9256	1.069	0.877	1.304	0.5089
Depression and/or anxiety	0.810	0.683	0.962	0.0163	0.875	0.782	0.978	0.0187
Diabetes	1.009	0.799	1.274	0.9406	0.965	0.837	1.112	0.6182
Gastroesophageal reflux disease	0.721	0.597	0.869	0.0006	0.783	0.695	0.883	< 0.0001

Table 3 continued

Covariate	Odds ratio	At COPD diagnosis 95% CI		p value	Odds ratio	At index 95% CI		p value
		Lower	Upper			Lower	Upper	
Asthma	0.676	0.546	0.837	0.0003	0.755	0.676	0.844	< 0.0001

CI confidence interval, *COPD* chronic obstructive pulmonary disease, *ICS* inhaled corticosteroids, *LABA* long-acting β$_2$-agonist, *LAMA* long-acting muscarinic antagonist, *OCS* oral corticosteroids, *SABA* short-acting β$_2$-agonist, *SAMA* short-acting muscarinic antagonist

[a] Logarithm transformation applied

Table 4 GOLD 2011 categories (calculated or estimated depending on data availability)

GOLD category	UK (*n* = 82,300)	Germany (GPs, *n* = 22,178)	Germany (pneumologists, *n* = 6816)	Italy (*n* = 10,443)	France (*n* = 6514)	Australia (*n* = 2478)
GOLD 2011 at COPD diagnosis[a]						
A or B, *n* (%),	22,126 (38.7)	13,776 (86.1)	2072 (94.1)	5032 (61.6)	2527 (76.6)	635 (50.6)
95% CI	(38.0–39.3)	(85.5–86.6)	(93.1–95.0)	(60.5–62.6)	(75.1–78.0)	(48.8–53.4)
C or D, *n* (%),	35,097 (61.3)	2230 (13.9)	131 (5.9)	3138 (38.4)	772 (23.4)	620 (49.4)
95% CI	(60.8–61.8)	(13.4–14.5)	(5.0–6.9)	(37.4–39.5)	(22.0–24.9)	(46.6–52.2)
Missing at baseline, *n* (%)	25,077 (30.5)	6172 (27.8)	4613 (67.7)	2273 (21.8)	3215 (49.4)	1223 (49.4)
GOLD 2011 pre-triple therapy[b]						
A or B, *n* (%),	17,933 (25.3)	12,498 (67.2)	5379 (88.4)	2128 (25.4)	3045 (70.9)	526 (34.0)
95% CI	(24.7–25.9)	(66.6–67.9)	(87.6–89.2)	(24.5–26.3)	(69.5–72.2)	(31.7–36.4)
C or D, *n* (%),	52,939 (74.7)	6090 (32.8)	704 (11.6)	6254 (74.6)	1253 (29.2)	1020 (66.0)
95% CI	(74.3–75.1)	(32.1–33.4)	(10.8–12.4)	(73.7–75.5)	(27.8–30.5)	(63.6–68.3)
Missing at index date, *n* (%)	11,428 (13.9)	3590 (16.2)	733 (10.8)	2061 (19.7)	2216 (34.0)	932 (37.6)

CI confidence interval, *COPD* chronic obstructive pulmonary disease, *GOLD* Global Initiative for Chronic Obstructive Lung Disease

[a] Only for patients with at least 6 months of data prior to the date of COPD diagnosis

[b] Only for patients with index date after COPD diagnosis and at least 6 months between diagnosis and index

severity in line with GOLD 2020 recommendations [10].

Cross-country comparisons of health systems and policies have become more common [23], and this analysis shows a method for capturing disease severity in populations where it is not implicitly recorded. Cross-country comparisons enable an understanding of differences

between countries and provide evidence that may assist clinical management and guidelines.

The models developed in this study performed relatively well, particularly for patients in the more severe group, and performed better than the model used in a US claims study, which accurately predicted COPD severity for 62.7% of patients, with a PPV of 67.0% for severe/very severe patients [24]. The differences in accuracy between the models could be due to the US study excluding patients with asthma.

An algorithm to measure severity was developed for a recent respiratory trial. The performance of this was poorer than the method used in the present study, with the trial model correctly identifying 53% of patients with severe COPD, of which 8% had very severe COPD [25]. The predictors included were similar; however, the trial excluded prescriptions such as antibiotics, which exhibited high odds ratios in this analysis and a strong association with COPD severity classification.

In this study, patients with comorbidities at COPD diagnosis/index (cardiovascular disease, depression and/or anxiety and gastroesophageal reflux disease) were less likely to be categorised into the severe group compared with those who did not have these conditions; this finding is similar to a previous study [26]. However, another European study found that comorbidities were significantly associated with COPD severity [27]. This discrepancy could be due to the grouping of A/B and C/D applied in this study, as the previous study showed that comorbidities were more frequent in GOLD B and D categories, whereas in this study the categories were combined.

At diagnosis, most patients, apart from those in the UK, were in GOLD A or B categories; this could be due to misclassification of patients with estimated GOLD criteria, due to the lack of spirometry-confirmed testing outside of the UK [28–30]. Also, differences in risk factors across countries can influence physician decisions and lead to misdiagnosis [30]. For example, patients with less smoking history are more likely to be misdiagnosed [30].

In the 12 months prior to triple therapy, most British, Italian and Australian patients were in GOLD C or D categories, in accordance with GOLD criteria for triple therapy initiation [10]. However, in Germany (both groups) and France, most patients were in GOLD A or B categories at triple therapy initiation, indicating the possibility that they were being overtreated [31]. The observed differences in classification between these countries may be due to regional variations in the measurement of the variables needed to estimate disease severity.

Pneumologist-treated patients in Germany were more likely to have their COPD diagnosis as their first record, with little or no history prior to this event, as pneumologists only see patients with respiratory diseases. The missing variables needed to estimate GOLD categories, due to a lack of clinical history data, might explain the higher level of patients with missing severity at diagnosis compared to GP-treated patients in Germany.

When examining GOLD categories pre-triple therapy, the proportion of patients estimated in GOLD C or D categories increased from diagnosis in all countries, suggesting patient risk increases over time prior to initiation of triple therapy. However, it should be noted that there is a higher proportion of patients with missing GOLD categories at diagnosis compared to at index.

Limitations

Patients with missing key parameters required to calculate GOLD categories were not included in the model development, which could lead to selection bias.

Due to the nature of this retrospective EMR study, it was not possible to confirm diagnoses for all patients that met appropriate COPD criteria; however, the high validity of diagnoses in the CPRD (in the UK), particularly in terms of the PPV of diagnostic codes, has been demonstrated in validation studies [32].

GOLD categories were calculated in patients with complete data in the UK, with covariates chosen based on availability in both the UK and European primary care databases. True GOLD categories could only be calculated in a subset of UK patients where complete primary and secondary EMR data were available. Therefore, while this method was validated in the UK, it

could not be validated in other countries. Also, records contributing to the calculation of GOLD categories, or used in our method, may be subject to under- or mis-recording. Therefore, if inaccuracies exist in the calculation of GOLD categories, these may be reflected in the accuracy of our method. For example, there is an underestimation (by over half) of exacerbations in patients with COPD [33]. We used an algorithm by Rothnie et al. in which patients with less severe COPD exacerbations may be further under-represented, given that mild exacerbations self-managed by patients were not captured [21, 22].

The model had high PPV and relatively low NPV values, making the estimation of non-severe patients less accurate; however, it correctly identified 74.4% and 75.9% of patients at COPD diagnosis and at index, respectively.

A binary outcome was used as a proxy for patient risk (C/D 'severe' and A/B 'non-severe'), as opposed to the four GOLD 2011 categories, since models predicting all four GOLD categories showed poor predictive power, with most patients estimated to belong to GOLD category D. Therefore, a binary outcome was used to increase accuracy.

While it is acknowledged that there are some differences between the GOLD 2011 criteria applied in this study and more recent GOLD updates, largely in terms of treatment options and approach to treatment, these differences are likely to have limited impact on our research question. Furthermore, spirometric test of pulmonary function—the item removed from patient risk assessment groups, which is not present in the latest GOLD criteria—is not commonly used in primary care decision-making.

This study found that the severity of COPD was lower in patients diagnosed with both COPD and asthma. COPD and asthma have been viewed as distinct conditions; however, evidence suggests they exhibit similar characteristics [34–37]. This study did not address this, and further research is needed to understand the severity of COPD among patients diagnosed with both COPD and asthma.

CONCLUSIONS

This study reports the development and validation of a method to categorise COPD severity using a GOLD 2011 calculation that can potentially be used to estimate COPD severity in patients without all the key parameters needed for this calculation in a real-world setting. This method may be used in future EMR retrospective studies to estimate COPD severity. It may also be used in future studies where linked data are not available due to the fact that severity is strongly associated with outcomes, but is not always readily available. Furthermore, a future goal could be that the models in this study provide a framework for the integration of this information into electronic healthcare records to ultimately inform decision making in the management of patients with COPD. Further research into machine learning algorithms and artificial intelligence applications is ongoing [38].

ACKNOWLEDGEMENTS

Authorship. All named authors meet the International Committee of Medical Journal Editors (ICJME) criteria for authorship for this article, take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.

Authorship Contributions. All authors made substantial contributions to the conception and design of the work. Caroline O'Leary and Alessandra Venerus made substantial contributions to the acquisition of data and to the analysis of data. All authors made substantial contributions to the interpretation of data and the preparation and review of the manuscript, approved the final version to be submitted and agree to be accountable for all aspects of the work.

Disclosures. Jennifer K Quint has received grants from IQVIA pertaining to this work, and grants and personal fees from AstraZeneca, Bayer, Boehringer Ingelheim and GlaxoSmithKline, and grants from Asthma UK, the British Lung Foundation, Chiesi, Medical Research Council and The Health Foundation, none of which relate to this work. Caroline O'Leary and Alessandra Venerus are employees of IQVIA who received consulting fees for conducting the study. Ulf Holmgren and Precil Varghese are employees of AstraZeneca. Claudia Cabrera is an employee of AstraZeneca and has an adjunct research position at the Karolinska Institute of Biostatistics and Epidemiology. The authors report no other conflicts of interest in this work.

Compliance with Ethics Guidelines. An independent scientific advisory committee approved the use of CPRD data (16_298R) and an independent scientific review committee approved the use of IMRD data (16THIN097). No ethics approval was required for the DA or LPD databases, as German law allows the use of anonymous electronic medical records for research purposes under certain conditions. Data was collected and processed in full compliance with General Data Protection Regulation (GDPR) and local privacy regulations requirements.

Data Availability. The datasets used and analysed during the current study are available on reasonable request in accordance with AstraZeneca's data sharing policy described at https://astrazenecagrouptrials.pharmacm.com/ST/Submission/Disclosure.

REFERENCES

1. European Respiratory Society. The economic burden of lung disease. 2019. https://www.erswhitebook.org/chapters/the-economic-burden-of-lung-disease/. Accessed 25 March 2019.

2. Lozano R, Naghavi M, Foreman K, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012;380:2095–128.

3. GBD 2016 Causes of Death Collaborators. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet. 2017;390:1151–210.

4. World Health Organization. Chronic Obstructive Pulmonary Disease (COPD). 2017. https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd). Accessed 4 November 2019.

5. GBD 2015 Chronic Respiratory Disease Collaborators. Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet Respir Med. 2017;5:691–706.

6. UpToDate. Management of refractory chronic obstructive pulmonary disease. 2019. https://www.uptodate.com/contents/management-of-refractory-chronic-obstructive-pulmonary-disease?search=severe%20copd%20outcome&source=search_result&selectedTitle=2~150&usage_type=default&display_rank=2. Accessed 20 March 2019.

7. Raherison C, Girodet P-O. Epidemiology of COPD. Eur Respir Rev. 2009;18:213–21.

8. Mannino DM, Buist AS, Petty TL, Enright PL, Redd SC. Lung function and mortality in the United States: data from the First National Health and Nutrition Examination Survey follow up study. Thorax. 2003;58:388–93.

9. Berry CE, Wise RA. Mortality in COPD: causes, risk factors, and prevention. COPD. 2010;7:375–82.

10. Global Initiative for Chronic Obstructive Lung Disease. 2020 report: global strategy for the diagnosis, management and prevention of COPD. 2020. https://goldcopd.org/gold-reports/. Accessed 17 March 2020.

11. Soriano JB, Lamprecht B, Ramírez AS, et al. Mortality prediction in chronic obstructive pulmonary disease comparing the GOLD 2007 and 2011 staging systems: a pooled analysis of individual patient data. Lancet Respir Med. 2015;3:443–50.

12. Vestbo J, Hurd SS, Agustí AG, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. Am J Respir Crit Care Med. 2013;187:347–65.

13. Lewis JD, Schinnar R, Bilker WB, Wang X, Strom BL. Validation studies of The Health Improvement Network (THIN) database for pharmacoepidemiology research. Pharmacoepidemiol Drug Saf. 2007;16:393–401.

14. Denburg MR, Haynes K, Shults J, Lewis JD, Leonard MB. Validation of The Health Improvement Network (THIN) database for epidemiologic studies of chronic kidney disease. Pharmacoepidemiol Drug Saf. 2011;20:1138–49.

15. Herrett E, Gallagher AM, Bhaskaran K, et al. Data resource profile: clinical practice research datalink (CPRD). Int J Epidemiol. 2015;44:827–36.

16. Becher H, Kostev K, Schröder-Bernhardi D. Validity and representativeness of the "Disease Analyzer" patient database for use in pharmacoepidemiological and pharmacoeconomic studies. Int J Clin Pharmacol Ther. 2009;47:617–26.

17. Laforest L, Licaj I, Devouassoux G, et al. Prescribed therapy for asthma: therapeutic ratios and outcomes. BMC Fam Pract. 2015;16:49.

18. Pacurariu A, Plueschke K, McGettigan P, et al. Electronic healthcare databases in Europe: descriptive analysis of characteristics and potential for use in medicines regulation. BMJ Open. 2018;8: e023090.

19. Gibson PG, McDonald VM. Asthma–COPD overlap 2015: now we are six. Thorax. 2015;70:683–91.

20. Roca J, Burgos F, Sunyer J, et al. References values for forced spirometry. Group of the European Community Respiratory Health Survey. Eur Respir J. 1998;11:1354–62.

21. Rothnie KJ, Müllerová H, Hurst JR, et al. Validation of the recording of acute exacerbations of COPD in UK primary care electronic healthcare records. PLoS ONE. 2016;11:e0151357.

22. Rothnie KJ, Müllerová H, Thomas SL, et al. Recording of hospitalizations for acute exacerbations of COPD in UK electronic health care records. Clin Epidemiol. 2016;8:771–82.

23. Cacace M, Ettelt S, Mays N, Nolte E. Assessing quality in cross-country comparisons of health systems and policies: towards a set of generic quality criteria. Health Policy. 2013;112:156–62.

24. Macaulay D, Sun SX, Sorg RA, et al. Development and validation of a claims-based prediction model for COPD severity. Respir Med. 2013;107:1568–77.

25. Goossens LMA, Baker CL, Monz BU, Zou KH, Rutten-van Mölken MPMH. Adjusting for COPD severity in database research: developing and validating an algorithm. Int J Chron Obstruct Pulmon Dis. 2011;6:669–78.

26. Greulich T, Weist BJD, Koczulla AR, et al. Prevalence of comorbidities in COPD patients by disease severity in a German population. Respir Med. 2017;132:132–8.

27. Raherison C, Ouaalaya E-H, Bernady A, et al. Comorbidities and COPD severity in a clinic-based cohort. BMC Pulm Med. 2018;18:117.

28. Fernández-Villar A, Soriano JB, López-Campos JL. Overdiagnosis of COPD: precise definitions and proposals for improvement. Br J Gen Pract. 2017;67: 183–4.

29. Sator L, Horner A, Studnicka M, et al. Overdiagnosis of COPD in subjects with unobstructed spirometry: a BOLD analysis. Chest. 2019;156:277–88.

30. Spero K, Bayasi G, Beaudry L, Barber KR, Khorfan F. Overdiagnosis of COPD in hospitalized patients. Int J Chron Obstruct Pulmon Dis. 2017;12:2417–23.

31. White P, Thornton H, Pinnock H, Georgopoulou S, Booth PP. Overtreatment of COPD with inhaled corticosteroids–implications for safety and costs: cross-sectional observational study. PLoS ONE. 2013;8:e75221.

32. Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the general practice research database: a systematic review. Br J Clin Pharmacol. 2010;69:4–14.

33. Jones PW, Lamarca R, Chuecos F, et al. Characterisation and impact of reported and unreported exacerbations: results from ATTAIN. Eur Respir J. 2014;44:1156–65.

34. Moore WC, Fitzpatrick AM, Li X, et al. Clinical heterogeneity in the severe asthma research program. Ann Am Thorac Soc. 2013;10(Suppl): S118–24.

35. Vestbo J, Agusti A, Wouters EFM, et al. Should we view chronic obstructive pulmonary disease differently after ECLIPSE? A clinical perspective from the study team. Am J Respir Crit Care Med. 2014;189: 1022–30.

36. Alshabanat A, Zafari Z, Albanyan O, Dairi M, FitzGerald JM. Asthma and COPD overlap syndrome (ACOS): a systematic review and meta analysis. PLoS ONE. 2015;10:e0136065.

37. Bateman ED, Reddel HK, van Zyl-Smit RN, Agusti A. The asthma–COPD overlap syndrome: towards a revised taxonomy of chronic airways diseases? Lancet Respir Med. 2015;3:719–28.

38. Nikolaou V, Massaro S, Fakhimi M, Stergioulas L, Price D. COPD phenotypes and machine learning cluster analysis: A systematic review and future research agenda. Respir Med. 2020;171:106093.

39. Battisti WP, Wager E, Baltzer L, et al. Good publication practice for communicating company-sponsored medical research: GPP3. Ann Intern Med. 2015;163:461–4.