ORIGINAL RESEARCH

Development and Validation of a Healthcare Utilization-Based Algorithm to Identify Acute Exacerbations of Chronic Obstructive Pulmonary Disease

Douglas W Mapel [1],*
Melissa H Roberts [1],*
Susan Sama [2]
Priyanka J Bobbili [3]
Wendy Y Cheng [3]
Mei Sheng Duh [3]
Catherine Nguyen [3]
Philippe Thompson-Leduc [3]
Melissa K Van Dyke [4]
Kieran J Rothnie [4]
Devi Sundaresan [2]
Julia M Certa [5]
Thomas S Whiting [5]
Jennifer L Brown [5]
Douglas W Roblin [5]

[1]College of Pharmacy, University of New Mexico, Albuquerque, NM, USA; [2]Reliant Medical Group, Inc., Worcester, MA, USA; [3]Analysis Group, Inc., Boston, MA, USA; [4]GlaxoSmithKline plc., London, UK; [5]Kaiser Permanente Mid-Atlantic States (KPMAS), Mid-Atlantic Permanente Research Institute (MAPRI), Rockville, MD, USA

*These authors contributed equally to this work

Correspondence: Melissa H Roberts
College of Pharmacy, University of New Mexico, 2502 Marble Ave, Albuquerque, NM, 87106, USA
Tel +1 505 925-0953
Email MHRoberts@salud.unm.edu

Introduction: Acute exacerbations of chronic obstructive pulmonary disease (AECOPD) are important events that may precipitate other adverse outcomes. Accurate AECOPD event identification in electronic administrative data is essential for improving population health surveillance and practice management.

Objective: Develop codified algorithms to identify moderate and severe AECOPD in two US healthcare systems using administrative data and electronic medical records, and validate their performance by calculating positive predictive value (PPV) and negative predictive value (NPV).

Methods: Data from two large regional integrated health systems were used. Eligible patients were identified using International Classification of Diseases (Ninth Edition) COPD diagnosis codes. Two algorithms were developed: one to identify potential moderate AECOPD by selecting outpatient/emergency visits associated with AECOPD-related codes and antibiotic/systemic steroid prescriptions; the other to identify potential severe AECOPD by selecting inpatient visits associated with corresponding codes. Algorithms were validated via patient chart review, adjudicated by a pulmonologist. To estimate PPV, 300 potential moderate AECOPD and 250 potential severe AECOPD events underwent review. To estimate NPV, 200 patients without any AECOPD identified by the algorithms (100 patients each without moderate or severe AECOPD) during the two years following the index date underwent review to identify AECOPD missed by the algorithm (false negatives).

Results: The PPVs (95% confidence interval [CI]) for both moderate and severe AECOPD were high: 293/298 (98.3% [96.1–99.5]) and 216/225 (96.0% [92.5–98.2]), respectively. NPV was lower for moderate AECOPD (75.0% [65.3–83.1]) than for severe AECOPD (95.0% [88.7–98.4]). Results were consistent across both healthcare systems.

Conclusion: This study developed healthcare utilization-based algorithms to identify moderate and severe AECOPD in two separate healthcare systems. PPV for both algorithms was high; NPV was lower for the moderate algorithm. Replication and consistency of results across two healthcare systems support the external validity of these findings.

Keywords: claims database, algorithm, validation, electronic medical records, predictive values, AECOPD events

Plain Language Summary

Acute exacerbations of chronic obstructive pulmonary disease (AECOPD events) are a sudden worsening of a patient's respiratory symptoms that is beyond normal day-to-day variations and leads to a change in medication. However, there remains no clear consensus

on how best to identify these exacerbations using administrative data, such as electronic records of medical claims. This is an important issue, as accurate identification of AECOPD events in electronic administrative data is essential for improving population health surveillance and practice management.

We developed two algorithms that could be applied to administrative data; one of these algorithms aimed to identify moderate AECOPD events, while the other aimed to identify severe AECOPD events. In this study spanning between January 1, 2010 and September 30, 2015, we used these newly developed algorithms to identify moderate and severe AECOPD events within two large regional integrated healthcare systems. We then assessed the performance of the algorithms by calculating both their positive and negative predictive values (PPV and NPV). These statistical values represent the probability of a positive or negative identification being accurate. To calculate predictive values, the results of each algorithm were compared with AECOPD events identified in medical records – a process that was adjudicated by an expert in the field.

Our results show that, in terms of both PPV and NPV, the developed algorithms can identify moderate and severe AECOPD events in administrative databases with sufficient validity. The proposed algorithms could therefore prove beneficial in comparative effectiveness research, health plan quality improvement projects, and practice management.

Introduction

Chronic obstructive pulmonary disease (COPD), a condition characterized by persistent airflow limitation and associated with enhanced inflammatory response of the lungs to irritants,[1] is a significant burden to both individuals and healthcare systems. The overall prevalence of COPD among US adults is estimated to be 6.2%[2] and is approximately 12% among Medicare beneficiaries.[3]

The Global Initiative for Chronic Obstructive Lung Disease (GOLD) committee defines an acute exacerbation of COPD (AECOPD) as "an acute worsening of respiratory symptoms that results in additional therapy."[4] Exacerbations are often classified by severity: typically, "moderate" exacerbations are defined as those that require a COPD-related emergency department (ED) visit or outpatient visit with systemic corticosteroid treatment and/or antibiotic therapy, whilst "severe" exacerbations are those requiring hospitalization.[4] Exacerbations are associated with a significant increase in mortality and healthcare resource utilization in the form of hospitalizations, ED visits, urgent care visits, and prescription medications.[5–7] Patients suffering from more frequent exacerbations can experience acute irreversible pulmonary function loss and

reduced quality of life.[8–11] Preventing exacerbations is therefore a primary goal in COPD treatment.[1]

Administrative data – computerized systems of healthcare events where healthcare utilization is summarized using standardized codes (eg, diagnosis, procedure, prescriptions) – are the predominant source of real-world healthcare data in the US and are used for health services, as well as pharmacological and epidemiological studies of COPD. These systems include claims data for fee-for-service billing and encounter data for managed care entities receiving capitated payments. While limited to insured individuals who access healthcare,[12,13] administrative data provide valuable information about the healthcare utilization experience of individuals with COPD and real-world management. The notion of AECOPD as a worsening of underlying COPD has its underpinnings in changes in healthcare utilization;[14,15] however, there remains no consensus on how best to identify exacerbations using administrative data.[14,16] This lack of consensus, across the range of mild to severe, represents a significant research barrier[16–18] for a retrospective analysis of large healthcare databases where detailed clinical outcomes of the study population are not always available as indicators of exacerbation events. The capability of large databases to identify patients who may benefit from COPD interventions depends on correct identification of AECOPD. Practical methods are needed to enable this identification to be performed with confidence.[19]

The aims of this study were to develop two utilization-based algorithms that could be applied to administrative data to identify moderate and severe AECOPD and to validate their performance.

Materials and Methods
Study Design

This was a retrospective longitudinal cohort study spanning January 1, 2010 – September 30, 2015. The index date was defined as the first prescription of GOLD[1] guideline-recommended therapy for COPD (Supplementary Table 1), during or following at least one hospitalization (minimum length of stay ≥2 days), one ED visit, or two outpatient visits, associated with COPD diagnoses codes as identified from utilization data. There was a baseline period of 12 months prior to the index date during which inclusion and exclusion criteria were applied (Figure 1). A patient's observation period began with the index date and ended with the earliest date of death, health plan disenrollment, or end of study
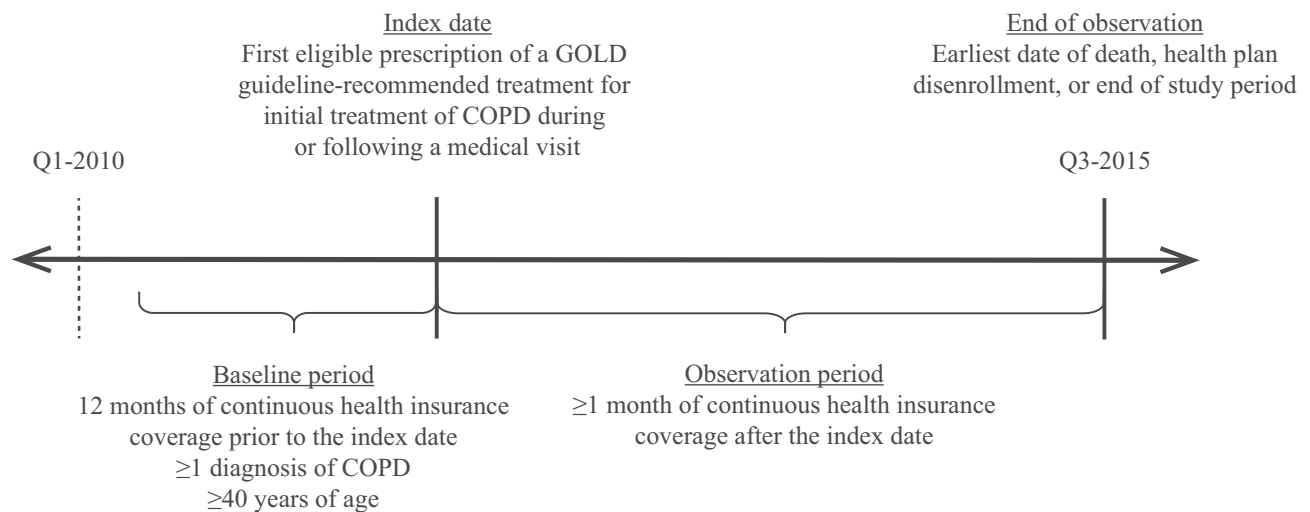
Index date
First eligible prescription of a GOLD
guideline-recommended treatment for
initial treatment of COPD during
or following a medical visit

End of observation
Earliest date of death, health plan
disenrollment, or end of study period

Q1-2010

Q3-2015

Baseline period
12 months of continuous health insurance
coverage prior to the index date
≥1 diagnosis of COPD
≥40 years of age

Observation period
≥1 month of continuous health insurance
coverage after the index date

Figure 1 Study design.
Note: This diagram is not to scale.
Abbreviations: COPD, chronic obstructive pulmonary disease; GOLD, Global Initiative for Chronic Obstructive Lung Disease; Q, quarter.

period. Potential AECOPD events were identified during the observation period using diagnosis codes from the International Classification of Disease, Ninth Revision, Clinical Modification (ICD-9-CM).

Study Population

Patients aged ≥40 years with ≥1 hospitalization, ≥1 ED visit, or ≥2 outpatient visits with a primary or secondary COPD diagnosis (ICD-9-CM codes: 491.xx, 492.xx, and 496) during the 12-month baseline period were included in the study. Additional inclusion criteria were: ≥1 prescription for any GOLD guideline-recommended maintenance therapy for COPD (stand-alone or in combination); ≥12 months of continuous health insurance coverage prior to the index date; ≥1 month of continuous health insurance coverage after the index date; and availability of electronic medical records (EMR; inclusive of claims for payment from contract healthcare providers) data. Patients were excluded if they had any of the following comorbid conditions during baseline: active pulmonary tuberculosis, sarcoidosis, respiratory cancer, cystic fibrosis, extrinsic allergic alveolitis, pneumoconiosis, pulmonary congestion and hypostasis, other alveolar and parietoalveolar pneumonopathy, lung involvement in conditions classified elsewhere, pulmonary collapse, pulmonary eosinophilia, or allergic bronchopulmonary aspergillosis.

Data Sources

Health insurance encounters and EMR data were obtained from two independent health systems: Kaiser Permanente Mid-Atlantic States (KPMAS) and Reliant Medical Group, Inc. (Reliant). KPMAS is an integrated health system, comprised of over 1600 providers practicing in 33 medical centers in Baltimore, Washington, D.C., Northern Virginia, and Southern Maryland, USA, serving more than 770,000 members. Reliant is the largest private, multi-specialty group in central Massachusetts, USA, providing comprehensive care for more than one million patient visits a year with more than 300 providers practicing in over 20 locations.

Study Outcomes

Baseline Characteristics and Algorithm Development

Baseline characteristics, including patient demographics and medical history, were collected for patients meeting inclusion criteria. In addition, the prevalence of diagnosis and procedure codes, as well as healthcare utilization data commonly used to define COPD and AECOPD events, were tabulated for the observation period. Based on the research team's expertise in data related to COPD, previously published algorithms[19–25] were modified to develop separate algorithms to identify "moderate" and "severe" AECOPD events in administrative data.

Algorithm Performance by Positive Predictive Value

Of the potential AECOPD events identified with the utilization-based algorithms, random samples of 550 (300 moderate and 250 severe) were selected for detailed medical chart review to validate whether the potential AECOPD was a true event. Validation of event identification and classification was based on review and abstraction of computerized medical records by trained research associates at KPMAS and experienced pulmonary nurses at Reliant.

A standardized abstraction instrument was designed to systematically capture COPD symptom changes, treatment changes and other utilization associated with the event, and spirometry and other clinical data supporting the diagnosis of COPD (Supplementary Table 2). Definitions of validated (true positive) AECOPD events were based on the GOLD COPD 2017 description of AECOPD and associated utilization. An event was classified as "definite" AECOPD if it met both major criteria: 1) acute worsening of COPD symptoms; and 2) a change in treatment such as a new prescription of antibiotics and/or corticosteroids or increase in respiratory medications. An event was classified as "probable" AECOPD if it met one of the major criteria, plus ≥4 minor criteria associated with utilization for AECOPD events (Supplementary Table 3). An event was classified as "possible" AECOPD if it only met ≥4 of the minor criteria.

Events were adjudicated by a pulmonologist with experience in COPD studies. EMR abstractors reviewed events that did not meet the definition of "definite" AECOPD events to identify potential reasons they were not fully confirmed. Positive predictive value (PPV) was computed using the true and false positives (where the unit of analysis was the AECOPD event) initially identified by medical record review and abstraction and then adjudicated by the pulmonologist.

Spirometry data were collected whenever available to further confirm validity of the COPD diagnosis. If multiple tests were performed, the one closest to the AECOPD event was selected. "COPD verified" patients had: a ratio of forced expiratory volume over 1 second (FEV_1) to forced vital capacity (FVC) of <0.70; "COPD unverified" patients had an FEV_1/FVC ratio ≥0.70 (these patients typically had reduced percent predicted FEV_1 and FVC); spirometry results could not be found for "COPD undefined" patients.

Algorithm Performance by Negative Predictive Value

Random samples of 100 patients who were identified by the algorithms as having no AECOPD during the observation period were selected to assess negative predictive value (NPV). Medical records were reviewed to find one potential moderate or severe event missed by the algorithm. If found, the AECOPD event was abstracted using the standardized form and adjudicated to determine if a true event had been missed. NPV was then calculated, with patients as the unit of analysis.

Results
Study Population

The sample selection flowcharts for Reliant and KPMAS data are presented in Figure 2. After applying inclusion and exclusion criteria, 2366 patients from Reliant and 5548 patients from KPMAS were selected.

Baseline Characteristics and Algorithm Development

Patient baseline characteristics are presented in Table 1. Patients were continuously enrolled in their healthcare plan for an average of 2.6 years following the index date at Reliant and 2.0 years at KPMAS. At both sites, the most common (≥20% of patients) medication classes initiated on the index date were short-acting β2 agonists (SABA), inhaled corticosteroids (ICS)/long-acting β2 agonists, and short-acting muscarinic antagonists.

Table 2 displays the frequencies of healthcare utilization and respiratory conditions during baseline. The proportions of patients who were hospitalized ≥2 days were very similar (Reliant: 24.0%, KPMAS: 24.6%). Between the two sites, however, utilization for ED visits, observation visits, and short hospital stays (length of stay = 1 day) varied greatly, presumably representing differences in how intermediate-level services were classified or delivered. Most patients had at least one outpatient visit (Reliant: 94.2%, KPMAS: 98.8%) and use of other services (Reliant: 96.4%, KPMAS: 97.3%) during baseline. The most common respiratory condition at baseline was "Chronic airway obstruction, not elsewhere classified (496)" (Reliant: 92.4%, KPMAS: 85.7%).

The final two algorithms contained utilization-based information deemed clinically relevant in identifying patients with moderate or severe exacerbations (Table 3). For example, key components of the algorithm used to

Table 1 Baseline Characteristics at Index Date

	Reliant	KPMAS
Number of patients	N=2366	N=5548
Females, n (%)	1338 (56.6)	3032 (54.7)
Age at index date (years), mean ± SD [median]	72.7 ± 11.5 [74.0]	68.9 ± 10.7 [68.7]
Length of follow-up (years), mean ± SD [median]	2.6 ± 1.5 [2.6]	2.0 ± 1.4 [1.7]
Year of index year, n (%) 2011 2012 2013 2014 2015	1222 (51.6) 461 (19.5) 292 (12.3) 196 (8.3) 195 (8.2)	1204 (21.7) 1063 (19.2) 1145 (20.6) 1002 (18.1) 1134 (20.4)
Short-acting inhaled β2-agonists, n (%)	1041 (44.0)	2490 (44.9)
ICS/LABA, n (%)	543 (23.0)	1414 (25.5)
Short-acting muscarinic antagonists, n (%)	479 (20.2)	1618 (29.2)
Long-acting muscarinic antagonists, n (%)	232 (9.8)	823 (14.8)
Inhaled corticosteroids, n (%)	283 (12.0)	898 (16.2)

Abbreviations: ICS, inhaled corticosteroids; KPMAS, Kaiser Permanente Mid-Atlantic States; LABA, long-acting beta-agonists; Reliant, Reliant Medical Group, Inc.; SD, standard deviation.

identify moderate events were ≥1 ED or outpatient visit relating to respiratory comorbidities or symptoms, along with ≥1 prescription of antibiotics or systemic steroid medications. For the algorithm to identify severe events, key components included ≥1 hospitalization relating to respiratory comorbidities or symptoms.

Common Diagnoses and Use of Services During Observation Period

The frequencies of healthcare utilization and respiratory conditions based on diagnosis codes during the period in which the algorithms were applied are presented in Table 4. Most patients had ≥1 outpatient visit (Reliant: 92.9%, KPMAS: 96.5%) and used other services (Reliant: 96.9%, KPMAS: 95.7%) during observation. Utilization related to acute events (ED visits and hospitalizations) was higher in the observation period (vs the baseline), and correspondingly more patients received the diagnoses summarized in Table 4.

Validation of Algorithm Using Positive Predictive Value

Validation results for moderate and severe AECOPD algorithms are presented in Tables 5 and 6. Among the moderate AECOPD events, 98.3% (293 of 298) identified by the algorithm were true positives (Table 5). Among the 225 suspected severe exacerbations identified by the algorithm, 216 (96.0%) were true positives (Table 6). The PPV (95% confidence interval [CI]) for severe exacerbations was lowest for the diagnosis "Chronic airway obstruction, not elsewhere classified" (PPV: 94.0% [95% CI: 86.7–98.0]). When evaluated by COPD status (verified, unverified, or undefined, dependent on spirometry results), the PPV was ≥93.8% across all categories for both the moderate and severe algorithms (Supplementary Tables 4 and 5, respectively).
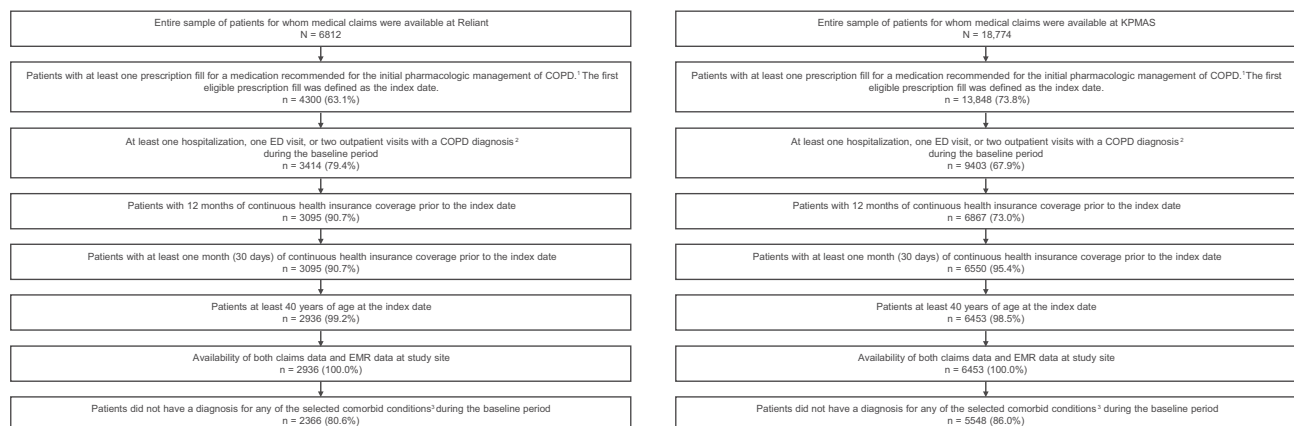
Figure 2 Patient flow diagrams.
Abbreviations: COPD, chronic obstructive pulmonary disease; ED, emergency department; EMR, electronic medical record; KPMAS, Kaiser Permanente Mid-Atlantic States; Reliant, Reliant Medical Group, Inc.

Table 2 Healthcare Utilization and Clinical Characteristics in the Baseline Period

	Reliant	KPMAS
	N=2366	N=5548
Number of patients (%) with ≥1		
Outpatient visit	2229 (94.2)	5481 (98.8)
Emergency department visit	294 (12.4)	2650 (47.8)
Observation visit	30 (1.3)	628 (11.3)
LoS = 1 day	1350 (57.1)	538 (9.7)
LoS ≥2 days	568 (24.0)	1364 (24.6)
Other services*	2280 (96.4)	5400 (97.3)
Number (%) of patients with ≥1 diagnosis (ICD-9-CM code)†		
Chronic airway obstruction, not elsewhere classified (496)	2187 (92.4)	4752 (85.7)
Chronic bronchitis (491.2x)	1021 (43.2)	1629 (29.4)
Pneumonia, organism unspecified (486)	445 (18.8)	892 (16.1)
Emphysema, not interstitial or compensatory (492.xx)	402 (17.0)	998 (18.0)
Acute bronchitis and bronchiolitis (466.xx)	208 (8.8)	385 (6.9)
Asthma, unspecified type with acute exacerbation (493.92)	182 (7.7)	382 (6.9)
Bronchitis, not specified as acute or chronic (490)	154 (6.5)	1225 (22.1)
Chronic obstructive asthma, unspecified (493.20)	103 (4.4)	521 (9.4)
Acute respiratory failure (518.81)	101 (4.3)	296 (5.3)
Chronic obstructive asthma with acute exacerbation (493.22)	67 (2.8)	495 (8.9)

Notes: *Other services included laboratory services, hospice claims, home health claims, skilled nursing facility claims, nursing home claims, telephone, and e-mail services. †Respiratory conditions that occurred in ≥5% of patients.
Abbreviations: ICD-9-CM, International Classification of Diseases, Ninth Revision, Clinical Modification; KPMAS, Kaiser Permanente Mid-Atlantic States; LoS, length of stay; Reliant, Reliant Medical Group, Inc.

Validation of Algorithm Using Negative Predictive Value

Performance of the algorithm was assessed by calculating the NPV (Table 7). NPV was lower for moderate exacerbations compared with severe exacerbations (75.0% [95% CI: 65.3–83.1] vs 95.0% [95% CI: 88.7–98.4], respectively). The "moderate" NPV for Reliant was 72.0% (95% CI: 57.5–83.8) and for KPMAS, 78.0% (95% CI: 64.0–88.5). The "severe" NPV value was 90.0% (95% CI: 78.2–96.7) for Reliant and 100% (95% CI: 92.9–100.0) for KPMAS.

Discussion

This study demonstrates that structured, computerized algorithms can identify moderate or severe AECOPD events in administrative databases with sufficient validity for use in health outcomes research and quality improvement projects.[26] We also show that the algorithms work equally well for COPD patients who had documented airflow obstruction on pulmonary function tests (PFTs), and those who did not have PFT results available or those who had a FEV_1/FVC ratio >0.70.

The final algorithms build on previously published algorithms and are based on coded values deemed clinically relevant in identifying patients with moderate or severe exacerbations.[19–25] As part of the process for our algorithm development, we incorporated a novel system for validating AECOPD events. Validation was based on the concept that an exacerbation is an acute worsening of respiratory symptoms that results in a change of treatment (major criteria), but also considers clinical factors associated with AECOPD events (minor criteria), such as signs of COPD on radiographs or PFTs. Healthcare utilization such as sputum cultures and blood gases (construct validity) are often associated with AECOPD events, and provider documentation verified that the event was an AECOPD episode and not some other disease process such as heart failure (clinical validation).

Our algorithms build on work from previous studies that examined the validity of using ICD-9 codes to identify AECOPD events in healthcare administrative databases. Stein et al[27] examined the performance characteristics of four different algorithms, using ICD-9 coding, to identify AECOPD hospitalization events at two urban academic medical centers. The study used a stratified probability sample of 200 hospitalizations with and without COPD. Interestingly, they found that the algorithms' PPV ranged from 81.2% to 97.2%, but the algorithm that was most comprehensive and similar to ours was at the low end of this range, while our PPV for severe events was 96.0%. The NPV for their severe algorithms ranged from 93.0% to 93.9%, similar to our 95.0%. Oelsner et al[28] also used ICD-9 coding to identify 216 patients who had likely "Chronic Lower Respiratory Disease" events, which encompassed exacerbations of COPD, chronic bronchitis, emphysema, or asthma that resulted in ED visits or hospitalizations. As with our study, they established a priori

Table 3 Algorithms for Moderate and Severe AECOPD Exacerbations Used in Round 1

Algorithm 1: Moderate Exacerbations
At least 1 office or outpatient non-emergency department visit with any of the following ICD-9-CM diagnosis codes as the primary diagnosis*:
Acute bronchitis and bronchiolitis (466)
Pneumonia, organism unspecified (486)
Bronchitis, not specified as acute or chronic (490)
Chronic bronchitis (491)
Emphysema, not interstitial or compensatory (492.xx)
Chronic obstructive asthma, unspecified (493.20)
Chronic obstructive asthma with acute exacerbation (493.22)
Asthma, unspecified type, with acute exacerbation (493.92)
Chronic airway obstruction, not elsewhere classified (496)
Acute respiratory failure (518.81)[†]
Other pulmonary insufficiency not elsewhere classified (518.82)[†]
Acute and chronic respiratory failure (518.84)[†]
Dyspnea and respiratory abnormalities (786.0)
Shortness of breath (786.05)
Cough (786.2)
Chest pain (786.5)
Wheezing (786.07)
Asphyxia and hypoxemia (799.0)
AND
At least 1 associated pharmacy dispensing for theophylline (intravenous) or aminophylline (intravenous) or systemic steroids (all routes of administration except topical, ear drops, and eye drops) or any the following antibiotics (all routes of administration except for topical, ear drops, and eye drops) commonly used for respiratory infections within 14 days of the first claim associated with an exacerbation
Amoxicillin (alone or in combination), beta-lactamase inhibitors, second to fourth-generation cephalosporins, macrolides, or doxycycline
Algorithm 2: Severe Exacerbations
At least 1 inpatient hospital stay of two or more days with any of the following ICD-9-CM diagnosis codes as the primary diagnosis:
Chronic bronchitis (491)
Emphysema (492)
Chronic obstructive asthma, unspecified (493.20)
Chronic obstructive asthma with (acute) exacerbation (493.22)
Chronic airway obstruction, not elsewhere classified (496)
Acute respiratory failure (518.81)[†]
Other pulmonary insufficiency not elsewhere classified (518.82)[†]
Acute and chronic respiratory failure (518.84)[†]

Notes: Based on algorithm in: Mapel et al, Identifying and characterizing COPD patients in US managed care. A retrospective, cross-sectional analysis of administrative claims data. BMC Health Services Research 2011 11:43. *COPD ICD-9-CM codes supplemented by: Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation (YNHHSC/CORE). 2016 Condition-Specific Measures Updates and Specifications Report: Hospital-Level 30-Day Risk-Standardized Readmission Measures. Centers for Medicare and Medicaid Services (CMS), 2016. Available at https://aann.org/uploads/Condition_Specific_Readmission_Measures.pdf. [†]Principal discharge diagnosis codes included in cohort if combined with a secondary diagnosis of COPD with exacerbation (491.21, 491.22, 493.21, or 493.22). [‡]Drugs prescriptions will be identified in claims data using NDC codes, which are too numerous to list in this table.
Abbreviations: AECOPD, acute exacerbation of chronic obstructive pulmonary disease; COPD, chronic obstructive pulmonary disease; ICD-9-CM, International Classification of Diseases, Ninth Revision, Clinical Modification; NDC, National Drug Code.

criteria for "probable", "highly probable", and "definite" events, with adjudication by physician review.[28] They found that the order in which the discharge diagnoses were listed had importance on the validity of the event; for example, the PPV of a "Chronic Lower Respiratory Disease" exacerbation resulting in hospitalization was >95% if a respiratory disease was the primary discharge diagnosis, but <55% if a secondary diagnosis. Stanford et al recently summarized results for a similar exacerbation algorithm effort.[29] Their algorithm used the Stein et al algorithm as a starting point. They did not separate moderate and severe exacerbations, but instead developed a combined algorithm. They reported a lower PPV value (67.5%) than found in our study. However, while our algorithm only considered events for which diagnosis codes were a primary diagnosis, the Stanford et al algorithm included events for which diagnosis codes may have been a primary or secondary diagnosis. Our study adds to current literature by demonstrating the validity of using healthcare utilization captured in large administrative databases, further outlining objective criteria that can be used to define and confirm AECOPD events.

Replication and consistency of the PPV results between sites is a strength of this study, bolstering external validity. The lower NPV values were indicative of the broader range of events encompassed by moderate exacerbations – from borderline mild to borderline severe – and the broader range of settings in which these events could be documented. Exacerbations missed by the algorithms in this study were reviewed using administrative and EMR data to understand potential reasons for omission. Reasons included: missing claims (ie, medications or visits not appearing in administrative data), timing (eg, medication dispensed before the medical visit, which the algorithm did not consider as being associated with the visit), or because the algorithm did not consider secondary diagnoses as indications of AECOPD. Missed moderate exacerbations often occurred near the index date, which was based on the first prescription fill for a COPD maintenance medication; if a respiratory problem was not listed

Table 4 Healthcare Utilization and Clinical Characteristics in the Observation Period

	Reliant	KPMAS
	N=2366	**N=5548**
Number of patients (%) with ≥1		
Outpatient visit	2199 (92.9)	5356 (96.5)
Emergency department visit	896 (37.9)	3591 (64.7)
Observation visit	113 (4.8)	1115 (20.1)
LoS = 1 day	1623 (68.6)	1017 (18.3)
LoS ≥2 days	1194 (50.5)	1959 (35.3)
Other services*	2293 (96.9)	5308 (95.7)
Number (%) of patients with ≥1 diagnosis (ICD-9-CM code)[†]		
Chronic airway obstruction, not elsewhere classified (496)	2180 (92.1)	4207 (75.8)
Chronic bronchitis (491.2x)	1458 (61.6)	1875 (33.8)
Pneumonia, organism unspecified (486)	1003 (42.4)	1424 (25.7)
Emphysema, not interstitial or compensatory (492.xx)	831 (35.1)	1238 (22.3)
Acute respiratory failure (518.81)	503 (21.3)	800 (14.4)
Chronic obstructive asthma, unspecified (493.20)	306 (12.9)	702 (12.7)
Acute and chronic respiratory failure (518.84)	280 (11.8)	313 (5.6)
Acute bronchitis and bronchiolitis (466.xx)	270 (11.4)	396 (7.1)
Bronchitis, not specified as acute or chronic (490)	240 (10.1)	1387 (25.0)
Other bacterial pneumonia (482)	192 (8.1)	207 (3.7)
Asthma, unspecified type with acute exacerbation (493.92)	195 (8.2)	430 (7.8)
Other pulmonary insufficiency, not elsewhere classified (518.82)	166 (7.0)	177 (3.2)
Chronic obstructive asthma with acute exacerbation (493.22)	169 (7.1)	625 (11.3)

Notes: *Other services included laboratory services, hospice claims, home health claims, skilled nursing facility claims, nursing home claims, telephone, and e-mail services. [†]Respiratory conditions that occurred in ≥5% of patients.
Abbreviations: ICD-9-CM, International Classification of Diseases, Ninth Revision, Clinical Modification; KPMAS, Kaiser Permanente Mid-Atlantic States; LoS, length of stay; Reliant, Reliant Medical Group, Inc.

as the primary diagnosis, then these events were missed. These observations suggest potential improvements that may be made to improve the overall performance of this classification system.

Another reason for the relatively low NPV in moderate exacerbations is the location of care (ie, home health services and telehealth services). For instance, Reliant has a COPD disease management group that tries to manage mild/moderate exacerbations via telephone or in person. Many patients treated by this group have pre-filled medications they can begin taking at the onset of symptoms. These patients are instructed to call their disease management nurse when initiating this treatment so that this can be confirmed and documented within the EMR. Services like these are becoming more prevalent, meaning that claims-based algorithms that do not include them have the potential to miss moderate AECOPD events.

PPV was calculated among patients with verified, unverified, and undefined COPD as per spirometry results. Results were robust to the main analysis and the PPV remained high across all three groups, suggesting the algorithm performs well without the need for spirometry results. Lack of PFT data in most administrative claims databases is frequently cited as a weakness, but this analysis suggests that lack has little appreciable impact on algorithm performance. It could be that the high use of PFTs and chest computed tomography scans in these two systems, compared to previous studies, has resulted in better defined research populations with fewer misclassifications.

Several limitations should be noted. EMRs are organized to collect and present data for clinical support, not necessarily research. Our algorithm was applied to structured codified data as is common in claims and encounter data. Different healthcare systems may prioritize electronic data collection for different purposes; therefore, algorithm performance may differ in other systems. Another limitation is that PPV and NPV were the only measures assessing algorithm performance. However, the sampling required to assess additional metrics (eg, sensitivity, c-statistic, net reclassification improvement, or the incremental discrimination index) were beyond the scope of the study. While the algorithms were robust – being validated in two different healthcare systems – it is not known whether the results can be generalized to more limited administrative databases in the US or elsewhere in the world. ICD-10-CM coding was not in use during the study time period,

Table 5 Validation Results for Moderate COPD Exacerbation Algorithm

	Abstracted	Confirmed*	PPV (95% CI)
Overall (any code)	298	293	98.3 (96.1, 99.5)
By specific ICD-9-CM code:			
Cough (786.2)	63	63	100.0 (94.3, 100.0)
Chronic bronchitis (491.x)	78	77	98.7 (93.1, 100.0)
Chronic airway obstruction, not elsewhere classified (496.x)	82	79	96.3 (89.7, 99.2)
Dyspnea and respiratory abnormalities (786.0)	49	48	98.0 (89.1, 99.9)
Shortness of breath (786.05)	24	23	95.8 (78.9, 99.9)
Acute bronchitis and bronchiolitis (466.x)	10	10	100.0 (69.2, 100.0)
Chest pain (786.5)	17	17	100.0 (80.5, 100.0)
Pneumonia, organism unspecified (486.x)	18	18	100.0 (81.5, 100.0)
Emphysema (492.x)	6	6	100.0 (54.1, 100.0)
Wheezing (786.07)	4	4	100.0 (39.8, 100.0)
Asthma, unspecified type, with acute exacerbation (493.92)	8	8	100.0 (63.1, 100.0)
Chronic obstructive asthma with acute exacerbation (493.22)	5	5	100.0 (47.8, 100.0)
Acute and chronic respiratory failure (518.84)	1	1	100.0 (2.5, 100.0)
Asphyxia and hypoxemia (799.0x)	4	4	100.0 (39.8, 100.0)
Chronic obstructive asthma, unspecified (493.20)	2	2	100.0 (15.8, 100.0)
Acute respiratory failure (518.81)	1	1	100.0 (2.5, 100.0)
Other pulmonary insufficiency not elsewhere classified (518.82)	-	-	-
Bronchitis, not specified as acute or chronic (490.x)	17	17	100.0 (80.5, 100.0)
By specific medication group based on GPI/CPT codes:			
Antibiotics	209	204	97.6 (94.5, 99.2)
Systemic corticosteroids	183	182	99.5 (97.0, 100.0)
Methylxanthines	2	2	100.0 (15.8, 100.0)

Notes: Patients with AECOPD events that had missing or unavailable electronic medical records were excluded from the analysis because adjudication of these events was not possible. *Confirmed true positive AECOPD events included definite, probable, or possible exacerbation events, as classified by the adjudicator, in the primary analysis.
Abbreviations: AECOPD, acute exacerbation of chronic obstructive pulmonary disease; CI, confidence interval; COPD, chronic obstructive pulmonary disease; CPT, current procedural terminology; GPI, generic product identifier; ICD-9-CM, International Classification of Diseases, Ninth Revision, Clinical Modification; PPV, positive predictive value.

Table 6 Validation Results for Severe COPD Exacerbation Algorithm

	Abstracted	Confirmed*	PPV (95% CI)
Overall (any code)	225	216	96.0 (92.5, 98.2)
By specific ICD-9-CM code:			
Chronic bronchitis (491.x)	101	99	98.0 (93.0, 99.8)
Chronic airway obstruction, not elsewhere classified (496.x)	84	79	94.0 (86.7, 98.0)
Acute respiratory failure (518.81)	58	57	98.3 (90.8, 100.0)
Emphysema (492.x)	30	29	96.7 (82.8, 99.9)
Acute and chronic respiratory failure (518.84)	18	18	100.0 (81.5, 100.0)
Other pulmonary insufficiency not elsewhere classified (518.82)	8	8	100.0 (63.1, 100.0)
Chronic obstructive asthma with acute exacerbation (493.22)	8	8	100.0 (63.1, 100.0)
Chronic obstructive asthma, unspecified (493.20)	2	2	100.0 (15.8, 100.0)

Notes: Patients with AECOPD events that had missing or unavailable electronic medical records were excluded from the analysis because adjudication of these events was not possible. *Confirmed true positive AECOPD events included definite, probable, or possible exacerbation events, as classified by the adjudicator, in the primary analysis.
Abbreviations: AECOPD, acute exacerbation of chronic obstructive pulmonary disease; CI, confidence interval; COPD, chronic obstructive pulmonary disease; ICD-9-CM, International Classification of Diseases, Ninth Revision, Clinical Modification; PPV, positive predictive value.

Table 7 Performance of Algorithms by Calculating NPV

Algorithm 1: Moderate Exacerbations			
	Abstracted	Missed	NPV (95% CI)
Compiled	100	25	75.0 (65.3, 83.1)
Reliant	50	14	72.0 (57.5, 83.8)
KPMAS	50	11	78.0 (64.0, 88.5)
Algorithm 2: Severe Exacerbations			
	Abstracted	Missed	NPV (95% CI)
Compiled	100	5	95.0 (88.7, 98.4)
Reliant	50	5	90.0 (78.2, 96.7)
KPMAS	50	0	100.0 (92.9, 100.0)

Notes: For the calculation of NPV, the patient was used as the unit of analysis. Therefore, NPV represents the proportion of patients whose lack of AECOPDs in the claims was confirmed following the adjudication.
Abbreviations: AECOPD, acute exacerbation of chronic obstructive pulmonary disease; CI, confidence interval; KPMAS, Kaiser Permanente Mid-Atlantic States; NPV, negative predictive value; Reliant, Reliant Medical Group, Inc.

which could limit the generalizability of the algorithms for future use. However, though not validated, the authors have created a preliminary crosswalk from ICD-9 to ICD-10 to address this point (Table 3 and Supplementary Tables 6 and 7). Underlying concepts that have been validated remain constant between ICD-9 and ICD-10, and the severe AECOPD codes listed in Supplementary Table 7 closely align with codes used for the CMS COPD Readmission Measure.[30] Finally, since there is no specific ICD-9 diagnosis code for COPD, patients were identified based on COPD-associated diagnoses and GOLD guideline-recommended maintenance therapy. This may have led to the misclassification of patients into the COPD cohort. However, the high PPV observed across all diagnoses (including "496 - Chronic airway obstruction, not elsewhere classified") and the relatively low and expected proportion of patients with concomitant asthma during the observation period indicate that misclassification was unlikely to have affected our conclusions.[31]

Conclusions

In this study, we developed and validated, through medical record review, improved utilization-based algorithms to identify moderate and severe AECOPD in administrative databases at two sites. PPV for both algorithms was very high; NPV was lower for the moderate exacerbation algorithm. Our proposed algorithms could prove useful for comparative effectiveness research, health plan quality improvement projects, practice management, and to increase knowledge on the frequency of COPD exacerbations using large administrative databases. Opportunities to improve our algorithms – such as inclusion of home health services and clinical data (radiology, laboratory, spirometry) which is increasingly being incorporated into computerized databases – should be explored.

Abbreviations

AECOPD, acute exacerbation of chronic obstructive pulmonary disease; CI, confidence interval; COPD, chronic obstructive pulmonary disease; ED, emergency department; EMR, electronic medical record; FEV_1, forced expiratory volume over 1 second; FVC, forced vital capacity; GOLD, Global Initiative for Chronic Obstructive Lung Disease; ICD-9-CM, International Classification of Diseases, Ninth Revision, Clinical Modification; ICD-10-CM, International Classification of Diseases, Tenth Revision, Clinical Modification; ICS, inhaled corticosteroid; KPMAS, Kaiser Permanente Mid-Atlantic States; NPV, negative predictive value; PFT, pulmonary function test; PPV, positive predictive value; Reliant, Reliant Medical Group, Inc.; SABA, short-acting beta-agonist.

Declarations

Trademarks are owned by or licensed to their respective owners (the GSK group of companies).

Ethics Approval

This retrospective study involving no more than minimal risk to patients received Institutional Review Board (IRB) approval inclusive of waiver of informed consent and waiver of HIPAA authorization from KPMAS IRB (#MA17-157) and Reliant Medical Group IRB (#1422EXT).

Acknowledgments

Douglas W Mapel and Melissa H Roberts are co-first authors for this study.

Author Contributions

All authors contributed to data analysis, drafting or revising the article, have agreed on the journal to which the article will be submitted, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

Funding

Disclosure

MKVD and KJR are employees of GlaxoSmithKline plc. and hold shares in the company. PJB, WYC, MSD, CN, and PT-L are employees of Analysis Group, which has received research funding from GlaxoSmithKline plc. to conduct the study. CN reports research entity from Sanofi, Pfizer, Ipsen, Celgene, Amgen, Taiho, Novartis, Takeda, Shire, Vertex, Janssen, Merck, Kiniksa, Intercept, AstraZeneca, and Novo Nordisk, outside the submitted work. SS and DS are employees of Reliant Medical Group, Inc. and part of their salary may have been supported by funding from GlaxoSmithKline plc. DWR, JMC, TSW, and JLB are employees of Kaiser Permanente Mid-Atlantic States and received research funding from GlaxoSmithKline plc. through Analysis Group to conduct the study. MHR and DWM received research funding from Analysis Group for this study, and in the past have received grant funding from AstraZeneca, Boehringer Ingelheim, Endo Pharmaceuticals, GlaxoSmithKline plc., Sunovion, and Pfizer for COPD-related studies. The authors report no other conflicts of interest in this work.

References

1. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease (2020 report). 2020; Available from: https://goldcopd.org/wp-content/uploads/2019/11/GOLD-2020-REPORT-ver1.0wms.pdf. Accessed 20 July, 2020.

2. Wheaton AG, Liu Y, Croft JB, et al. Chronic obstructive pulmonary disease and smoking status - United States, 2017. *MMWR Morb Mortal Wkly Rep*. 2019;68(24):533–538. doi:10.15585/mmwr.mm6824a1

3. Blanchette CM, Gutierrez B, Ory C, Chang E, Akazawa M. Economic burden in direct costs of concomitant chronic obstructive pulmonary disease and asthma in a Medicare Advantage population. *J Manag Care Pharm*. 2008;14(2):176–185. doi:10.18553/jmcp.2008.14.2.176

4. Vogelmeier CF, Criner GJ, Martinez FJ, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease 2017 report. GOLD executive summary. *Am J Respir Crit Care Med*. 2017;195(5):557–582. doi:10.1164/rccm.201701-0218PP

5. Dhamane AD, Moretz C, Zhou Y, et al. COPD exacerbation frequency and its association with health care resource utilization and costs. *Int J Chron Obstruct Pulmon Dis*. 2015;10:2609–2618. doi:10.2147/COPD.S90148

6. Rothnie KJ, Mullerova H, Smeeth L, Quint JK. Natural history of chronic obstructive pulmonary disease exacerbations in a general practice-based population with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2018;198(4):464–471. doi:10.1164/rccm.201710-2029OC

7. Soler-Cataluna JJ, Martinez-Garcia MA, Roman Sanchez P, Salcedo E, Navarro M, Ochando R. Severe acute exacerbations and mortality in patients with chronic obstructive pulmonary disease. *Thorax*. 2005;60(11):925–931. doi:10.1136/thx.2005.040527

8. Donaldson GC, Seemungal TA, Bhowmik A, Wedzicha JA. Relationship between exacerbation frequency and lung function decline in chronic obstructive pulmonary disease. *Thorax*. 2002;57(10):847–852. doi:10.1136/thorax.57.10.847

9. Hurst JR, Vestbo J, Anzueto A, et al. Susceptibility to exacerbation in chronic obstructive pulmonary disease. *N Engl J Med*. 2010;363(12):1128–1138. doi:10.1056/NEJMoa0909883

10. Miravitlles M, Ferrer M, Pont A, et al. Effect of exacerbations on quality of life in patients with chronic obstructive pulmonary disease: a 2 year follow up study. *Thorax*. 2004;59(5):387–395. doi:10.1136/thx.2003.008730

11. Seemungal TA, Donaldson GC, Paul EA, Bestall JC, Jeffries DJ, Wedzicha JA. Effect of exacerbation on quality of life in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 1998;157(5 Pt 1):1418–1422. doi:10.1164/ajrccm.157.5.9709032

12. Seemungal TA, Donaldson GC, Bhowmik A, Jeffries DJ, Wedzicha JA. Time course and recovery of exacerbations in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2000;161(5):1608–1613. doi:10.1164/ajrccm.161.5.9908022

13. Wilkinson TM, Donaldson GC, Hurst JR, Seemungal TA, Wedzicha JA. Early therapy improves outcomes of exacerbations of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2004;169(12):1298–1303. doi:10.1164/rccm.200310-1443OC

14. Burge S, Wedzicha JA. COPD exacerbations: definitions and classifications. *Eur Respir J Suppl*. 2003;41:46s–53s. doi:10.1183/09031936.03.00078002

15. Rodriguez-Roisin R. Toward a consensus definition for COPD exacerbations. *Chest*. 2000;117(5 Suppl 2):398S–401S. doi:10.1378/chest.117.5_suppl_2.398S

16. Montes de Oca M, Laucho-Contreras ME. Is it time to change the definition of acute exacerbation of chronic obstructive pulmornary disease? What do we need to add? *Med Sci (Basel)*. 2018;6:2. doi:10.3390/medsci6020050

17. Pauwels R, Calverley P, Buist AS, et al. COPD exacerbations: the importance of a standard definition. *Respir Med*. 2004;98(2):99–107. doi:10.1016/j.rmed.2003.09.001

18. Stein BD, Charbeneau JT, Lee TA, et al. Hospitalizations for acute exacerbations of chronic obstructive pulmonary disease: how you count matters. *COPD*. 2010;7(3):164–171. doi:10.3109/15412555.2010.481696

19. Mapel DW, Dutro MP, Marton JP, Woodruff K, Make B. Identifying and characterizing COPD patients in US managed care. A retrospective, cross-sectional analysis of administrative claims data. *BMC Health Serv Res*. 2011;11:43. doi:10.1186/1472-6963-11-43

20. Dalal AA, Shah MB, D'Souza AO, Lunacsek OE, Nagar SP, Crater GD. Observational study of the outcomes and costs of initiating maintenance therapies in patients with moderate exacerbations of COPD. *Respir Res*. 2012;13:41. doi:10.1186/1465-9921-13-41

21. Kern DM, Davis J, Williams SA, et al. Comparative effectiveness of budesonide/formoterol combination and fluticasone/salmeterol combination among chronic obstructive pulmonary disease patients new to controller treatment: a US administrative claims database study. *Respir Res*. 2015;16:52. doi:10.1186/s12931-015-0210-x

22. Pasquale MK, Sun SX, Song F, Hartnett HJ, Stemkowski SA. Impact of exacerbations on health care cost and resource utilization in chronic obstructive pulmonary disease patients with chronic bronchitis from a predominantly Medicare population. *Int J Chron Obstruct Pulmon Dis*. 2012;7:757–764. doi:10.2147/COPD.S36997

23. Pasquale MK, Xu Y, Baker CL, et al. COPD exacerbations associated with the modified Medical Research Council scale and COPD assessment test among Humana Medicare members. *Int J Chron Obstruct Pulmon Dis*. 2016;11:111–121. doi:10.2147/COPD.S94323

24. Perera PN, Armstrong EP, Sherrill DL, Skrepnek GH. Acute exacerbations of COPD in the United States: inpatient burden and predictors of costs and mortality. *COPD*. 2012;9(2):131–141. doi:10.3109/15412555.2011.650239

25. Suh DC, Lau H, La HO, Choi IS, Geba GP. Association between incidence of acute exacerbation and medication therapy in patients with COPD. *Curr Med Res Opin*. 2010;26(2):297–306. doi:10.1185/03007990903465926

26. Trevethan R. Sensitivity, specificity, and predictive values: foundations, pliabilities, and pitfalls in research and practice. *Front Public Health*. 2017;5:307. doi:10.3389/fpubh.2017.00307

27. Stein BD, Bautista A, Schumock GT, et al. The validity of International Classification of Diseases, Ninth Revision, Clinical Modification diagnosis codes for identifying patients hospitalized for COPD exacerbations. *Chest*. 2012;141(1):87–93. doi:10.1378/chest.11-0024

28. Oelsner EC, Loehr LR, Henderson AG, et al. Classifying chronic lower respiratory disease events in epidemiologic cohort studies. *Ann Am Thorac Soc*. 2016;13(7):1057–1066. doi:10.1513/AnnalsATS.201601-063OC

29. Stanford RH, Engel-Nitz NM, Bancroft T, et al. The identification and cost of acute chronic obstructive pulmonary disease exacerbations in a United States population healthcare claims database. *COPD*. 2020;17(5):499–508. doi:10.1080/15412555.2020.1817357

30. CMS. COPD readmission measure code specifications. 2020; Available from: https://qualitynet.cms.gov/inpatient/measures/readmission/methodology. Accessed June 03, 2021.

31. Hardin M, Silverman EK, Barr RG, et al. The clinical features of the overlap between COPD and asthma. *Respir Res*. 2011;12(1):127. doi:10.1186/1465-9921-12-127