

# Evaluating Triple Therapy Treatment Pathways in Chronic Obstructive Pulmonary Disease (COPD): A Machine-Learning Predictive Model

Michael Bogart<sup>1</sup>, Yuhang Liu<sup>2</sup>, Todd Oakland<sup>2</sup>, Marjorie Stiegler<sup>1,3</sup> 

<sup>1</sup>Value Evidence and Outcomes, GlaxoSmithKline, Research Triangle Park, NC, USA; <sup>2</sup>GNS Healthcare, Somerville, MA, USA; <sup>3</sup>University of North Carolina Chapel Hill, Chapel Hill, NC, USA

Correspondence: Michael Bogart, GlaxoSmithKline, Five Moore Drive, PO Box 13398, Research Triangle Park, NC, 27709-3398, USA, Tel +19198897413, Email michael.r.bogart@gsk.com

**Purpose:** Inhaled triple therapy (TT) comprising a long-acting muscarinic antagonist, long-acting  $\beta_2$  agonist, and inhaled corticosteroid is recommended for symptomatic chronic obstructive pulmonary disease (COPD) patients, or those at risk of exacerbation. However, it is not well understood which patient characteristics contribute most to future exacerbation risk. This study assessed patient predictors associated with future exacerbation time following initiation of TT.

**Patients and Methods:** This retrospective cohort study used data from the Optum™ Clinformatics™ Data Mart, a large health claims database in the United States. COPD patients who initiated TT between January 2008 and March 2018 (index) were eligible. Patients were required to be aged  $\geq 18$  years at index and have continuous enrollment for the 12 months prior to index (baseline) and the 12 months following index (follow-up). Patients who had received TT during baseline were excluded. Data from eligible patients were analyzed using a reverse engineering forward simulation machine learning platform to predict future COPD exacerbation time.

**Results:** Data from 73,625 patients were included. The model found that prior exacerbation was largely correlated with post-index exacerbation time; patients who had  $\geq 4$  exacerbation episodes during baseline had an average increase of 32.4 days post-index exacerbation, compared with patients with no exacerbations during baseline. Likewise,  $\geq 2$  inpatient visits (effect size 27.1 days), the use of xanthines (effect size 11.5 days), or rheumatoid arthritis (effect size 6.4 days) during baseline were associated with increased exacerbation time. Conversely, diagnosis of anemia (effect size -5.68 days), or oral corticosteroids in the past month (effect size -3.43 days) were associated with reduced exacerbation time.

**Conclusion:** Frequent prior exacerbations, healthcare resource utilization, xanthine use, and rheumatoid arthritis were the strongest factors predicting the future increase of exacerbations. These results improve our understanding of exacerbation risk among COPD patients initiating triple therapy.

**Keywords:** Bayesian modeling, chronic obstructive pulmonary disease, exacerbation, triple therapy, predictive modeling

## Introduction

Chronic obstructive pulmonary disease (COPD) is a progressive respiratory disease, characterized by a high symptom burden and a high risk of exacerbation.<sup>1</sup> Exacerbations, defined as an acute worsening of symptoms, are the main driver of the total COPD-related burden on healthcare systems in the United States (US).<sup>1-3</sup> Several patient factors are associated with an increased risk of COPD exacerbation, including history of previous exacerbation,<sup>4-7</sup> presence of chronic comorbidities,<sup>7-10</sup> COPD-related healthcare resource utilization,<sup>7</sup> and prescription of antibiotics or other medications.<sup>7,11</sup>

Pharmacological management of COPD involves the use of long-acting muscarinic antagonists (LAMA), long-acting  $\beta_2$  agonists (LABA), and/or inhaled corticosteroids (ICS; usually taken in combination with LABA [ICS/LABA]).<sup>1</sup> Pharmacological management typically begins with monotherapy and escalates to dual or triple therapy, as necessary.

Global Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines recommend that symptomatic patients, or those at risk of exacerbation, are escalated to triple therapy to help improve lung function and reduce future exacerbations.<sup>1</sup>

A recent meta-analysis of 21 randomized controlled trials demonstrated a significantly lower rate of moderate/severe exacerbations, improved lung function, and improved quality of life with the use of inhaled triple therapy, as compared with therapy with either LAMA, LAMA/LABA, or ICS/LABA.<sup>12</sup> Evidence has been generated with regard to the timing of triple therapy initiation,<sup>13</sup> adherence and persistence,<sup>14,15</sup> and its clinical efficacy.<sup>1</sup> However, some gaps in the current research remain, such as an understanding of the patient types most likely to respond to triple therapy. In this context, real-world data are important to improve our understanding of treatment pathways in COPD and inform whether earlier initiation of triple therapy may impact patient outcomes.

This study used a proprietary and patented reverse engineering forward simulation (REFS™) machine learning platform to provide further information on the ideal timing and patient types for triple therapy. The REFS™ platform utilizes a Bayesian and ensemble-based approach to rapidly transform real-world data into computer models that can robustly predict the number of COPD exacerbation days experienced by a patient. The Bayesian ensemble approach has several advantages compared with conventional analytical approaches that employ single- and multiple-hypothesis algorithms, which are often time consuming to develop and test. The ensemble approach of the machine learning platform includes a network of at least 128 models which allows improved sampling of the modeling space and an increased predictive power.

This study therefore aimed to identify patient factors associated with future exacerbation time among COPD patients initiating triple therapy by applying a Bayesian ensemble methodology to real-world data from the US. Factors associated with the decision to start triple therapy among COPD patients who initiated mono- or dual therapy were also identified as an exploratory objective.

## Materials and Methods

### Study Design and Objectives

A retrospective cohort study design was applied to data from the Optum™ Clinformatics™ Data Mart (Optum CDM), a US administrative health claims database. Optum CDM contains data on patient medical history, prescription drug history, laboratory data, and healthcare resource utilization. Data from Optum CDM were obtained for the period of January 1, 2007 to March 31, 2019 and analyzed using a proprietary and patented REFS™ machine learning platform.

The primary objective of this study was to identify factors that predict future exacerbation time among COPD patients initiating triple therapy using a REFS™ machine learning platform. The exploratory objective was to identify factors associated with the decision to start triple therapy in COPD patients who initiated mono- or dual therapy.

### Study Population

Patients with a diagnosis of COPD anytime during the available claims history were identified. A diagnosis of COPD was identified using ICD-9-CM codes: 491.xx, 492.x, 493.2x, 496, or ICD-10-CM codes: J41.x, J42, J43.x, J44.x. For the primary objective, COPD patients who initiated triple therapy between January 1, 2008 and March 31, 2018 were eligible. Triple therapy was defined as a  $\geq 1$  day overlap of dispensations for LAMA (tiotropium, aclidinium, or umeclidinium), LABA (arformoterol, formoterol, indacaterol, olodaterol, or salmeterol), and ICS (beclomethasone, budesonide, ciclesonide, flunisolide, fluticasone propionate, fluticasone furoate, or mometasone). The first date on which the patient received overlapping dispensations of LAMA/LABA/ICS was considered the initiation date of triple therapy (defined as the index date). For the exploratory objective, COPD patients who initiated monotherapy (LAMA, LABA, or ICS) or dual therapy (LAMA/LABA, LAMA/ICS, or LABA/ICS), in a single inhaler or multiple inhalers, were eligible. The index date for the exploratory objective was defined as the earliest date of mono- or dual-maintenance therapy initiation during the index date identification period (January 1, 2008 to March 31, 2018).

Patients were required to be  $\geq 18$  years at the index date and have 12 months of continuous enrollment pre-index (baseline period) and post-index (follow-up period), with evidence of COPD (defined as  $\geq 1$  COPD diagnosis code in their

medical claims) in the 12-month baseline period. Post-index continuous enrollment was not required for the time-to-event analyses in the exploratory objective. For the primary objective, patients were excluded from the analysis if they had overlapping dispensing for ICS/LABA/LAMA (either single or multiple inhalers) during the baseline period. For the exploratory objective, patients with prior maintenance therapy (mono or dual) during the baseline period (12 months pre-index) were excluded from the analysis.

## Study Outcomes

For the primary objective, the study outcome was COPD exacerbation time, measured in days, within the 12-month post-index period. The total number of moderate and severe COPD exacerbations within the 12-month post-index period were also identified. A COPD moderate exacerbation event was defined as: a COPD-related emergency department, urgent care, or outpatient visit, with a diagnosis of COPD in any position, and at least one dispensation for antibiotic or oral systemic corticosteroid within 5 days of the visit. A severe COPD exacerbation was defined as a COPD-related inpatient hospitalization (an inpatient hospital stay with a diagnosis of COPD in any position). For each exacerbation event, the episode began on the date of the first exacerbation event and ended after 14 days if no additional moderate or severe exacerbations occurred during the 14-day period. If another exacerbation occurred before the end of the episode, then the episode was extended and reclassified based on the most severe event type (severe > moderate).

For the exploratory objective, the study outcome was the time to initiation of inhaled triple therapy, defined as: at least 1 day of overlap for ICS, LABA, and LAMA delivered in any form (free-dose or fixed-dose combinations: ICS/LABA + LAMA, LAMA/LABA + ICS, ICS + LABA + LABA). The event date was defined as the first date of prescription fill. The outcome in the exploratory objective was treated as a time-to-event variable, with the first qualifying event following index being considered and no subsequent/multiple events taken into account. Censoring included disenrollment from medical/pharmacy benefit, death, or end of study period.

## Data Analysis and Modeling Approach

For descriptive analyses, patient information from the baseline period was assessed to describe the population prior to triple therapy initiation. Patient variables included in the analysis comprised: demographics, comorbidities, pharmaceutical utilization, medical procedures, laboratory, healthcare resource utilization, calendar year of index event, and additional parameters (for example smoking status and number of prior COPD exacerbations). Covariates were described using mean and standard deviation (SD) for continuous variables, or by frequency distribution for categorical variables. Differences in covariate distributions were examined via Fishers exact test for categorical variables, and Wilcoxon-Mann-Whitney test for continuous variables. Hazard ratios (HR) were used as the reporting statistic for the exploratory objective.

The REFS<sup>TM</sup> methodology is an ensemble-based approach (the resulting “model” is a set of models called an ensemble), which has previously been described.<sup>16</sup> The REFS<sup>TM</sup> platform enables unique insights by reverse engineering an ensemble of models directly from data, without a priori hypotheses. Bayesian network inference is used to build models directly from data and subsequently creates an ensemble of models, which typically consists of hundreds to thousands of models.

To produce each model in the ensemble, REFS<sup>TM</sup> evaluates a large number of putative models that each define a specific interaction form between variables. An interaction form can consist of any number of variables, and a wide range of interaction forms are available within the REFS<sup>TM</sup> platform, including linear, log-linear, logistic, multinomial interactions, Poisson, Gaussian, and survival models. Interaction forms are available to handle both discrete and continuous variables, as well as combinations of discrete and continuous variables and countless interactions between them. Tens of billions to trillions of models are proposed and scored for every model that is eventually accepted into the ensemble. To evaluate a model, REFS<sup>TM</sup> computes its Bayesian score, which is a parameter-free score linked to the likelihood of a particular model given the observed data. The score considers the goodness-of-fit of the model and its complexity. For a particular combination of variables and interaction form, the Bayesian score is computed by marginalizing out the model parameters and applying the Bayesian information criterion, which penalizes overly complex

models. REFS<sup>TM</sup> uses a Markov Chain Monte Carlo approach to generate samples from the equilibrium distribution of models weighted by their score. Each step in the chain corresponds to a small local transformation, such as adding or removing a variable or an interaction from the model.

For the predictive analyses, a unified set of covariates was built for all outcomes using patients' 12-month baseline information (including medical, pharmacy, and laboratory claims data). For the primary objective, an ensemble of predictive linear models of exacerbation days was built using the REFS<sup>TM</sup> platform. All available covariates were evaluated as potential predictors of each outcome. Two-way interactions and higher-order terms between covariates were also considered. The effects of covariates on the outcome were estimated by beta coefficients aggregated across the REFS<sup>TM</sup> linear models. Model prediction performance was measured via R-squared for continuous outcome, with the interpretation of the proportion of the variability in the outcome explained by the selected predictors.

For the exploratory objective, one REFS<sup>TM</sup> survival ensemble of models (time to initiation of triple therapy) was used to understand patient characteristics that predict or are associated with triple therapy initiation decision. The effects of covariates on the outcome were estimated by HRs aggregated across the REFS<sup>TM</sup> survival models. Model prediction performance was measured via Harrell's concordance index for both in-sample and out-of-sample data.

Eighty percent of the study cohort was selected at random to train the REFS<sup>TM</sup> models (training data). The remaining 20% of the study cohort was reserved for model validation (test data), to make predictions to obtain model performance.

## Results

### Primary Analysis

In total, data from 73,625 patients were included in the primary analysis. Patient attrition for the primary analysis is outlined in [Figure 1A](#). From a total of 727 potential predictors (and their interactions), the REFS<sup>TM</sup> platform identified 33 predictors and 12 interaction terms that correlated with post-index exacerbation time ([Table 1](#)).

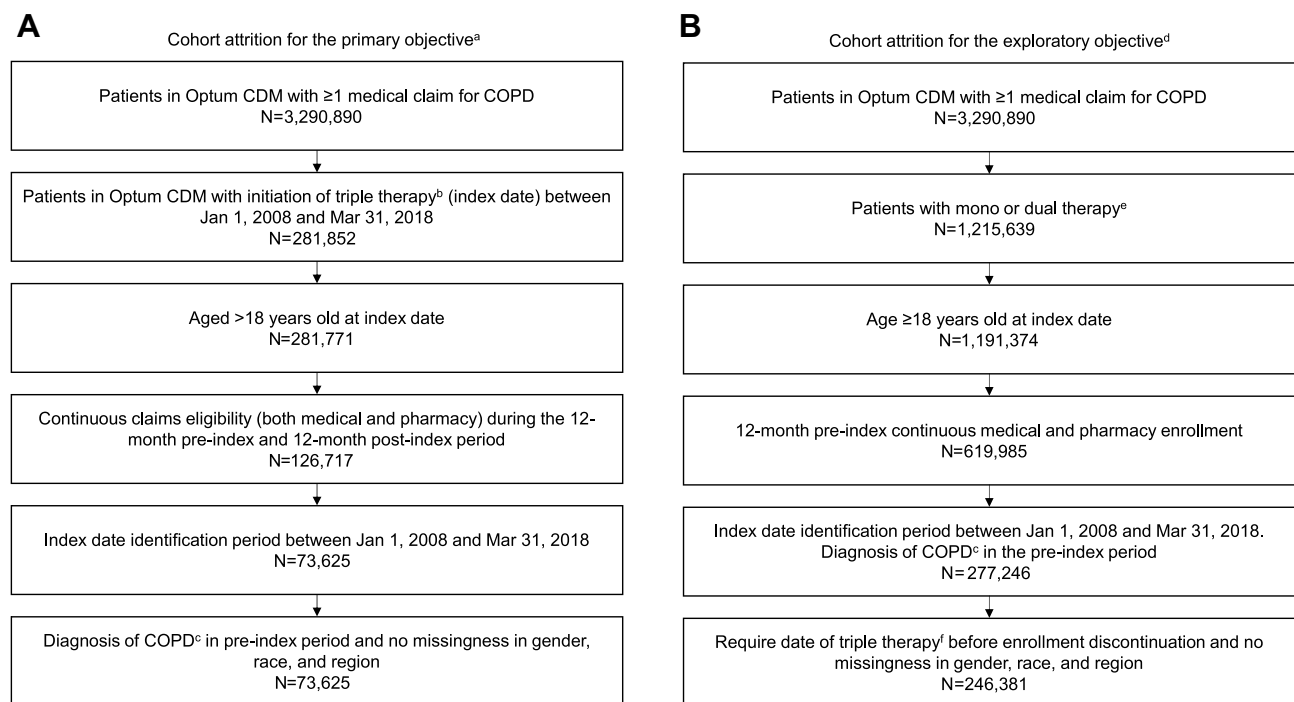
Prior COPD exacerbation was largely correlated with post-index exacerbation time. Patients who had  $\geq 4$  episodes of exacerbation in the baseline period were found to have an average increase of 32 days post-index exacerbation, compared with patients with no exacerbation history in the baseline period. The descriptive analysis reported a similar result, with post-COPD exacerbation time increasing with the number of prior exacerbation episodes ([Figure 2](#)). Similarly, healthcare resource utilization was associated with an increase in post-index COPD exacerbation days;  $\geq 2$  emergency room (ER) visits had an effect size of 2.88 days,  $\geq 2$  inpatient visits had an effect size of 27.05 days, and  $>35$  outpatient visits had an effect size 4.28 days. In comparison, one inpatient visit in the past year had little effect on exacerbation risk (effect size 0.28 days). Among medication usage, the use of xanthines in the baseline period (effect size 11.54 days) was associated with the greatest increase in the number of post-index exacerbation days. A diagnosis of rheumatoid arthritis (RA; effect size 6.39 days) and esophageal disorders (effect size 2.18 days) in the past year was also associated with an increase in post-index COPD exacerbation days.

Variables that were negatively correlated with post-index COPD exacerbation days are shown in [Table 2](#). Among these, diagnosis of acute post-hemorrhagic anemia in the baseline period ( $-5.68$  days) and use of oral corticosteroids in the past month ( $-3.43$  days) had the largest effect size.

Interactions between predictors with a significant effect on exacerbation time are presented in [Table 3](#). The interaction between prior exacerbation episodes ( $\geq 4$  episodes) and oral corticosteroid use in the past month had the largest effect size (average increase of 25.68 days). Although diagnosis of acute post-hemorrhagic anemia was associated with increased COPD exacerbation time, a decreasing trend of exacerbation time for patients with anemia given their inpatient visit levels was observed ([Figure 3](#)).

### Exploratory Objective

In total, data from 246,381 patients were included in the exploratory analysis. Patient attrition is outlined in [Figure 1B](#). From a total of 462 possible predictors (and their interactions), the REFS<sup>TM</sup> platform identified 12 predictors and six interaction terms that correlated with time to triple therapy initiation.



**Figure 1** Patient attrition for (A) the primary analysis and (B) the exploratory analysis.

**Notes:** <sup>a</sup>The primary objective was to identify predictors that contribute the most to COPD exacerbation time among patients initiating triple therapy. <sup>b</sup>Triple therapy defined as  $\geq 1$  day overlap of dispensing days for any LAMA (tiotropium, aclidinium, or umeclidinium), LABA (arformoterol, formoterol, indacaterol, olodaterol, or salmeterol), or ICS (beclomethasone, budesonide, ciclesonide, flunisolide, fluticasone propionate, fluticasone furoate, or mometasone). The first date of a dispensing overlap of the three drug classes was defined as the index date. <sup>c</sup>Defined as  $\geq 1$  COPD diagnosis code in medical claims in the 12-month baseline period. <sup>d</sup>The exploratory objective was to identify factors that are associated with the treatment decision of initiation of triple therapy in patients with COPD who initiate mono- or dual maintenance therapy. <sup>e</sup>Defined as  $\geq 1$  prescription claim for a qualifying COPD maintenance medication (mono or dual therapy) any time during a patient's available claims history between January 1, 2008 and March 31, 2018. For dual therapy, the index date was defined as the first date of a dispensing overlap of any two of the three drug classes. <sup>f</sup>Defined as at least 1 day of overlap for ICS, LABA, and LAMA, delivered in any form: ICS/LABA + LAMA, LABA/LABA + ICS, ICS + LABA + LAMA, or initiation of single-inhaler triple therapy (ICS/LABA/LAMA).

**Abbreviations:** Optum CDM, Optum™ Clinformatics™ Data Mart; COPD, chronic obstructive pulmonary disease.

The variables found to correlate with initiation of triple therapy are shown in Table 4. Factors such as smoking (HR = 1.53) and computed tomography scan of chest (HR = 1.13) were associated with an increased chance of triple therapy initiation. In terms of baseline medication, patients who were receiving LAMA at index had a 3.8 times higher chance of initiating triple therapy compared with patients receiving ICS at index. The use of antibiotics (HR = 0.85), and a diagnosis of asthma (HR = 0.77) were associated with a decreased chance of triple therapy initiation. Interaction terms with significant effects on switching to triple therapy, identified by the model, are shown in Table 5. The interaction of an asthma diagnosis and initiation of either LAMA (HR = 1.55) or LABA/LAMA (HR = 1.82) was associated with an increased chance of triple therapy initiation.

## Discussion

To the authors' knowledge, this is the first study to systematically use a Bayesian ensemble model to corroborate key factors associated with exacerbation time among COPD patients who initiate triple therapy using a large real-world cohort. In line with the findings from previous studies,<sup>5–7</sup> prior COPD exacerbation was shown to be a strong factor in predicting future increase of exacerbation time in patients initiating triple therapy; there was a positive correlation between the number of prior exacerbations and effect size. Healthcare resource utilization, use of xanthines, and concomitant RA were also strong predictors.

The healthcare resource utilization variables with the largest effect size on future exacerbation days were  $\geq 2$  ER visits,  $\geq 2$  inpatient visits, and  $>35$  outpatient visits. Healthcare resource utilization is typically higher among patients with more severe COPD and/or a history of frequent or severe exacerbations, and is associated with a high economic burden and inconvenience to patients.<sup>17</sup> Our findings are consistent with the results from another study that used a machine learning algorithm to predict patient hospitalization due to COPD exacerbation, using real-world data from Swedish



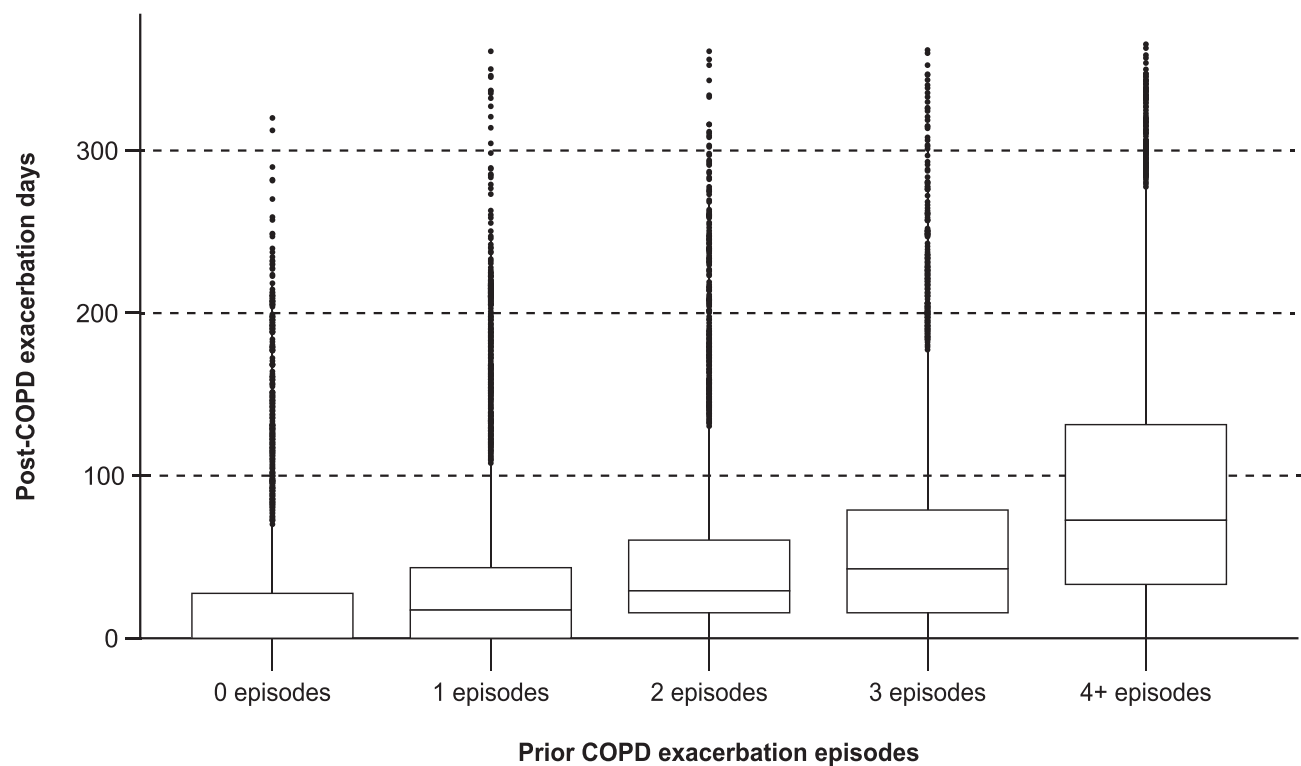
primary and secondary healthcare centers. In this study, the model identified the number of COPD-related healthcare contacts in the patients' entire medical history as the third most important predictor of imminent hospitalization due to exacerbation (following severe exacerbation in the past 180 days and severe exacerbation at any point in the patients' medical history).<sup>7</sup>

**Table 1** Variables with a Positive Effect on Exacerbation Days

Variable Name	Variable Explanation	Effect Size, Days
ER visit	1 ER visit in the past year (compared with 0 visits)	1.66
ER visit	2+ ER visits in the past year (compared with 0 visits)	2.88
Inpatient visit	1 inpatient visit in the past year (compared with 0 visits)	0.28
Inpatient visit	2+ inpatient visits in the past year (compared with 0 visits)	27.05
Outpatient visit	Outpatient visit counts – quartile 2 (13–21) compared with quartile 1: ≤12 in the past year	1.63
Outpatient visit	Outpatient visit counts – quartile 3 (22–35) compared with quartile 1: ≤12 in the past year	3.05
Outpatient visit	Outpatient visit counts – quartile 4 (>35) compared with quartile 1: ≤12 in the past year	4.28
Prior COPD exacerbation episode <sup>a</sup>	1 COPD exacerbation episode <sup>a</sup> in the past year (compared with 0 episodes)	5.51
Prior COPD exacerbation episode <sup>a</sup>	2 COPD exacerbation episodes <sup>a</sup> in the past year (compared with 0 episodes)	11.69
Prior COPD exacerbation episode <sup>a</sup>	3 COPD exacerbation episodes <sup>a</sup> in the past year (compared with 0 episodes)	18.80
Prior COPD exacerbation episode <sup>a</sup>	4+ COPD exacerbation episodes <sup>a</sup> in the past year (compared with 0 episodes)	32.42
Age	Age (in years) at index date	0.24
Antibiotics (1 month)	Antibiotics medication in the past month	5.32
Pneumonia	Pneumonia (except that caused by tuberculosis or sexually transmitted disease) diagnosis in the past year	1.94
Respiratory failure	Respiratory failure; insufficiency; arrest (adult) diagnosis in the past year	3.17
Esophageal disorders	Esophageal disorders diagnosis in the past year	2.18
Rheumatoid arthritis	Rheumatoid arthritis and related disease diagnosis in the past year	6.39
Bacterial infection	Bacterial infection; unspecified site diagnosis in the past year	3.67
Anxiety disorders	Anxiety disorders in the past year	3.58
Chest x-ray (1 month)	Chest x-ray procedure in the past month	1.49
LABA	LABA medication in the past year	2.90
Anti-infective misc.	Anti-infective miscellaneous – combinations use in the past year	0.88
Glucocorticosteroids	Glucocorticosteroid use in the past year	1.00
Bisphosphonates	Bisphosphonates (bone-density regulators) use in the past year	3.13
Xanthines	Xanthines use in the past year	11.54
Leukotrine receptor antagonists	Leukotrine receptor antagonists use in the past year	1.35
Anticonvulsants misc.	Anticonvulsants miscellaneous – use in the last year	2.73
Potassium	Potassium	2.53
Anti-infectives – throat	Anti-infectives – throat use in the past year	5.22
Oxygen therapy	Oxygen therapy in the past year	5.48
Oxygen therapy (1 month)	Oxygen therapy in the past month	8.55
SABA	SABA medication in the past year	2.26
SABA (1 month)	SABA medication in the past month	2.51
SABA/SAMA	SABA/SAMA medication in the past year	3.15
SABA/SAMA (1 month)	SABA/SAMA medication in the past month	4.70

**Notes:** <sup>a</sup>COPD exacerbation episode defined by COPD-related health encounters, particularly a COPD-related ER, urgent care, or outpatient visit with a diagnosis of COPD in any position and at least one dispensing for antibiotic or oral systemic corticosteroids within 5 days of the visit, or a COPD-related inpatient hospitalization (an inpatient hospital stay with a diagnosis of COPD in any position).

**Abbreviations:** COPD, chronic obstructive pulmonary disease; ER, emergency room; LABA, long-acting  $\beta_2$  agonist; SABA, short-acting  $\beta_2$  agonist; SAMA, short-acting muscarinic agonist.



**Figure 2** Correlation between prior COPD exacerbation episodes and post-index exacerbation time.

**Abbreviation:** COPD, chronic obstructive pulmonary disease.

When considering medication use, the prescription of xanthines in the past year (effect size 11.5 days), oxygen therapy in the past month (effect size 8.6 days), and antibiotics in the past month (effect size 5.3 days) were all associated with the greatest increase in exacerbation time. As xanthines are commonly used to treat asthma, this variable serves as a predictor of patients who had asthma in the baseline period. This finding provides evidence to support the clinical expectation that patients with concomitant asthma will suffer more exacerbation days during their treatment. Supplemental oxygen is administered to treat hypoxemia in patients hospitalized for an exacerbation of COPD.<sup>18</sup> We found that oxygen therapy in the past month increased exacerbation length, suggesting that this variable may represent patients who recently experienced a severe exacerbation. Antibiotic use was also identified as an important predictor of exacerbation risk in two studies that utilized machine learning algorithms.<sup>7,11</sup> Tavakoli et al applied several statistical and machine learning algorithms to Canadian administrative health data to predict the risk of severe exacerbation during the following 2 months. Gradient boosting was identified as the best predictive algorithm; this model found that total antibiotic prescriptions and number of dispensed all-cause medications were important predictors of severe exacerbation.<sup>11</sup>

**Table 2** Variables with a Negative Effect on Exacerbation Days

Variable Name	Variable Explanation	Effect Size, Days
Airflow test	Airflow test in the past year	-2.03
From the West region	West region (compared with non-West) at index date	-1.65
Anemia	Acute post-hemorrhagic anemia	-5.68
HMG-CoA reductase inhibitors	HMG-CoA reductase inhibitors use in the past year	-1.61
Oral corticosteroid (1 month)	Oral corticosteroid use in the past month	-3.43

**Abbreviation:** HMG-CoA,  $\beta$ -hydroxy  $\beta$ -methylglutaryl-CoA.

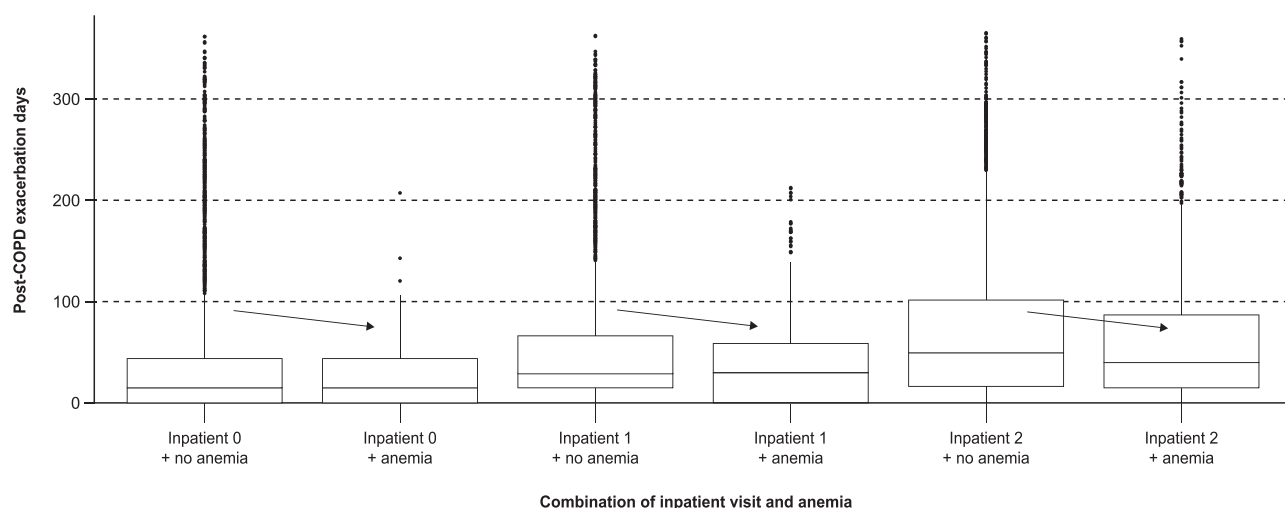
**Table 3** Interaction Terms and Effect Size on Exacerbation Days

Interaction Term	Effect Size, Days
1 inpatient visit * Age	−0.05
1 inpatient visit * Chest x-ray (1 month)	6.71
2 inpatient visits * Age	−0.46
2 inpatient visits * Chest x-ray (1 month)	8.57
1 COPD exacerbation episode <sup>a</sup> * Oral corticosteroid (1 month)	15.53
2 COPD exacerbation episodes <sup>a</sup> * Oral corticosteroid (1 month)	13.13
3 COPD exacerbation episodes <sup>a</sup> * Oral corticosteroid (1 month)	17.73
4+ COPD exacerbation episodes <sup>a</sup> * Oral corticosteroid (1 month)	25.68
Antibiotics (1 month) * Oral corticosteroid (1 month)	−9.82
Respiratory failure * Oral corticosteroid (1 month)	9.88
Chest x-ray (1 month) * Oral corticosteroid (1 month)	−7.61
Pneumonia * Anti-infective misc.	7.11
Outpatient visit counts – quartile 2 (13–21) * Oxygen therapy (1 month)	0.40
Outpatient visit counts – quartile 3 (22–35) * Oxygen therapy (1 month)	1.74
Outpatient visit counts – quartile 4 (>35) * Oxygen therapy (1 month)	8.66
1 COPD exacerbation episode <sup>a</sup> * Oxygen therapy	−3.12
2 COPD exacerbation episodes <sup>a</sup> * Oxygen therapy	−0.51
3 COPD exacerbation episodes <sup>a</sup> * Oxygen therapy	−1.32
4+ COPD exacerbation episodes <sup>a</sup> * Oxygen therapy	15.05
Oral corticosteroid (1 month) * Oxygen therapy (1 month)	11.75
Outpatient visit counts – quartile 2 (13–21) * Glucocorticosteroids	0.42
Outpatient visit counts – quartile 3 (22–35) * Glucocorticosteroids	2.81
Outpatient visit counts – quartile 4 (>35) * Glucocorticosteroids	7.98

**Notes:** <sup>a</sup>COPD exacerbation episode defined by COPD-related health encounters, particularly a COPD-related ER, urgent care, or outpatient visit with a diagnosis of COPD in any position and at least one dispensing for antibiotic or oral systemic corticosteroids within 5 days of the visit, or a COPD-related inpatient hospitalization (an inpatient hospital stay with a diagnosis of COPD in any position).

**Abbreviations:** COPD, chronic obstructive pulmonary disease; ER, emergency room.

Several previous cohort studies have confirmed that comorbidity burden is associated with an increased risk of COPD exacerbation.<sup>8–10</sup> In addition, the machine learning algorithm used by Stållberg et al demonstrated that Charlson Comorbidity Index score, a measure of the number and severity of comorbidities, was a good predictor of hospitalization due to exacerbation.<sup>7</sup>

**Figure 3** Correlation between the combination of inpatient visit and acute post-hemorrhagic anemia and post-index exacerbation time.

**Abbreviation:** COPD, chronic obstructive pulmonary disease.



**Table 4** Hazard Ratios for Variables Identified in Exploratory Analysis Indicating Risk of Triple Therapy Initiation

Variable	HR
<b>Variables with a positive hazard ratio for initiation of triple therapy</b>	
Index drug ICS/LABA (categorical, compared with ICS)	2.42
Index drug LABA (categorical, compared with ICS)	1.40
Index drug LAMA (categorical, compared with ICS)	3.82
Index drug LAMA/LABA (categorical, compared with ICS)	2.36
CT scan chest	1.13
COPD-related inpatient count	1.05
Prior COPD exacerbation <sup>a</sup> in the past year	1.55
Age at index date	1.14
Acute bronchitis (binary, yes or no)	1.08
Quadratic term of age	1.00
Smoking history (binary, yes or no)	1.53
<b>Variables with a negative HR for initiation of triple therapy</b>	
Inpatient visit in the last year	0.75
Antibiotics (binary, yes or no)	0.85
Medication – injections, infusions, and other forms (binary)	0.99
Index year	0.98
Asthma (binary)	0.77

**Notes:** <sup>a</sup>COPD exacerbation episode defined by COPD-related health encounters, particularly a COPD-related ER, urgent care, or outpatient visit with a diagnosis of COPD in any position and at least one dispensing for antibiotic or oral systemic corticosteroids within 5 days of the visit, or a COPD-related inpatient hospitalization (an inpatient hospital stay with a diagnosis of COPD in any position).

**Abbreviations:** COPD, chronic obstructive pulmonary disease; CT, computed tomography; ER, emergency room; HR, hazard ratio; ICS, inhaled corticosteroid; LABA, long-acting  $\beta_2$  agonist; LAMA, long-acting muscarinic antagonist.

Of the comorbidities considered by our model, RA had the largest positive effect on exacerbation time (6.4 days). Patients with RA have an increased risk of developing COPD compared with controls,<sup>19</sup> and increased mortality has been reported among patients with RA and COPD versus with RA patients without COPD.<sup>20</sup> A previous population-based cohort study reported that the risk of hospitalization due to COPD exacerbation was 47% higher among patients with RA compared with the general population, after adjusting for potential confounders.<sup>21</sup> It is possible that the increased risk of exacerbation among patients with RA and COPD may be linked to the role of inflammation in both diseases.<sup>21</sup>

Concurrent esophageal disorders were also found to have a positive effect on future exacerbation time, adding to previous reports that abnormal swallowing reflexes (including gastroesophageal reflux disease) are prevalent among patients with COPD and are predictive of exacerbation.<sup>10,22</sup> In our model, respiratory failure, bacterial infection, and anxiety disorders were also found to have a positive effect on future exacerbation time.

Conversely, lung function testing and oral corticosteroid use in the past month predicted a reduced exacerbation time. In addition to airway inflammation, systemic inflammation also plays a role in the pathogenesis of COPD, with elevated inflammatory mediators observed in stable COPD and during exacerbations.<sup>23</sup> GOLD treatment guidelines recommend the use of short-term systemic corticosteroids during a COPD exacerbation to improve recovery time, lung function, and length of hospitalization.<sup>1</sup> It is therefore likely that patients who received oral corticosteroids in the past month did so for a recent exacerbation; this group of patients may be closely monitored by their healthcare provider and treatment adjusted in a timely manner to prevent further exacerbation.

Of the patients included in the model training dataset (n=58,901), 3.1% had a diagnosis of acute post-hemorrhagic anemia in the baseline period. Prior literature has reported that COPD can lead to anemia and worsen COPD exacerbations.<sup>24</sup> Consistent with previous reported data, the authors performed a univariate test that confirmed a diagnosis of anemia was associated with an increase of exacerbation days (data not shown). However, the REFS<sup>TM</sup> model reported that a diagnosis of acute post-h hemorrhagic anemia in the baseline period was associated with

**Table 5** Hazard Ratios for Interaction Terms Identified in the Exploratory Analysis Indicating Risk of Triple Therapy Initiation

Variable	HR
Antibiotics * Prior COPD exacerbation <sup>a</sup>	1.26
ICS/LABA * Inpatient visit	1.15
LABA * Inpatient visit	1.09
LAMA/LABA * Inpatient visit	1.04
LAMA * Inpatient visit	1.18
Inpatient visit * Prior COPD exacerbation <sup>a</sup>	0.96
ICS/LABA * Prior COPD exacerbation <sup>a</sup>	0.97
LABA * Prior COPD exacerbation <sup>a</sup>	1.01
LAMA/LABA * Prior COPD exacerbation <sup>a</sup>	1.01
LAMA * Prior COPD exacerbation <sup>a</sup>	0.97
Antibiotics * Acute bronchitis	0.82
ICS/LABA * Asthma	0.99
LABA * Asthma	1.08
LAMA/LABA * Asthma	1.82
LAMA * Asthma	1.55

**Notes:** <sup>a</sup>COPD exacerbation episode defined by COPD-related health encounters, particularly a COPD-related ER, urgent care, or outpatient visit with a diagnosis of COPD in any position and at least one dispensing for antibiotic or oral systemic corticosteroids within 5 days of the visit, or a COPD-related inpatient hospitalization (an inpatient hospital stay with a diagnosis of COPD in any position).

**Abbreviations:** COPD, chronic obstructive pulmonary disease; ER, emergency room; HR, hazard ratio; ICS, inhaled corticosteroid; LABA, long-acting  $\beta_2$  agonist; LAMA, long-acting muscarinic antagonist.

a reduction in exacerbation time (5.7 days) following triple therapy initiation. This reversed finding from our analysis is possibly due to joint effects; there was a decreasing trend of COPD exacerbation days for patients with anemia given their inpatient visit level, hence the model identified anemia as having a negative effect size. As the study focused on the prediction of exacerbations, any joint estimations that improve the prediction (in this case anemia and inpatient visit) were adopted by the REFS<sup>TM</sup> model.

In the exploratory analysis, the model was also able to identify potential triggers for the initiation of triple therapy among COPD patients on mono- or dual therapy. It was found that patients who initiate LAMA, ICS/LABA, LAMA/LABA, or LABA were more likely to escalate to triple therapy compared with patients who initiate ICS only. However, initiating ICS monotherapy may indicate that a patient has mild disease or a concurrent asthma diagnosis, and may therefore be less likely to initiate triple therapy. Diagnosis of asthma led to a 23% reduction in the likelihood to switch to triple therapy, although a diagnosis of asthma combined with initiation of LAMA, or LAMA/LABA, resulted in an increased chance of triple therapy initiation. Patients with both asthma and COPD must balance their medication carefully in order to control both diseases, which may explain the lower chance of a switch to triple therapy with an asthma diagnosis.

The REFS<sup>TM</sup> model used in the primary analysis may have real-world applications. A web application is currently in development which will allow physicians to input variables across five categories (demographics, visits, diagnosis, medication, and medical procedures) to simulate a patient profile and observe the predicted exacerbation profile. It is hoped that, in the future, this application could be used as a reference tool in clinical practice.

A major strength of this study is use of the REFS<sup>TM</sup> platform, which offers advantages over traditional transgressional modeling. The platform is suited for the exploration of large hypothesis spaces; predictor discovery occurs in an unbiased manner that can lead to novel predictive insights, and thousands of predictive models are created (allowing the strength of an individual predictor to be assessed as a function of the proportion of the models in which it appears). The REFS<sup>TM</sup> model has been successfully used in various other disorders.<sup>25–28</sup>

Limitations of this study are mostly typical for studies using administrative claims data. Coding inaccuracies – a lack of diagnosis for comorbid conditions or missing or incorrect ICD-9/10 codes – could lead to an underestimation or absence of some conditions. In addition, patient compliance for medications cannot be guaranteed, as the filling of a prescription does not guarantee that the patient took the medicine. Also, certain data are missing from the database, for example forced expiratory volume, patient-reported outcomes, and laboratory biomarkers. Furthermore, confounders for every predictor may not be accounted for if they are not selected into the model, although strong confounders associated with the outcomes of interest are likely selected. Finally, formal validation of the model findings was beyond the scope of the current study, however validation using another cohort or time period will be considered for future work.

## Conclusions

Using a reverse engineering forward simulation machine learning platform, prior exacerbation was shown as a strong factor in predicting the future increase of exacerbations in patients with COPD initiating triple therapy, as were healthcare resource utilization, xanthine use, and concomitant RA. Lung function testing and oral corticosteroid use predicted a reduction in the number of future exacerbation days. The results from this study improve our understanding of exacerbation risk among treated COPD patients as well as triggers for triple therapy initiation.

## Abbreviations

COPD, chronic obstructive pulmonary disease; CT, computed tomography; ER, emergency room; HMG-CoA,  $\beta$ -hydroxy  $\beta$ -methylglutaryl-CoA; HR, hazard ratio; ICS, inhaled corticosteroid; LABA, long-acting  $\beta_2$  agonist; LAMA, long-acting muscarinic antagonist; Optum CDM, Optum™ Clinformatics™ Data Mart; RA, rheumatoid arthritis; REFS™, reverse engineering forward simulation; SABA, short-acting  $\beta_2$  agonist; SAMA, short-acting muscarinic agonist; SD, standard deviation.

## Data Sharing Statement

The data analyzed in this manuscript are contained in a database owned by Optum Clinformatics and therefore are not publicly available. Access to the data may be available on license from Optum ([optum.com/life-sciences-solutions](https://optum.com/life-sciences-solutions)).

## Ethics Approval and Informed Consent

Ethics committee approval was not required for this study as the results were presented as aggregate analyses that omit patient identification. Patient informed consent was not required as there was no direct patient contact or primary collection of individual patient data.

## Acknowledgments

Editorial support (in the form of writing assistance, including preparation of the draft manuscript under the direction and guidance of the authors, collating and incorporating authors' comments for each draft, assembling tables and figures, grammatical editing, and referencing) was provided by Fiona Goodwin and Rebecca Cunningham of Aura, a division of Spirit Medical Communications Group Limited, and was funded by GlaxoSmithKline. These data have been presented in abstract/poster form at the American Thoracic Society – 117th International Conference (Bogart M, Oakland T, Liu Y, Enev T. Triple therapy treatment pathways in chronic obstructive pulmonary disease (COPD): a real-world predictive model – American Thoracic Society – 117th International Conference. *Am J Respir Crit Care Med*. 2021;203:A2313.

## Author Contributions

All authors made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; took part in drafting the article or revising it critically for important intellectual content; agreed to submit to the current journal; gave final approval of the version to be published; and agree to be accountable for all aspects of the work.

## Funding

This study was funded by GlaxoSmithKline plc (study number 213319). The sponsor was involved in study conception and design, data interpretation, and the decision to submit the article for publication. The sponsor was also given the opportunity to review the manuscript for medical and scientific accuracy as well as intellectual property considerations.

## Disclosure

MB is an employee of and/or holds stocks/shares in GlaxoSmithKline. YL is an employee of and/or holds stocks/shares in GNS Healthcare. TO is an employee of and/or holds stocks/shares in GNS Healthcare. MS is an employee of and/or holds stocks/shares in GlaxoSmithKline. The authors report no other conflicts of interest in this work.

## References

1. Global Initiative for Chronic Obstructive Lung Disease. 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 August 25, 2021.
2. Blasi F, Cesana G, Conti S, et al. The clinical and economic impact of exacerbations of chronic obstructive pulmonary disease: a cohort of hospitalized patients. *PLoS One*. 2014;9(6):e101228. doi:10.1371/journal.pone.0101228
3. Dalal AA, Shah M, D'Souza AO, Rane P. Costs of COPD exacerbations in the emergency department and inpatient setting. *Respir Med*. 2011;105(3):454–460. doi:10.1016/j.rmed.2010.09.003
4. Guerra B, Gaveikaite V, Bianchi C, Puhon MA. Prediction models for exacerbations in patients with COPD. *Eur Respir Rev*. 2017;26(143):160061. doi:10.1183/16000617.0061-2016
5. Hoogendoorn M, Feenstra TL, Boland M, et al. Prediction models for exacerbations in different COPD patient populations: comparing results of five large data sources. *Int J Chron Obstruct Pulmon Dis*. 2017;12:3183–3194. doi:10.2147/COPD.S142378
6. 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
7. Stållberg B, Lisspers K, Larsson K, et al. Predicting hospitalization due to COPD exacerbations in Swedish primary care patients using machine learning – based on the ARCTIC study. *Int J Chron Obstruct Pulmon Dis*. 2021;16:677–688. doi:10.2147/COPD.S293099
8. Stållberg B, Janson C, Larsson K, et al. Real-world retrospective cohort study ARCTIC shows burden of comorbidities in Swedish COPD versus non-COPD patients. *NPJ Prim Care Respir Med*. 2018;28(1):33. doi:10.1038/s41533-018-0101-y
9. Kaszuba E, Odeberg H, Rastam L, Halling A. Impact of heart failure and other comorbidities on mortality in patients with chronic obstructive pulmonary disease: a register-based, prospective cohort study. *BMC Fam Pract*. 2018;19(1):178. doi:10.1186/s12875-018-0865-8
10. Westerik JA, Metting EI, van Boven JF, Tiersma W, Kocks JW, Schermer TR. Associations between chronic comorbidity and exacerbation risk in primary care patients with COPD. *Respir Res*. 2017;18(1):31. doi:10.1186/s12931-017-0512-2
11. Tavakoli H, Chen W, Sin DD, FitzGerald JM, Sadatsafavi M. Predicting severe chronic obstructive pulmonary disease exacerbations. Developing a population surveillance approach with administrative data. *Ann Am Thorac Soc*. 2020;17(9):1069–1076. doi:10.1513/AnnalsATS.202001-070OC
12. Zheng Y, Zhu J, Liu Y, et al. Triple therapy in the management of chronic obstructive pulmonary disease: systematic review and meta-analysis. *BMJ*. 2018;363:k4388. doi:10.1136/bmj.k4388
13. Bogart M, Glassberg MB, Reinsch T, Stanford RH. Impact of prompt versus delayed initiation of triple therapy post COPD exacerbation in a US-managed care setting. *Respir Med*. 2018;145:138–144. doi:10.1016/j.rmed.2018.10.013
14. Yu AP, Guérin A, Ponce de Leon D, et al. Therapy persistence and adherence in patients with chronic obstructive pulmonary disease: multiple versus single long-acting maintenance inhalers. *J Med Econ*. 2011;14(4):486–496. doi:10.3111/13696998.2011.594123
15. Bogart M, Wu B, Germain G, et al. Real-world adherence to single-inhaler vs multiple-inhaler triple therapy among patients with COPD in a commercially insured US population. *Chest*. 2020;158(4):Supplement A1773–A1774. doi:10.1016/j.chest.2020.08.1548
16. Xing H, McDonagh PD, Bienkowska J, et al. Causal modeling using network ensemble simulations of genetic and gene expression data predicts genes involved in rheumatoid arthritis. *PLoS Comput Biol*. 2011;7(3):e1001105. doi:10.1371/journal.pcbi.1001105
17. Iheanacho I, Zhang S, King D, Rizzo M, Ismaila AS. Economic burden of chronic obstructive pulmonary disease (COPD): a systematic literature review. *Int J Chron Obstruct Pulmon Dis*. 2020;15:439–460. doi:10.2147/COPD.S234942
18. Brill SE, Wedzicha JA. Oxygen therapy in acute exacerbations of chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis*. 2014;9:1241–1252. doi:10.2147/COPD.S41476
19. Ma Y, Tong H, Zhang X, et al. Chronic obstructive pulmonary disease in rheumatoid arthritis: a systematic review and meta-analysis. *Respir Res*. 2019;20(1):144. doi:10.1186/s12931-019-1123-x
20. Hyldgaard C, Bendstrup E, Pedersen AB, et al. Increased mortality among patients with rheumatoid arthritis and COPD: a population-based study. *Respir Med*. 2018;140:101–107. doi:10.1016/j.rmed.2018.06.010
21. McGuire K, Aviña-Zubieta JA, Esdaile JM, et al. Risk of incident chronic obstructive pulmonary disease in rheumatoid arthritis: a population-based cohort study. *Arthritis Care Res (Hoboken)*. 2019;71(5):602–610. doi:10.1002/acr.23410
22. Terada K, Muro S, Ohara T, et al. Abnormal swallowing reflex and COPD exacerbations. *Chest*. 2010;137(2):326–332. doi:10.1378/chest.09-0482
23. Falk JA, Minai OA, Mosenifar Z. Inhaled and systemic corticosteroids in chronic obstructive pulmonary disease. *Proc Am Thorac Soc*. 2008;5(4):506–512. doi:10.1513/pats.200707-096ET
24. McMahon TJ, Prybylowski AC. Anemia in the patient with chronic lung disease. In: Provenzano R, Lerma E, Szczec L, (editors). *Management of Anemia*. New York: Springer; 2018:143–155. doi:10.1007/978-1-4939-7360-6\_8.

25. Latourelle JC, Beste MT, Hadzi TC, et al. Large-scale identification of clinical and genetic predictors of motor progression in patients with newly diagnosed Parkinson's disease: a longitudinal cohort study and validation. *Lancet Neurol.* **2017**;16(11):908–916. doi:10.1016/S1474-4422(17)30328-9
26. Berger JS, Haskell L, Ting W, et al. Machine learning methodology predicts comorbidities are associated with increased total healthcare costs among patients with severe peripheral artery disease. Poster presented at: The American Heart Association Quality of Care and Outcomes Research Scientific Sessions; April 2–3, 2017, Arlington, Virginia. *Circulation: Cardiovascular Quality and Outcomes.* **2017**;10(Suppl\_3):A150. doi:10.1161/circoutcomes.10.suppl\_3.150
27. Ivanov V, Torgovitsky R, Tcheggen ET, et al. Using clinical trial and real world data to bridge efficacy to effectiveness of fingolimod in multiple sclerosis patients. *Value Health.* **2016**;19(7):PA426. doi:10.1016/j.jval.2016.09.461
28. Anderson JP, Parikh JR, Shenfeld DK, et al. Reverse engineering and evaluation of prediction models for progression to type 2 diabetes: an application of machine learning using electronic health records. *J Diabetes Sci Technol.* **2016**;10(1):6–18. doi:10.1177/1932296815620200

## International Journal of Chronic Obstructive Pulmonary Disease

Dovepress

### Publish your work in this journal

The International Journal of COPD is an international, peer-reviewed journal of therapeutics and pharmacology focusing on concise rapid reporting of clinical studies and reviews in COPD. Special focus is given to the pathophysiological processes underlying the disease, intervention programs, patient focused education, and self management protocols. This journal is indexed on PubMed Central, MedLine and CAS. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-chronic-obstructive-pulmonary-disease-journal>