CONFIDENTIAL

HEALTH OUTCOMES STUDY PROTOCOL

UNIQUE IDENTIFIER	HO-18-18992 ETrack: 208782
FULL TITLE	Claims-linked Survey Study to Assess Burden of Illness among Patients Treated with LAMA/LABA vs ICS/LABA Single Inhaler Dual Therapy
ABBREVIATED TITLE	COPD Dual Therapy Burden of Illness
FINAL PROTOCOL APPROVED	17-4-2018
SPONSORSHIP	Sponsored
DIVISION	Pharma
BUSINESS UNIT	US Medical Affairs
DEPARTMENT	USHO
STUDY ACCOUNTABLE PERSON	PPD ScD, Director, CEVEO
CONTRIBUTING AUTHORS	PPD PhD, Director, CEVEO PPD PharmD, Senior Director, CEVEO PPD , MA, Associate Director, Optum

ASSET ID	17379
GSK ASSET	Fluticasone/Salmeterol 250/50 mcg (Advair) Umeclidinium/Vilanterol 62.5/25 mcg (Anoro)
INDICATION	COPD

SPONSOR SIGNATORY

PPD

Mart PPD Marciniak
VP Customer Engagement Values Evidence and Outcomes,
Medical Affairs NA

4-18-2018
Date

PROTOCOL SYNOPSIS

Unique Identifier	HO-18-18992
Abbreviated Title	COPD Dual Therapy Burden of Illness
GSK Product	Fluticasone/Salmeterol 250/50 mcg (FLUT/SAL; Advair) Umeclidinium/Vilanterol 62.5/25 mcg (UMEC/VI; Anoro)
Rationale	The purpose of the study is to assess COPD burden of illness using both patient-reported symptom burden and claims-based economic burden among patients treated with FLUT/SAL or UMEC/VI to support GOLD category B recommendations.
Objectives (Primary, Secondary)	<u>Primary Objective</u> • The primary objective is to compare patient-reported COPD symptom burden among patients treated with umeclidinium/vilanterol (UMEC/VI) or fluticasone/salmeterol (FLUT/SAL) <u>Secondary Objectives</u> • Describe baseline all-cause, COPD-related, and exacerbation-related health care resource utilization, costs, and clinical characteristics (e.g., comorbidities, treatment history) in patients treated with UMEC/VI or FLUT/SAL • Identify the GOLD classification of patients treated with FLUT/SAL or UMEC/VI.
Study Design	This is a claims-linked cross-sectional observational survey study of Medicare Advantage (MA) enrollees of a large US health plan with evidence of COPD and single-inhaler dual therapy treatments FLUT/SAL or UMEC/VI. Patients will be recruited using Optum's health plan recruitment strategy. For this approach, patients will be identified in the Optum Research Database (ORD) following receipt of the appropriate approvals. Patients will complete a mailed cross-sectional survey to collect information on patients' condition history, current treatment, smoking history, symptoms and symptom severity (mMRC and CAT), and demographic and sociodemographic characteristics. Patient survey and diary responses are linked to medical and pharmacy claims for the 12-month period prior to the survey (baseline). A 30-35% response rate is estimated based on previous patient survey studies. Two waves of data collection are anticipated to meet the target sample size based on patient counts and estimated response rate. Descriptive and multivariable techniques are used for analysis.
Study Population and Sampling Methods	<u>Inclusion Criteria:</u> • ≥2 ICD-10-CM diagnosis codes for COPD at least 30 days apart during the 12 month period prior to sample identification ○ Diagnosis codes J40-J44 will be included • ≥1 pharmacy claim for UMEC/VI or FLUT/SAL single-inhaler dual therapy during baseline • Age ≥ 65 • Self-reported health care provider diagnosis of COPD • Self-reported prescription for FLUT/SAL or UMEC/VI • 12 months of continuous enrollment during the baseline period • Ability to complete the study survey in English <u>Exclusion criteria:</u>

	12 month period prior to sample identification • Claims for both UMEC/VI and FLUT/SAL in the 6 months closest to sample identification • Claims for triple therapy (ICS + LAMA + LABA) during the baseline period • Evidence of lung cancer diagnosis and/or treatment
Data Source	The ORD includes enrollment, pharmacy, and medical claims data from a large U.S. health plan affiliated with Optum to identify eligible adult participants. The individuals covered by this health plan are geographically diverse across the US.
Data Analysis Methods	Study endpoints and all other study variables will be analyzed descriptively. Numbers and percents will be provided for dichotomous and polychotomous variables. Means, medians, and standard deviations will be provided for continuous variables. Means and standard deviations will be provided for cost and utilization measures. <u>Primary Endpoints</u> *Patient-reported symptom burden.* Scores will be presented from validated instruments measuring dyspnea (mMRC) and condition-related burden of illness (CAT). Mean total, summary, and/or domain scores and standard deviations will be presented for each scale/item). Results will be presented for the overall study sample and will compare self-reported FLUT/SAL to UMEC/VI treatment cohorts. To address the objective of comparing symptom burden of patients by treatment cohort (FLUT/SAL as compared to UMEC/VI), inverse probability of treatment weighting (IPTW) will be applied to reduce confounding on observed clinical, demographic, and sociodemographic characteristics. *Proportion of patients reporting COPD symptoms while treated with FLUT/SAL or UMEC/VI.* The proportion of patients who report COPD symptoms will be calculated as the number of COPD patients who report COPD symptoms as measured by the CAT, divided by the total number of respondents. Results will be presented for the overall study sample and by treatment cohort (FLUT/SAL or UMEC/VI). <u>Secondary Endpoints</u> *Baseline health care resource utilization and costs.* All-cause and COPD-related health care costs will be presented for the 12-month baseline period. Mean, median, and standard deviations will be presented for total costs and utilization and by each category of costs. Costs and counts will be presented overall and for the subset of patients utilizing the selected type of service. The proportion of patients receiving services for each type of service and the proportion of patients with non-zero costs for each type of cost will be calculated. Results will be presented for the overall study sample and by treatment cohort (FLUT/SAL or UMEC/VI). *Patient clinical characteristics.* The proportion of patients with evidence of selected clinical characteristics will be presented. The top 20 claims-based comorbidities and 10 COPD-related comorbidities of interest will be identified. The percent of patients with each condition will be identified. Mean and standard deviation will be calculated for the Quan-Charlson comorbidity score. Results will be presented for the overall study sample and may be stratified by relevant subgroups for analysis (e.g., by symptom severity as measured by the mMRC and CAT, and by treatment). *Patient GOLD classification.* The proportion of patients in each GOLD category of symptom burden and exacerbation (A, B, C, D) will be identified using the baseline

	count of COPD exacerbations combined with CAT total score and mMRC score. GOLD categories are mutually exclusive. Results will be presented overall and by treatment cohort (FLUT/SAL or UMEC/VI). Multivariable modeling of the primary objective will be conducted following review of descriptive and IPTW results. Each model will examine the relationship between symptom burden and treatment with FLUT/SAL or UMEC/VI. The dependent variable for the models will be CAT total score and mMRC score, respectively; covariates for the models will be based on clinical and analytic significance.
Sample Size and Power	The study sample is estimated to include 385 patients per cohort with evidence of COPD diagnosis and single-inhaler dual therapy treatment with FLUT/SAL or UMEC/VI. A sample size of n=385 per cohort is required to assess statistical significance with a 95% confidence interval at ±0.05 with 80% power. A sampling frame of 2,700 patients is estimated to be used to reach the target evaluable sample size, n=770 patients (385/cohort). Based on patient counts and response rate estimates, it is anticipated that two waves of data collection will be needed to meet the target sample size. This estimated study sample size is sufficient to detect proportional differences in PRO outcomes between two cohorts of patients (i.e., patients treated with FLUT/SAL or UMEC/VI who have severe symptoms as measured by a CAT total score ≥10 or an mMRC score ≥2); other stratifications tied to primary and secondary endpoints may be considered for statistically powered analyses as sample size allows.
Limitations	Because claims data are collected for the purpose of payment and not research, there are certain limitations associated with the use of claims data. First, presence of a claim for a filled prescription does not indicate that the medication was consumed or taken as prescribed. Second, medications filled over-the-counter or provided as samples by a physician are not observed in the claims data. Third, presence of a diagnosis code on a medical claim is not positive presence of disease, as the diagnosis code may be included as rule-out criteria or may be incorrectly coded. Use of multiple claims with COPD diagnosis and of patient self-reported treatment mitigates these limitations. Limitations of survey data may include sampling error, coverage error, and measurement error. Because the study population is selected from MA enrollees, the results may not be generalizable to uninsured populations, those covered by Medicare fee for service plans, and younger patients with COPD.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

ABBREVIATIONS

AE	Adverse Event
AHRQ	Agency for Healthcare Research and Quality
BOI	Burden of Illness
CAT	COPD Assessment Test
CE	Continuous enrollment
CMS	Centers for Medicare and Medicaid Services
COPD	Chronic Obstructive Pulmonary Disease
CPT	Current Procedural Terminology
FDA	Food and Drug Administration
FDC	Fixed dose combination
FLUT	Fluticasone
GLM	Generalized Linear Model
GOLD	Global Initiative for Chronic Lung Disease
HCPCS	Healthcare Common Procedure Coding
HCRU	Health Care Resource Utilization
HIPAA	Health Insurance Portability and Accountability Act
ICD	International Statistical Classification of Diseases and Related Health Problems
ICS	Inhaled Corticosteroid
IPTW	Inverse probability of treatment weighting
IRB	Institutional Review Board
LABA	Long-acting ß-agonist
LAMA	Long-acting Antimuscarinic
MA	Medicare Advantage (Medicare)
mMRC	Modified Medical Research Council Dyspnea Scale
NMD	National Medical Director
ORD	Optum Research Database
PHI	Protected Health Information
PII	Personally Identifiable Information
PRO	Patient-reported Outcome
PS	Propensity score
RRB	Research Review Board
SAE	Serious Adverse Event
SAL	Salmeterol
SAP	Statistical Analysis Plan
T2DM	Type 2 diabetes mellitus
UMEC	Umeclidinium
VI	Vilanterol

1 INTRODUCTION/BACKGROUND

Chronic obstructive pulmonary disease (COPD) is a prevalent respiratory condition characterized by a gradual loss of lung function and progressive airflow limitation that is not fully reversible.[1] COPD includes chronic bronchitis, chronic obstructive bronchitis, emphysema, and combinations of these conditions and is characterized by airflow obstruction that progressively worsens with time, leading to dyspnea, coughing, wheezing, sputum production, and/or exercise intolerance.[2]

COPD is estimated to affect 24 million people in the U.S.[3] and is the third-leading cause of death in the U.S.[4] Among the top six causes of death, COPD is the only one that continues to increase in incidence.[5] The prevalence of COPD is rising with the increasing proportion of the population over the age of 65, and the burden of COPD is projected to rise considerably over the next 15-20 years, driven largely by increasing prevalence and population demographics. Recent reports indicate that COPD accounts for 1.5 million emergency department (ED) visits and 636,000 inpatient hospitalizations (IP) annually in the United States.[6,7] The quality of life and economic impact of COPD are substantial; treatments that can improve management of COPD and reduce use of healthcare services can both improve patient quality of life and reduce health care expenditures.

Studies of COPD patients have demonstrated that those with higher adherence to their prescribed treatment regimens experienced fewer hospitalizations and lower medical costs than patients who exhibited lower adherence behaviors.[8] Ease and frequency of use are important considerations in disease management. Products that are easier for patients to use consistently may improve treatment adherence, thereby resulting in improved symptom management. By helping to better manage symptoms, improved adherence may also improve health-related quality life, exercise tolerance, and other pro-health behaviors, thereby helping to mitigate disease progression.

Guidelines for COPD management include early disease identification; slowing lung function decline; relieving symptoms (e.g., dyspnea and cough); improving daily lung function; reducing exacerbations; and improving quality of life.[9] The Global Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines emphasize treatment regimens based on disease severity, including both symptom burden and exacerbation risk.[10] Current treatment guidelines recommend an incremental approach, involving the use of combinations of drug classes with different or complementary mechanisms of action, and

---

[1] Celli BR, MacNee W. ATS/ERS Task Force: standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. Eur Respir J. 2004; 23:932-46.

[2] Torpy JM, Burke AE, Glass RM. JAMA patient page. Chronic obstructive pulmonary disease. JAMA 2010;303:2430.

[3] Mintz ML, Yawn BP, Mannino DM, et al. Prevalence of airway obstruction assessed by lung function questionnaire. Mayo Clin Proc 2011;86:375-81.

[4] Heron M, Anderson RN. Changes in the Leading Cause of Death: Recent Patterns in Heart Disease and Cancer Mortality. NCHS Data Brief 2016:1-8.

[5] Rosenberg SR, Kalhan R, Mannino DM. Epidemiology of Chronic Obstructive Pulmonary Disease: Prevalence, Morbidity, Mortality, and Risk Factors. Semin Respir Crit Care Med 2015;36:457-69.

[6] Dalal AA, Shah M, D'Souza AO, Rane P. Costs of COPD exacerbations in the emergency department and inpatient setting. Respir Med 2011;105:454-60.

[7] Ford ES. Hospital discharges, readmissions, and ED visits for COPD or bronchiectasis among US adults: findings from the nationwide inpatient sample 2001-2012 and Nationwide Emergency Department Sample 2006-2011. Chest 2015;147:989-98.

[8] Simoni-Wastila L et al., Association of Chronic Obstructive Pulmonary Disease Maintenance Medication Adherence With All-Cause Hospitalization and Spending in a Medicare Population *Am J Geriatr* Pharmacother. 2012;10:201–210.

[9] Rabe KF, Hurd S, Anzueto A. et al. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease: GOLD Executive Summary. Am J Respir Crit Care Med. 2007; 176: 532-555.

[10] Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease, 2018 Report. http://goldcopd.org/wp-content/uploads/2017/11/GOLD-2018-v6.0-FINAL-revised-20-Nov_WMS.pdf. Accessed February 2018.

regular treatment with one or more long-acting bronchodilators as the disease progresses.[11] Inhaled long-acting bronchodilators, which include long-acting anticholinergics and long-acting β-agonists (LABAs), improve the airflow obstruction that characterizes COPD.

Dual therapies combined into single-inhaler, fixed dose delivery systems are important to examine in the context of simplifying treatment, improving adherence and quality of life, and managing disease progression. Umeclidinium/vilanterol (UMEC/VI 62.5/25 mcg; Anoro® Ellipta®) is a once-daily single inhaler dual LAMA/LABA therapy that was FDA-approved for the treatment of COPD in December 2013. Fluticasone/salmeterol (FLUT/SAL 250/50 mcg; Advair® Diskus®) is a twice-daily single inhaler dual therapy ICS/LABA treatment approved by the FDA for COPD treatment in November 2003. To better understand variation in burden of illness associated with these treatments, this study will provide evidence about associations between symptoms, clinical, economic, and sociodemographic characteristics, and GOLD classification for patients treated with UMEC/VI as compared to those treated with FLUT/SAL. Results of the study may be of interest to health care providers treating COPD patients and to third party payers.

The overall purpose of this study is to describe patient-reported symptom severity and claims-based healthcare utilization and costs among COPD patients treated with UMEC/VI or FLUT/SAL single-inhaler dual therapies in a real-world setting.

2 OBJECTIVES

The goal of the study is to assess the symptom burden and baseline clinical and economic characteristics in COPD patients treated with the single-inhaler dual therapies UMEC/VI or FLUT/SAL.

2.1 Primary

- The primary objective is to compare patient-reported COPD symptom burden among patients treated with UMEC/VI to those treated with FLUT/SAL.

2.2 Secondary

The secondary objectives are to:

- Describe baseline all-cause, COPD-related, and exacerbation-related health care resource utilization, costs, and clinical characteristics (e.g., comorbidities, treatment history) in patients treated with UMEC/VI or FLUT/SAL.
- Identify the GOLD classification of patients treated with UMEC/VI or FLUT/SAL.

3 RESEARCH METHODOLOGY

3.1 Study Design

This is a cross-sectional observational survey study of adult Medicare Advantage enrollees with evidence of COPD diagnosis and treatment. Cross-sectional survey data will be linked with retrospective claims data for the 12-month period (baseline) prior to the survey. An image illustrating the study process is provided in the Appendix.

---

[11] Celli BR, Decramer M, Wedzicha JA, et al. An Official American Thoracic Society/European Respiratory Society Statement: Research Questions in Chronic Obstructive Pulmonary Disease. *Am J Respir Crit Care Med*. 2015;191:e4-e27.

The study will use a health plan recruitment strategy. Use of this strategy requires that the health plan affiliated with Optum approve the study concept. Optum will also work with a central institutional review board (IRB) to obtain appropriate approvals for the study.

Study subjects will be identified using enrollment, medical, and pharmacy records from Optum's proprietary research claims database (ORD). A target population of 2,700 potential study participants will be contacted directly by mail. Patients will be asked to return the study survey if they consent to participate. The survey will collect information on patients' COPD duration, current treatment, symptoms, burden of illness, smoking history, and demographic and sociodemographic characteristics. Based on informal testing of the study instrument, the estimated time for survey completion is 10 minutes. Based on previous cross-sectional patient survey studies conducted by Optum, an overall response rate of 30-35% is expected.[12]

Following completion of data collection, results of the survey will be merged with claims data covering the 12-month baseline period for analysis. Pharmacy and medical claims data will be used to calculate all-cause and COPD-related health care utilization and costs, treatment patterns, and baseline clinical characteristics. Descriptive, IPTW, and multivariable statistical methodologies will be used.

3.1.1 Health Plan Notification

Optum will send a research study notification letter to the health plan market medical directors to alert them that a survey research study is being conducted in the field. In addition, Optum will send a courtesy letter to affiliated health plan physicians currently treating the identified patients to notify them that one or more of their patients is being invited to take part in the study.

3.1.2 Patient Invitation

Patient names and addresses will be obtained from the most recent contact information recorded in the enrollment data. Identified patients will receive a patient invitation letter on health plan and Optum letterhead, signed by the National Medical Director (NMD) of the health plan and senior research director of Optum, inviting them to participate in the survey study. The letter will describe the goals of the study, explain that participation is voluntary, and provide instructions for completing and returning the patient questionnaire.

A total of $25 in post-paid compensation in the form of a pre-paid debit card will be mailed to subjects who return the study survey.

3.1.3 Study Start-Up Data Collection Services

A survey research vendor will be utilized to handle data collection and management activities. The survey research vendor will be required to comply with the established practices of Optum as specified in their contract. Procedures relating to data handling and data transfers, database set-up, ID encryption, confidentiality, data quality management, and documentation will be specified by Optum. The vendor will be responsible for the following:

- Printing subject materials
- Shipping and tracking all subject materials
- Tracking, mailing, and reconciling post-paid debit card distribution to participants
- Complying with confidentiality procedures
- Developing the study database and entering data in accordance with the Optum's specifications
- Securely transferring the database to Optum per Optum's SOPs

---

[12] Goolsby Hunter A, Brenneman S, Brekke L. Response Rates in Direct-to-Patient Surveys. Presented at the 20th Annual Meeting of the International Society of Pharmacoeconomic Research (ISPOR); Philadelphia, May 17-20, 2015.

3.1.4 Patient Survey Implementation

The study survey packet containing the invitation letter printed on letterhead, the IRB-approved statement of informed consent, the approximately 6-page survey instrument printed in booklet form, and postage-paid business reply envelope will be sent to approximately 2,700 patients in total. All study participants will be mailed a reminder postcard two weeks following the initial mailing. The postcard will thank those participants who have completed and returned their survey and encourage those who have not yet completed their survey to do so.

Four weeks following the initial mailing, a second study packet will be sent to survey non-responders. This packet will include a second version of the study invitation letter, the IRB-approved consent statement, the survey instrument booklet, and a postage-paid business reply envelope. Some respondents whose responses are missing for key elements of the survey will be contacted by mail to request completion of missing materials.

The mailings will be sent in a personalized envelope (i.e., the address is printed directly on the envelope) and first class postage will be used. Personalized survey materials and first-class postage have all been shown to increase response rates in mail surveys.[i,ii]

Data collection will occur over the course of eight weeks[iii]. It is estimated that two waves of data collection will be needed to reach the target sample size. The same process and materials will be used for both waves of survey data collection (Wave 1 and Wave 2). While the survey is in the field, the number of completed surveys will be tracked on a weekly basis.

Overall, it is estimated 30-35 percent of patients will respond to the survey, yielding a final sample of 770 total completed surveys (approximately 385 per treatment group) based upon a sampling frame of 2,700 MA health plan members. Personalized thank-you letters will be mailed to every patient who completes the survey. The post-paid patient compensation of $25 in the form of a pre-paid debit card will be included with the thank-you mailing.

3.1.5 Study Period

An illustrative example of the study period is provided in Figure 1, which includes the sample identification, claims observation, and data collection periods described below. The 12-month period used for claims sample identification corresponds to the 12-month baseline claims analysis period for each sample.

Study period dates are estimated based on the current timeline for sample identification; changes in timelines for study tasks prior to sample identification will have a corresponding impact on the timing of later tasks.

Sample Identification Period

The sampling frame of COPD patients will be identified in Optum's research database using administrative claims data from the most recently available 12-month period at the time of sample identification, defined as the *claims sample identification period*. This period is estimated to begin 01 Jun 2017 and end 31 May 2018 for Wave 1 survey administration. This will ensure that the study sample is identified based on the most recently available pharmacy data in Optum's research database at the estimated time of identification (estimated as mid-June for Wave 1).

Medical claims for the full 12-month period will be accessed for sample identification; however, only a portion of those are complete, fully adjudicated claims. However, due to the approximate 6-month medical claims lag and 6-week pharmacy claims lag, fully complete medical and pharmacy claims data during this time period will not be available for all patients at the time of sample identification. An estimated 6 months of complete, fully adjudicated medical claims data will be available at the time of identification (01 June 2017 to 30 November 2017), with partially adjudicated claims available for the

latter 6 months of the sample identification period (01 December 2017 – 31 May 2018). Following completion of the claims lag, fully adjudicated medical claims will be accessed for the Claims Observation Period as described below. The sample identification period and the claims observation period will cover the same 12-month period for each wave of the survey.

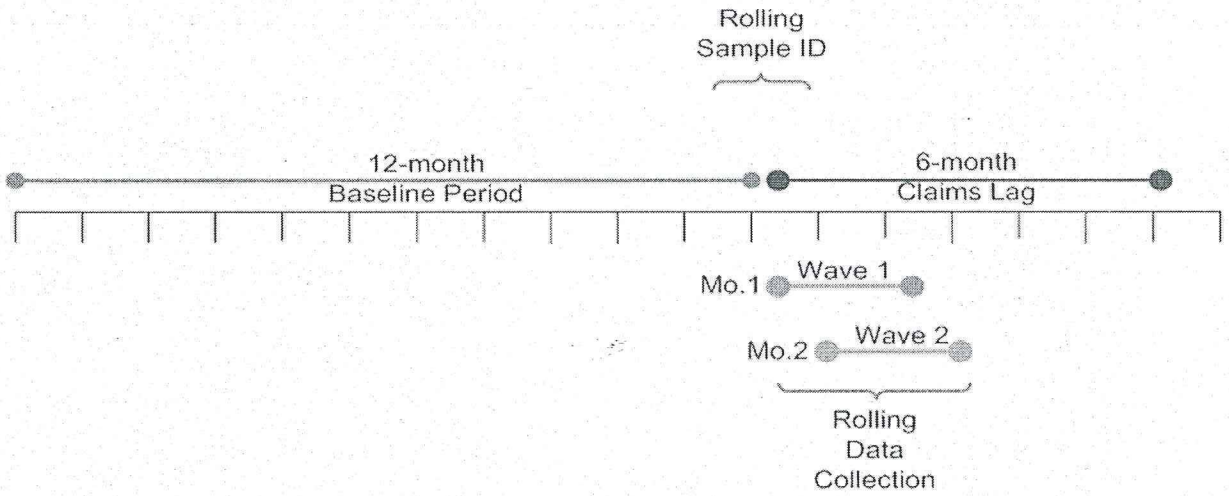
Claims Observation Period
The baseline claims observation period will be the 12-month period closest to and prior to sample identification. The observation period for Wave 1 is estimated to be 01Jun2017-31May2018.

Eligible survey respondents' complete medical and pharmacy claims data during the 12-month baseline period will be extracted for creation of the merged analytic file following completion of the 6-month medical claims lag for the latter administration (Wave 2).

Dates provided above are estimates for Wave 1; the same process and lags will apply for the subsequent sample identification period and claims observation period for Wave 2. Final identification and observation period dates will be provided in the statistical analysis plan (SAP).

Survey Fielding Period
Assuming sample identification is completed in June 2018, the 8-week survey fielding period will begin in late June-early July 2018 continue through late August-early September 2018 (Wave 1). Final survey fielding dates for Wave 1 and Wave 2 will be provided in the SAP.

Figure 1 - Study Period: Rolling Sample Identification for 12-month Baseline Claims Data Linked to Cross-sectional Survey Data



3.1.6 Link to Administrative Claims Data

Following completion of survey data collection activities for both waves and the 6-month medical claims lag, 12 months of surveyed patients' pharmacy and medical claims data for the baseline period will be re-extracted. Continuous health plan enrollment and other exclusion criteria will be applied. Survey and health care claims data will be combined into a merged analytic file.

3.2 Study Population

3.2.1 Eligibility Criteria

The study will include MA health plan members with evidence of COPD in the 12 months prior to sample identification and evidence of treatment with UMEC/VI or FLUT/SAL single-inhaler dual therapy in the

six months prior to sample identification. To be included in the target study population, patients must meet the study inclusion/exclusion criteria. Potential inclusion criteria for consideration are listed below.

3.2.1.1 Inclusion Criteria

- ≥2 ICD-10-CM COPD diagnosis codes at least 30 days apart during the 12-month baseline period. Diagnosis codes are shown below in Figure 2.

Figure 2 – Diagnosis Codes for COPD

Code	Type	Description
J40	ICD-10 Dx	Bronchitis, not specified as acute or chronic
J410	ICD-10 Dx	Simple chronic bronchitis
J411	ICD-10 Dx	Mucopurulent chronic bronchitis
J418	ICD-10 Dx	Mixed simple and mucopurulent chronic bronchitis
J42	ICD-10 Dx	Unspecified chronic bronchitis
J430	ICD-10 Dx	Unilateral pulmonary emphysema (MacLeod's syndrome)
J431	ICD-10 Dx	Panlobular emphysema
J432	ICD-10 Dx	Centrilobular emphysema
J438	ICD-10 Dx	Other emphysema
J439	ICD-10 Dx	Emphysema, unspecified
J440	ICD-10 Dx	Chronic obstructive pulmonary disease
J441	ICD-10 Dx	Chronic obstructive pulmonary disease with (acute) exacerbation
J449	ICD-10 Dx	Chronic obstructive pulmonary disease, unspecified

- ≥1 pharmacy claim in the 6 months closest to sample identification for either UMEC/VI or FLUT/SAL fixed-dose single-inhaler dual therapy COPD treatments shown in Figure 3 *and* no claims for these treatments in the first 6 months of the sample identification period.

Figure 3 – COPD Medications for Sample Inclusion

Subclass	Medication	NDC Codes
LABA/anticholinergic	umeclidinium/vilanterol 62.5/25 mcg single-inhaler therapy	00173086906 00173086910
ICS/LABA	fluticasone/salmeterol 250/50 mcg single-inhaler therapy	68258303101 00173069604 23490754201 49999081960 21695019601 55045368601 68115092460 58016460401 54569524200 54868451700 00173069602 00173069600

- Age 65 and older
- Continuous enrollment during the 12-month baseline at the time of sample identification
- Self-reported health care provider diagnosis of COPD
- Self-reported treatment with FLUT/SAL or UMEC/VI
- Ability to complete the study survey in English

3.2.1.2 Exclusion Criteria

- Patients with ≥2 diagnosis codes for asthma at least 30 days apart in the 12-month baseline period will be excluded. ICD-10-CM codes for identifying asthma are shown in Figure 4.

Figure 4 – Diagnosis Codes for Asthma

Code	Type	Description
J4520	ICD-10 Dx	Mild intermittent asthma, uncomplicated
J4521	ICD-10 Dx	Mild intermittent asthma with (acute) exacerbation
J4522	ICD-10 Dx	Mild intermittent asthma with status asthmaticus
J4530	ICD-10 Dx	Mild persistent asthma, uncomplicated
J4531	ICD-10 Dx	Mild persistent asthma with (acute) exacerbation
J4532	ICD-10 Dx	Mild persistent asthma with status asthmaticus
J4540	ICD-10 Dx	Moderate persistent asthma, uncomplicated
J4541	ICD-10 Dx	Moderate persistent asthma with (acute) exacerbation
J4542	ICD-10 Dx	Moderate persistent asthma with status asthmaticus
J4550	ICD-10 Dx	Severe persistent asthma, uncomplicated
J4551	ICD-10 Dx	Severe persistent asthma with (acute) exacerbation
J4552	ICD-10 Dx	Severe persistent asthma with status asthmaticus
J45901	ICD-10 Dx	Unspecified asthma with (acute) exacerbation
J45902	ICD-10 Dx	Unspecified asthma with status asthmaticus
J45909	ICD-10 Dx	Unspecified asthma, uncomplicated
J45990	ICD-10 Dx	Exercise induced bronchospasm
J45991	ICD-10 Dx	Cough variant asthma
J45998	ICD-10 Dx	Other asthma

- Patients with claims for both UMEC/VI and FLUT/SAL in the 6 months closest to sample identification will be excluded.
- Patients with claims for triple therapy will be excluded.
    - Triple therapy is defined as claims for formulations of ICS/LABA/LAMA, including combined monotherapy formulations and fixed dose combinations (FDC).
    - The impact of using same day fills, or 30 days, or 60 days between fills will be assessed at the time of sample identification.
- Patients with claims-based evidence of diagnosis or treatment for lung cancer during the 12-month baseline will be excluded from the final analytic sample following the claims lag.

Some exclusion criteria cannot be applied until after survey data collection and/or the medical claims lag are complete (e.g., patient-reported diagnosis and treatment, evidence of lung cancer). Criteria that can be applied at the time of sample identification will be applied at that time.

3.2.2 Sampling

The study sample is not randomly selected. A convenience (e.g., non-random) sampling approach will be used. It is estimated that all eligible patients will be invited to participate in the study. Sufficient counts are not available to match patients on demographic characteristics at the time of sample identification; however, IPTW will be used to address potential confounding due to non-treatment characteristics (see Section 4.4).

3.3 Data Sources

3.3.1 Patient-reported Measures: Cross-sectional Survey

The cross-sectional survey will collect the following measures:

3.3.1.1 Patient Reported Clinical Characteristics

- **Diagnosis**—Patients will be asked to confirm that a doctor or other health care provider has told them they have COPD. Patients will be instructed to return but not complete the survey if they do not meet this criterion.
- **General health**—A single item general health measure will be used to measure patient health on a five-point Likert scale ranging from excellent to poor.
- **COPD duration**—The length of time since the patient was diagnosed with COPD will be recorded from less than 1 year, 1-5 years, 6-10 years, 11-19 years, or 20 or more years ago.
- **COPD treatment**—Patients will be asked to indicate their current COPD treatment.
- **Height**—Patients will be asked to provide their height in feet and inches. Total height in inches will be calculated.
- **Weight**—Weight in pounds will be collected.
- **BMI**—Body Mass Index will be calculated from patient report of height and weight. Patients will be assigned to one of four categories: Underweight (<18.5 kg/m^2), Normal Weight (18.5-24.9 kg/m^2), Overweight (25-29.9 kg/m^2), or Obese (≥30 kg/m^2).
- **Smoking behavior**
    - **Smoking status**—Status will be recorded as current smoker, former smoker, never smoked and lives with a smoker, or never smoked and no one in their household smokes.
    - **Cigarettes smoked per day, former smokers**—Former smokers will be asked to report how many cigarettes, on average, they smoked per day as an adult.
    - **Cigarettes smoked per day, current smokers**—Current smokers will be asked to report how many cigarettes, on average, they smoke per day.
    - **Pack year**—Pack year will be calculated by multiplying the total number of cigarettes smoked per day by the total number of years smoked, then dividing by 20 (1 pack has 20 cigarettes).
    - **Age started smoking**—Age in years.
    - **Age stopped smoking**—Age in years will be collected from former smokers.
    - **Total years smoked**—The number of years the patient has smoked will be calculated by subtracting patient-reported age when started smoking from age when quit smoking, or for current smokers, then subtracted from age at time of survey.

3.3.1.2 Patient Reported Demographic and Sociodemographic Variables

- **Age**—Age will be calculated from the 1-item survey question asking patients to provide their year of birth; claims data may supplement patient-report for subjects who do not respond to this item.
- **Age groups**—Subjects will be assigned to an age group based on the distribution of the age measure, for example: 65-69, 70-74, and 75+.
- **Gender**—Gender will be captured from a single-item survey question; claims data may supplement patient report for subjects who do not respond to this item.
- **Race**—Whether the respondent considers themselves to be American Indian or Alaskan Native, Asian, Black or African American, Native Hawaiian or Pacific Islander, White, and/or Other race.
- **Ethnicity**—Whether the respondent identifies as Hispanic/Latino.

- **Population density**—Defined by a single question asking patients to identify whether the community they live in is urban, suburban, or rural.
- **Marital status**—Defined by a single question asking patients to identify whether they are single/never married, living with partner, married, separated, divorced, or widowed.
- **Educational attainment**—The highest level of education the patient has completed will be measured as some high school, high school or equivalent (e.g., GED), some college but no degree, two-year college, four-year college, or graduate school.
- **Income**—Annual income for the last completed calendar year will be collected as less than $30,000, $30,000-$49,999, $50,000-$74,999, $75,000-$99,999, $100,000 or greater, or I choose not to answer.

3.3.1.3 Patient-reported COPD Symptom Burden

- **Dyspnea**—Breathlessness will be assessed using the Modified Medical Research Council Dyspnea scale (mMRC), a single item (0-4 scale) to assess patients' current level of dyspnea.[iv,v] The mMRC is comprised of five statements that describe almost the entire range of respiratory disability from none (Grade 0) to almost complete incapacity (Grade 4). The score (Grade) is the number that best fits the patient's level of activity.
- **Burden of illness**—Condition-related well-being will be assessed using the COPD Assessment Test (CAT), an eight-item validated questionnaire developed by GSK to determine the impact of COPD on the patient's well-being and daily life.[vi, vii] The items of CAT measure the most salient symptoms of COPD, including cough, chest tightness, breathlessness, and activity limitation attributed to COPD symptoms. Each of the items of CAT is scaled on a 0-5 point scale with higher values indicative of worse health status. The item response values of CAT are summed to produce a single score that ranges from 0 (best health status) to 40 (worst health status). The reference period is current.

3.3.2 Claims-based Measures

General descriptions of claims data based study variables are provided below. All variable definitions, including diagnosis codes, procedures codes, medication lists, and algorithms will be provided in detail in the SAP. Patients' clinical and treatment characteristics will be measured in the baseline period, defined as the 12-month period prior to survey sample identification.

3.3.2.1 Claims-based Clinical and Treatment Characteristics

- **Baseline Quan-Charlson comorbidity score**—A comorbidity score will be calculated based on the presence of diagnosis codes on medical claims in the baseline period.[viii,ix] The Quan-Charlson comorbidity score will also be categorized into the following groups: zero, one to two, three to four, and five or more. The mean comorbidity score will be presented.
- **Baseline comorbid conditions**—General comorbid conditions will be defined using the Clinical Classifications Software managed by the Agency for Healthcare Research and Quality (AHRQ)[x]. This measure generates indicator variables for specific disease conditions based on ICD-10-CM diagnoses codes. The top 20 comorbid conditions will be presented.
- **COPD-related comorbid conditions**—In addition, whether or not patients have evidence of specific COPD-related comorbidities will be assessed for approximately 10 COPD-related conditions. Conditions of interest may include, for example, upper and lower respiratory infections, type 2 diabetes mellitus (T2DM), pulmonary hypertension, and obstructive sleep apnea. A binary indicator will be created for each condition to identify whether or not the patient has claims-based evidence of that condition. Counts and percents will be presented. Codes for conditions of interest will be specified in the SAP.

- **Baseline medications**—A count of unique medications filled in the baseline period will be calculated.
- **Baseline medication dispensings**—A count of unique dispensings for all medications filled in the baseline period will be calculated.

3.3.2.2 Claims-based Health Care Resource Utilization and Costs

- **All-cause baseline health care resource utilization**—Binary indicators and counts of ambulatory (physician office and hospital outpatient) visits, ED visits, and inpatient stays will be calculated in the 12-month baseline period.
- **All-cause baseline health care costs**—Consumer Price Index-adjusted costs will be assessed for the 12-month baseline period. Costs will be calculated as total costs, pharmacy costs, and medical costs; medical costs include ambulatory (physician office and hospital outpatient) costs, emergency costs, inpatient costs, and other costs.
- **COPD-related baseline health care resource utilization**—Binary indicators and counts of ambulatory (physician office and hospital outpatient) visits, ED visits, and inpatient stays will be calculated in the 12-month baseline period. Utilization will be defined as COPD-related when a diagnosis code for COPD (Figure 2) is in the primary position.
- **COPD-related baseline health care costs**—Consumer Price Index-adjusted costs will be assessed for the 12-month baseline period. Costs will be calculated as total costs, pharmacy costs, and medical costs; medical costs include ambulatory (physician office and hospital outpatient) costs, emergency costs, inpatient costs, and other costs. Costs will be defined as COPD-related if the claim has a diagnosis code for COPD (Figure 2) in the primary position or is for a medication used to treat COPD. Codes for COPD treatment will be provided in the SAP (and will include those in Figure 3).

  Payments from Medicare are estimated based on coordination of benefits information obtained by the health plan in its usual course of business. This study will incorporate the amounts estimated to be paid by other payers (i.e., Medicare) for a total paid or allowable amount[xi].
- **COPD exacerbations**—Whether the patient has evidence of a COPD exacerbation during the 12-month baseline period will be captured. The count of exacerbations will be presented. GSK's harmonized exacerbation definition will be used; the definition will be included in the SAP.
- **COPD exacerbation-related baseline health care costs**—CPI-adjusted costs will be assessed for the 12-month baseline period. Total costs (pharmacy + medical) will be reported. Costs will be attributed to an exacerbation per GSK's harmonized exacerbation definition.

3.3.2.3 Claims-based Patient Demographic Measures

- **Age**—Age will be defined as of the first date of the survey fielding period.
- **Age groups**—Subjects will be assigned to an age group as described in Section 3.3.1.2
- **Gender**—Gender will be captured from enrollment data; subjects with undefined gender will be removed from the study sample.
- **Geographic region**—The U.S. region in which the study subject is enrolled in a health plan will be determined and reported and states will be categorized into four geographic regions in accordance with the U.S. Census Bureau's region designations; the four health plan regions are presented in Figure 5 below.

Figure 5 – Health Plan Regions

Region	Division	State
Northeast	New England	CT, MA, ME, NH, RI, VT
	Mid Atlantic	NJ, NY, PA
Midwest	East North Central	IL, IN, MI, OH, WI
	West North Central	IA, KS, MN, MO, ND, NE, SD
South	South Atlantic	DC, DE, FL, GA, MD, NC, SC, VA, WV
	East South Central	AL, KY, MS, TN
	West South Central	AR, LA, OK, TX
West	Mountain	AZ, CO, ID, MT, NM, NV, UT, WY
	Pacific	AK, CA, HI, OR, WA

3.3.3 Medicare Advantage Data Source

Medical and pharmacy claims data are available for approximately 4.2 million enrollees since 2006 that are enrolled in Medicare Part C (commonly referred to as the Medicare Advantage program) through an offering associated with Optum. Prior to 2006, data were available for approximately 500,000 of these enrollees. Medicare Advantage enrollees choose to receive all of their health care services through a provider organization in lieu of Medicare Part A/B coverage (commonly referred to as Medicare Fee for Service).

Claims for pharmacy services are typically submitted electronically by the pharmacy at the time prescriptions are filled. The claims history is a profile of all outpatient prescription pharmacy services provided and covered by the health plan. Pharmacy claims data include drug name, dosage form, drug strength, fill date, days of supply, financial information, and de-identified patient and prescriber codes, allowing for longitudinal tracking of medication refill patterns and changes in medications.

Medical claims or encounter data are collected from all available health care sites (inpatient hospital, outpatient hospital, ED, physician's office, surgery center, etc.) for virtually all types of provided services, including specialty, preventive and office-based treatments. Medical claims and coding conform to insurance industry standards. Claims for ambulatory services submitted by individual providers, e.g., physicians, use the Health Care Financing Administration (HCFA)-1500 or Centers for Medicare and Medicaid Services (CMS)-1500 formats. Claims for facility services submitted by institutions, e.g., hospitals, use the Uniform Billing (UB)-82, UB-92, UB-04, or CMS-1450 formats. Medical claims include: multiple diagnosis codes recorded with the International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) diagnosis codes; procedures recorded with ICD-10-CM procedure codes, Current Procedural Terminology (CPT), or Healthcare Common Procedure Coding System (HCPCS) codes; site of service codes; provider specialty codes; revenue codes (for facilities); paid amounts; and other information. Facility claims do not generally include medications dispensed in hospital.

Pharmacy claims are typically added to the research database within six weeks of dispensing. Approximately six months following the delivery of services are required for complete medical data for MA enrollees.

3.4 Endpoints

Patient-reported outcomes, which include all primary endpoints, will be measured using the cross-sectional survey at the time of survey administration. Claims-based characteristics, including economic endpoints, will be measured in the baseline period, which is the 12-month period prior to survey sample identification.

3.4.1 Primary Endpoint

- **Patient-reported symptom burden**. Scores will be presented from validated instruments measuring dyspnea (mMRC) and condition-related burden of illness (CAT). Mean total, summary, and/or domain scores and standard deviations will be presented for each scale/item (as described in Section 3.3.1.3). The proportion of patients with high and low symptom burden on both PRO measures will be calculated. Results will be presented for the overall study sample and by treatment cohort (FLUT/SAL or UMEC/VI).
- **Proportion of patients reporting COPD symptoms while treated with FLUT/SAL or UMEC/VI.** The proportion of patients who report COPD symptoms will be computed for survey respondents. Two calculations will be presented: the number of COPD patients who report COPD symptoms as measured by a CAT total score ≥10, divided by the total number of respondents with a CAT total score; and the number of COPD patients with an mMRC score of 2-4, divided by the total number of respondents with an mMRC score. Results will be presented for the overall study sample and by treatment cohort.
- The mMRC and CAT instruments are included to correspond with current GOLD guidelines.

3.4.2 Secondary Endpoint(s)

- **Baseline health care resource utilization and costs.** All-cause and COPD-related health care costs will be presented for the 12-month baseline period. Mean, median, and standard deviations will be presented for total costs and utilization and for each category of costs (i.e., pharmacy + medical, and the subcategories of medical costs as described in Section 3.3.2.2). Costs and counts will be presented overall and for the subset of patients utilizing the selected type of service. The proportion of patients receiving services for each type of service and the proportion of patients with non-zero costs for each type of cost will be calculated. Results will be presented for the overall study sample and may be stratified by relevant subgroups for analysis (e.g., by low versus symptom burden as measured by patient mMRC and CAT scores).
- **Patient clinical characteristics.** The proportion of patients with evidence of selected clinical characteristics will be presented. The top 20 claims-based comorbidities and up to 10 COPD-related comorbidities will be identified. The percent of patients with each condition will be identified. Mean and standard deviation will be calculated for the Quan-Charlson comorbidity score. Results will be presented for the overall study sample and may be stratified by relevant subgroups for analysis (e.g., by low versus high symptom burden as measured by patient mMRC and CAT scores).
- **Patient GOLD classification.** The proportion of patients in each GOLD category of symptom burden and exacerbation will be presented. The calculation will use count of COPD exacerbations during the baseline combined with CAT total score and mMRC score to create mutually exclusive counts for each category (A, B, C, D) (Section 10, Appendix 3).

3.5 Sample Size / Power Calculations

The sample size required for the study is based on several requirements. First, the study's primary outcome measures will be proportions (e.g., the proportion of patients experiencing symptoms who report low versus high symptom severity as measured by the mMRC and CAT total score). The estimated sample size is determined by the value and desired precision of the measured proportions.

Figure 6 displays the sample sizes required for various proportions, assuming a 95% confidence interval and a precision of ±0.05. Because the actual population proportions (e.g., the proportion of patients in each treatment cohort with mMRC scores ≥2 and the proportion of patients in each treatment cohort with CAT total scores ≥10) are unknown, the largest sample size indicated with that precision was selected. This largest value is for a proportion of 0.5 and assures a 95% confidence interval of ±0.05 or

better for all proportions observed. Thus, a sample size of 385 completed patient surveys per treatment cohort was selected.

Figure 6 – Sample Sizes Required for Various Proportions

95% Confidence Interval	Proportion								
	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
±0.05	139	246	323	369	385	369	323	246	139

Since the proportion of patients experiencing symptoms will be determined from the survey and a variety of exclusion criteria can be applied only after data collection, a sampling frame of 2,700 patients is needed for an estimated evaluable sample size of 385 patients per treatment cohort. The sampling frame will be comprised of equal proportions per medication cohort (n=1,350 per cohort); however, proportional response rate by treatment cohort cannot be guaranteed. Based on the feasibility counts, Optum estimates that two waves of sample identification and survey data collection will be required to meet the target sample size. Sample identification for wave 2 will occur as soon as sufficient sample is available, which is estimated to be the month following wave 1 sample identification. The availability of the second wave will be assessed during protocol development.

Figure 7 – Estimated Final Evaluable Sample Size

Sampling frame per treatment cohort	n=1,350
Estimated response rate (30-35%)	405-470
Exclusion criteria applied following claims lag (2-5% ineligible)	8-24
Estimated final evaluable sample size per treatment cohort	385-446
Estimated total final evaluable sample size	770-892

A recent feasibility assessed counts of MA patients with ≥2 COPD diagnoses initiating UMEC/VI or FLUT/SAL and found sufficient counts to develop a sampling frame. Based on these counts, there will likely not be enough sample who meet all eligibility criteria to complete data collection in one wave. Therefore, it is estimated that two waves of sample identification will be required to meet the target sample size based on these estimates of counts of eligible patients and response rate.

Table 1 – Feasibility Counts

Sample Attrition	Medicare Advantage n (%)
Patients with evidence of >=2 codes for treatment with Anoro Ellipta OR >=2 codes for treatment with Advair Diskus 250mg from 01Mar2017-31Aug2017	28,157
Patients with >=2 COPD diagnosis from 01Sep2016-31Aug2017	4,391 (15.6%)
Patients aged >=65 on 01Sep2016	3,224 (11.5%)
Patients with 12 months continuous enrollment 01Sep2016-31Aug2017	2,070 (7.4%)
Patients with claims for only Anoro Ellipta from 01Mar2017- 31Aug2017 and no evidence of Anoro Ellipta or Advair Diskus use from 01Sep2016-28Feb2017	**977** (47.2%)
Patients with claims for only Advair Diskus 250 mg from 01Mar2017-31Aug2017 and not evidence of Advair Diskus or Anoro Ellipta use from 01Sep2016-28Feb2017	**1,064** (51.4%)
Patients with >= 1 claim for Anoro Ellipta AND Advair Diskus 250mg from 01Mar2017-31Aug2017*	29 (1.4%)

*Patients with claims for both treatments in the 6 months prior to sample identification will be excluded.

3.6 Hypotheses

The primary objective is to assess patient-reported COPD symptom burden among patients treated with UMEC/VI as compared to FLUT/SAL. The hypothesis for the primary objective is:

Null hypothesis:
1A) Patient-reported symptom severity as measured by CAT total score will not differ for UMEC/VI patients as compared to FLUT/SAL patients.
1B) Patient-reported symptom severity as measured by mMRC score will not differ for UMEC/VI patients as compared to FLUT/SAL patients.

Alternative hypothesis:
1A) Patient-reported symptom severity as measured by CAT total score will differ for UMEC/VI patients as compared to FLUT/SAL patients.
1B) Patient-reported symptom severity as measured by mMRC score will differ for UMEC/VI patients as compared to FLUT/SAL patients.

Hypotheses are presented for the primary objective since the sample size / power calculations are based on assessment of the primary objective.

4 DATA ANALYSIS CONSIDERATIONS

4.1 Statistical Methods

The final analysis for the study will comprise descriptive analyses of cross-sectional survey data linked to baseline claims data and multivariable modeling of the primary study outcome. Analyses will be conducted using the SAS statistical software package. Illustrative table shells and figures showing the format for the analysis results will be provided separately and will include specific statistical tests that will be undertaken in support of each study objective.

4.2 Data Editing and Missing Data

Data management activities will comply with all applicable regulatory requirements and established practices for collection and validation of data and preparation of a reliable database for statistical analysis. The project specific procedures will encompass all of the data management activities associated with the implementation of the study observational plan.

The scores for each validated survey measure (i.e., mMRC, CAT) will be calculated according to the individual scoring instructions provided by the license holders. For the mMRC, a missing value is assigned if none of the five response options is selected by the patient. A CAT total score will be calculated if 2 or fewer items are missing. If 1 or 2 items are missing, the missing items will be set to the average of the non-missing items. If more than 2 items are missing, the CAT total score will be set to missing.[xii]

Scoring instructions for all survey measures will be described in the SAP. Additionally, the specifications for a survey to be considered complete (vs. partially complete or incomplete) will be defined in the SAP. Key elements for survey completion will include responses to the mMRC and CAT, in addition to confirmation of COPD diagnosis and treatment. Additional key measures may be identified. Responses that are invalid or out of range will be considered missing. Respondents who have completed a survey that is missing a few non-key items will be included in the study analysis, but will be removed from analysis of the specific measures for which they have missing data. Removing patients from specific analyses will change the available sample size for certain measures. Because the available N may change, the valid N associated with each study measure will be reported. Select demographic claims data (e.g., age, gender) may supplement missing self-report information where possible in order to maintain sample size. Because this is a one-time, cross-sectional survey, loss of participants during subsequent observations will not be an issue.

4.3 Descriptive Analyses of Survey and Claims Data

Administrative Claims Data

Health service utilization and costs for the baseline period will be identified and presented descriptively for all-cause and COPD-related total HCRU and costs. For dichotomous variables, N (%) will be presented. For continuous variables, means, standard deviations, medians, ranges, and percentiles will be presented as appropriate. Costs will be defined as costs to the health plan and estimated costs paid by Medicare; costs will not include costs to the patient (copays, deductibles). Detailed descriptions of cost and utilization measures are provided in 3.3.2.2.

Survey Data

All study variables will be analyzed descriptively. Numbers and percents will be provided for dichotomous and polychotomous variables. Relevant measures of centrality such as means and medians will be presented for continuous measures, as well as variance measures such as standard deviations and percentiles (e.g., p25, p75). Comparisons of demographic characteristics and outcome measures will be provided. Appropriate tests (e.g., t-test, Mann Whitney-U test, chi-square test) will be used based on the distribution of the measure. For example, a one sample t-test could be used to show difference from the sample mean on a given PRO (such as CAT score) as compared to an expected value (such as CAT scores among older adults from the published literature). Pearson correlation coefficients are utilized for bivariate comparisons of CAT total score and mMRC score with selected clinical and sociodemographic characteristics.

In addition to descriptive analyses of all survey variables, survey responders and non-responders will be compared on select demographic characteristics in order to assess the generalizability of the evaluable survey sample.

Primary Objective

In order to fulfill the primary objective, the proportion of patients with low and high symptom severity will be identified based on responses to the mMRC and CAT measures. Patients may be categorized into multiple levels of severity based on the distribution of responses to one or both measures (e.g., patients with a high degree of dyspnea per the mMRC score and severe symptom burden per the CAT score may be categorized as most severe for relevant analyses).

Classification of symptom burden is based on established cutpoints for validated instruments. For the mMRC, scores ≥2 reflect high levels of dyspnea, and CAT scores >10 reflect high levels of COPD burden, with a 5-point difference between stable and exacerbating patients established during instrument validation[xiii]. Correlation between these measures of symptom burden will be compared descriptively using t-tests.

Secondary Objectives

One secondary objective is to describe baseline all-cause, COPD-related, and exacerbation-related health care resource utilization, costs, and clinical characteristics. Using cross-tabulations and chi-square tests, we will describe characteristics of patients with high and low utilization and costs in terms of their symptom burden (e.g., mMRC and CAT scores), clinical characteristics (e.g., time since COPD diagnosis, treatment, comorbidity score, smoking history), and demographic and sociodemographic characteristics (age, gender, race, ethnicity, marital status, educational attainment, area of residence, and income level). Utilization and cost levels may be based on the distribution of scores among the survey sample and/or relevant literature.

The additional secondary objective is to describe patient GOLD classification. The proportion of patients in each GOLD category of combined symptom burden and exacerbation risk will be presented by treatment cohort. The calculation will use count of COPD exacerbations during the baseline combined with CAT total score and mMRC score to create mutually exclusive counts for each category (A, B, C, D)(see Section 10, Appendix 3).

Sensitivity Analysis

To assess whether differences are observed in patients based on their history of COPD exacerbations, sensitivity analysis will be conducted. This analysis will describe patient-reported symptom burden, demographic and sociodemographic characteristics, and baseline clinical characteristics for patients with and without claims-based evidence of exacerbations during the 12-month baseline period.

4.4 Inverse Probability of Treatment Weighting

To address the objective of comparing symptom burden of patients by treatment cohort (FLUT/SAL as compared to UMEC/VI), inverse probability of treatment weighting will be applied to reduce confounding on observed clinical, demographic, and sociodemographic characteristics shown in Appendix 4 (Section 11). The goal of adjusting for differences in non-treatment characteristics between treatment cohorts is to balance the distribution of covariates that would occur if participants were randomly assigned to treatment cohorts *a priori.*[xiv] The weights in IPTW are based on each individual's probability of receiving a specific treatment given the confounders, which is known as the propensity score (PS). Use of methods that make adjustments via propensity scores is an appropriate way to create balance in the covariate distribution among cohorts.[xv,xvi, xvii]

A logistic regression will be employed with a dependent or response that is conditioned on a set of predetermined covariates. Each subject will be assigned a propensity score weight and the adjustment for potential treatment selection bias will be made by weighting the expected value of an outcome or outcomes of interest in the corresponding multivariable model by a function of the inverse propensity score and sample size of the cohort.[xviii]

The study team will examine the quality of the propensity score determination via descriptive techniques (e.g., examination of the overlap in histograms of the respective cohort propensity score distributions, and assessment of goodness of fit of the logistic regression) to ensure the weighting approach is sound. Following review of these diagnostics, the study team will identify the characteristics of any outliers and will determine whether those subjects should remain in the model based on their clinical and analytic characteristics and whether these characteristics would unduly impact subsequent multivariable analysis. The p values determined from the weighted and unweighted descriptive analysis of covariates included in the IPTW analysis will be compared.

4.5 Multivariable Analysis

Multivariable analysis of the primary objective will be conducted using appropriate regression models (e.g., ordinary least squares [OLS], logistic regression, generalized linear models [GLM], negative binomial regression, Poisson [log-linear] regression) based on the distribution of the measure.

Two multivariable models are planned for the study. Final outcomes for the multivariable analysis will be selected jointly by Optum and GSK after review of descriptive results, including PRO scores, IPTW results, utilization, and costs. Specific predictors to be included in the models will be determined based upon clinical rationale and statistical significance. In general, the statistical significance of each predictor variable will be examined in the univariate regression analysis. Variables that demonstrate at least marginal statistical significance ($p<0.1$) in univariate analysis will be included in the multivariate models, along with relevant clinical variables regardless of their statistical significance.

Multivariable analyses to assess the primary objective are outlined here. The dependent variable of the first model will be the CAT total score, and the dependent variable of the second model will be the mMRC score. Each model will examine the relationship between symptom burden and treatment (FLUT/SAL or UMEC/VI). Each model will include clinically and analytically meaningful covariates based on the IPTW analysis; it is assumed that the covariates will be the same for each model.

Depending on the distribution of the measure, the CAT and mMRC outcomes may be analyzed as a continuous or dichotomous measure. In general, continuous variables are analyzed using ordinary least square regression models. If transformed into a dichotomous measure (e.g., patients with a CAT total score >10 or an mMRC score ≥2), the probability of an event will be modeled using logistic regression. Logistic regression models fit a maximum-likelihood logit model. Following standard procedure, regression diagnostics will be performed for each model to assess goodness of fit and violations of model assumptions (e.g., multicollinearity, heteroskedasticity). When there are violations of the model, they will be noted and appropriate corrections made to the data (i.e., typically through transformation of either the independent or dependent variables) or in the method of estimation. In addition, the fitted and the observed data will be examined to uncover outliers, their effect on the analysis, and possible misspecification of the initial equation.

5 LIMITATIONS

While claims data are extremely valuable for the efficient and effective examination of health care outcomes, treatment patterns, health care resource utilization, and costs, claims data are collected for the purpose of payment and not research. Therefore, there are certain limitations associated with the

use of claims data. First, presence of a claim for a filled prescription does not indicate that the medication was consumed or that it was taken as prescribed. Second, medications filled over-the-counter or provided as samples by the physician will not be observed in the claims data. Third, presence of a diagnosis code on a medical claim is not positive presence of disease, as the diagnosis code may be incorrectly coded or included as rule-out criteria rather than actual disease. Finally, certain information is not readily available in claims data that could inform study outcomes, especially certain clinical and disease-specific parameters; therefore, spirometry results are not included. To mitigate these limitations, multiple claims for diagnosis and treatment are required for sample identification and patients are asked about their COPD diagnosis and current treatment in the survey. Limitations of survey data may include sampling error, coverage error (including lower response rates for commercially versus publicly sponsored studies), and measurement error. Because the population will be selected from Medicare Advantage enrollees, the results may not be generalizable to patients with Medicare Fee for Service coverage, uninsured populations, or to younger patients with COPD. To limit the proportion of patients who may have dual insurance coverage, patients are required to be age 65 or older at the beginning of the sample identification period; however, it is possible some patients may have dual coverage, which may underestimate their costs and utilization.

6 STUDY CONDUCT MANAGEMENT & ETHICS

6.1 Ethics Committee/IRB Approval

6.1.1 Health Plan Approval

Use of a health plan recruitment strategy requires that the study concept be approved by the large U.S. health plan affiliated with Optum since patients will be identified in the plan's health care claims data. Optum will submit the study design to the health plan's Research Review Board (RRB) for approval. Approval by the RRB is not guaranteed. Since the study involves Medicare Advantage patients, the study concept must also be approved through the compliance process of the large U.S. health plan affiliated with Optum. Optum will submit the study concept to the health plan's Medicare Advantage compliance review team for approval. Approval by the health plan's Medicare Advantage compliance review team is not guaranteed.

6.1.2 Health Plan Medicare Advantage Plan Approval

Use of a health plan recruitment strategy with Medicare Advantage plan patients requires all patient data collection and communication materials (e.g., survey instrument, informed consent statement, study invitation letters, reminder postcard, missing materials letter, and thank you letter) to be approved by the health plan's compliance review team to ensure appropriate communications with Medicare Advantage plan patients. After GSK has approved the study protocol and survey instrument, Optum will submit all patient communication materials to the health plan's MA compliance review team for approval. Approval by the health plan's MA review team is not guaranteed.

6.1.3 Institutional Review Board Approval

Optum will work with a central IRB to obtain appropriate approvals for the study, including patient recruiting and consenting procedures. Regulatory, independent ethics, and any other review and approval will be obtained and maintained in the study file. The subject informed consent statement will be approved by a central IRB prior to the start of the study.

Optum will prepare and submit the appropriate documents to a central IRB for review and approval. The submission package will contain documents specified by the central IRB, which will include:

- Complete or sections of the study protocol

- Data collection instruments
- Subject consent materials
- Documentation of forms of communication with the study subject

Optum will communicate directly with the central IRB to address any questions and provide any additional information requested in connection with the central IRB's review. Optum will not begin subject recruitment until IRB approval of all components of the study is obtained.

6.2 Informed Consent

As part of the IRB submission, Optum will request a waiver of documentation of informed consent. If approved, the study packet will contain IRB-approved informed consent statement, which does not require signature. The consent form will ask patients to return the study survey if they elect to participate in the study. Consent is implied when patients return study materials, and signed consent will not be obtained. Patients may withdraw consent at any time.

Consent documentation informs participants that they may withdraw consent by contacting the Optum study project manager, including their contact information. After aggregate results have been published, subject data cannot be removed from published results. In the rare event that a subject withdraws consent after agreeing to participate (i.e., by returning a completed survey), it is usually proximal to the survey data collection period and prior to results dissemination. In that case, the subject will be removed from all analyses and their survey destroyed.

6.3 Data Protection

Optum will adhere to ethical and legal requirements for protection of participants' privacy and confidentiality, Optum's established practices for collection and validation of data, and preparation of a reliable database for statistical analysis.

Physical Security

Optum has established strict physical and environmental controls to protect information assets and information technology systems from unauthorized physical access, safeguard against reasonable environmental hazards, and preserve the safety of personnel.

Network Security

Optum employees complete a two-tiered security awareness-training program. The first tier requires all employees to understand the basic principles of security. The second tier involves security awareness training and additional technical training for our information technology staff. Employees receive ongoing training.

Optum stores its customers' subject identifiable medical information in secure computer systems with strict access controls. Access to the data is permitted only as needed to perform company responsibilities. Data from the automated database can only be accessed through a secure protocol involving personal logon information to establish the identity of the person accessing the data. Only staff approved for access are provided logon access, and data access activity is monitored.

Optum also employs functional and technical security processes to maintain the confidentiality of customer data. All data accessed by approved staff are identified with an ID number that is not directly translatable into a medical record number or social security number. A key to the ID number is maintained at Optum that allows the identification of a specific person from the ID number. This key is kept separate from the database and only accessed when needed to identify a member's records. Additionally, data containing protected health information (PHI) can only be extracted from the

automated database after IRB or Privacy board approval has been received. After it is extracted, all PHI will be housed in a password protected electronic folder.

6.4 Personally Identifiable Information (PII)

Personally identifiable information (PII) and protected health information (PHI) will not be transmitted to GSK according to GSK policy.

6.5 Adverse Event (AE) Reporting

There are no pharmaceutical products or other treatments administered as part of this study. The informed consent form advises subjects to contact their physician directly if they experience any illness, health problems or concerns, but some subjects may instead spontaneously report a medical event in the margin of the written survey or via an unsolicited telephone conversation with Optum or the survey vendor.

If Optum or a vendor working on behalf of Optum receives an adverse event for a GSK product, the adverse event will be reported to GSK pharmacovigilance within one working day. The minimum criteria for reporting an adverse event include an identifiable subject, reporter, medicinal product, and event.

Should a serious adverse event not involving a GSK product occur and become known to Optum, Optum or a vendor working on behalf of Optum will complete a MedWatch form and send the form to the Food and Drug Administration (FDA) Safety Information and Adverse Event Reporting Program within one working day and to GSK pharmacovigilance. The FDA is then responsible for any further reporting obligations in accordance with applicable law.

A serious adverse event (SAE) is defined as any untoward medical occurrence that at any dose:

- Results in death.
- Is life-threatening—the term "life threatening" refers to an event in which the subject was at risk of death at the time of the event; it does not refer to an event that hypothetically might have caused death if it were more severe.
- Requires inpatient hospitalization or prolongation of existing hospitalization.
- Results in persistent or significant disability/incapacity.
- Leads to a congenital anomaly/birth defect.
- Is any other important medical event that satisfies may require intervention to prevent the items listed above.

It is recommended that as much information as possible be collected at the time of the SAE report submission; however, the minimum required data elements for an SAE case to be valid are an identifiable reporter, an identifiable patient, an adverse reaction, and a suspect product.

6.6 Data Storage/Archival

6.6.1 Record Retention

To enable evaluations and/or audits from publishers or GSK, Optum agrees to keep records, including the identity of all participants (sufficient information to link claims data and survey data) and electronic and paper copy data. Optum will store study records in a secure location for a period of six years.

7 EXTERNAL INVOLVEMENT

7.1 Third Party Supplier

The survey data collection vendor intended to be contracted for the study is:

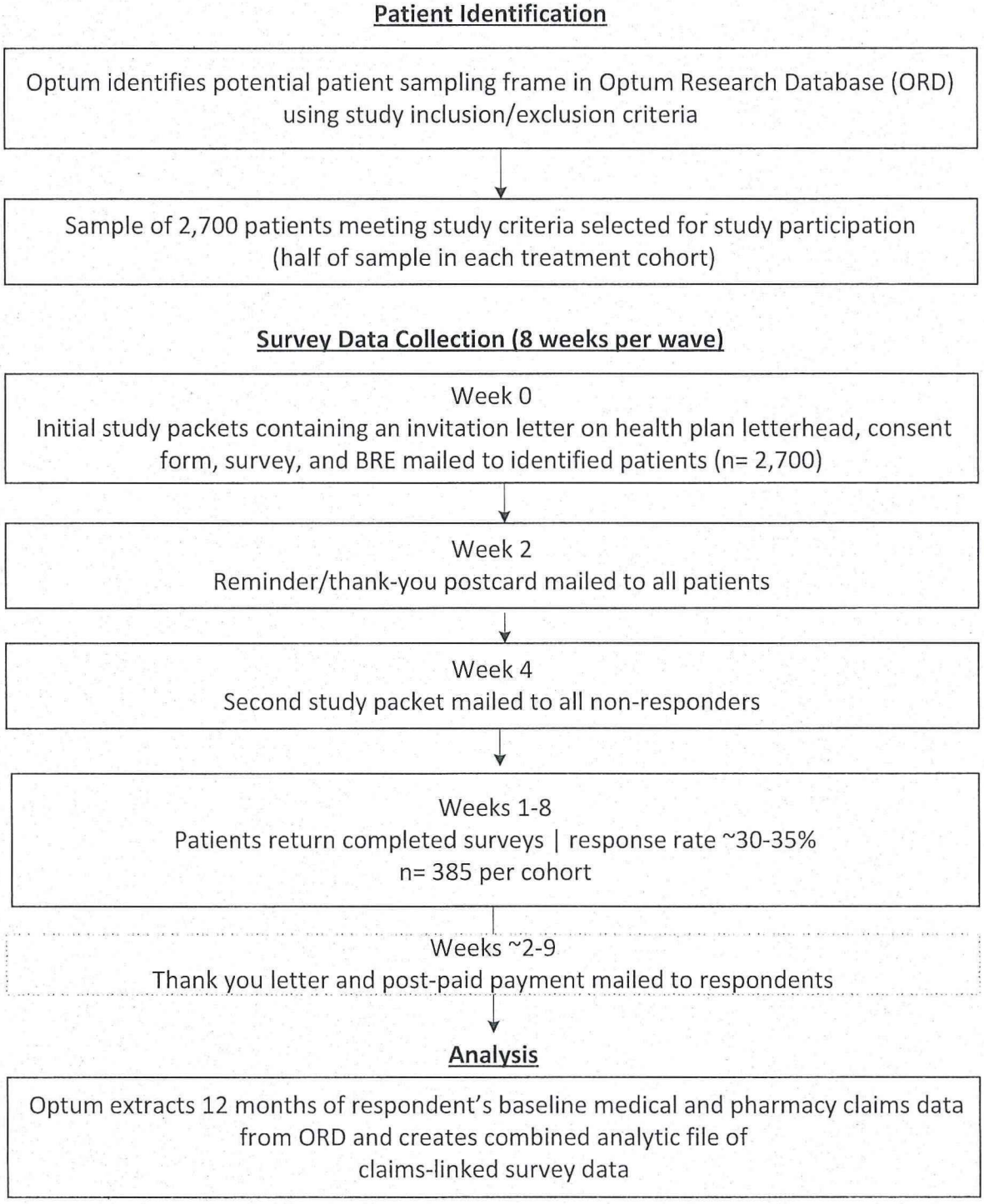
ANA Research
5155 East River Road, Suite 409
Minneapolis, MN 55421
Phone: PPD
ana-inc.com

ANA project staff will be assigned following contracting, which is currently estimated to occur in Q2 2018.

7.2 External Expert/Health Care Professionals (Consultants & Research PIs)

Not applicable.

8 APPENDIX 1. SCHEMATIC OF CLAIMS-LINKED CROSS-SECTIONAL SURVEY STUDY



9 APPENDIX 2. GSK HARMONIZED DEFINITION OF COPD EXACERBATION

The current GSK harmonized definition of COPD exacerbations is provided below. The final definition may be modified at the time of analysis if a more recent harmonized definition is available. The final definition will be provided in the SAP. Exacerbation events and episodes will be identified in the claims period for the 12-month claims observation period (baseline) following the claims lag.

1. EXACERBATION EVENTS

1A. SEVERE EXACERBATION EVENTS

Severe Exacerbation (Hospitalization) Events. COPD-related hospitalization is defined as a severe exacerbation event. Unique hospital admissions and COPD-related hospitalizations are defined with the ICD-10-CM diagnosis codes used to identify COPD-related hospitalizations. Codes will be provided in the SAP.

- If the patient had a qualifying COPD ED event that led to a COPD-related hospitalization, this will count only as a COPD-related hospitalization.
- Hospitalization "via" ED visit or ED visit "leading" to COPD-related hospitalization is defined as an inpatient event that occurred on the same service date or (service date + 1 day) as an ED visit. Allowing for (service date + 1 day) accounts for ED visits that bridge midnight.
- **NOTE:** A qualifying COPD ED event which leads to a non-qualifying hospitalization (i.e. the patient is hospitalized, but the hospitalization does not qualify as COPD-related) is counted as a COPD-related ED event only. The hospitalization in that case is not counted as COPD-related and does not contribute to the episode length.

1B. MODERATE/SEVERE EXACERBATION EVENTS

Moderate/Severe Exacerbation Events are identified by the presence of **emergency department (ED) and/or ambulatory events.**

ED Events: An ED visit with the presence of an ICD-10 code identifying COPD exacerbations **AND** a dispensing of antibiotics or systemic corticosteroids +/- 5 days from the date of the ED visit was considered a moderate/severe exacerbation event.

- Systemic corticosteroids included both oral formulations filled in pharmacies as well as injectable medications administered by a physician.
- A maximum of one ED event per day is counted.
- A COPD-related qualifying ED event leading to a COPD-related hospitalization (occurring on or between the admission and discharge dates of a hospitalization) is included *only* as a COPD-related hospitalization (the ED event was not considered a separate event).
- In claims data, systemic corticosteroids/antibiotic evidence includes prescriptions filled.
- ED visits with qualifying COPD diagnosis codes but NO systemic corticosteroids/antibiotic prescriptions were not included as COPD exacerbations and did *not* contribute to the episode logic.
- Every ED visit meeting the qualification for proximity to systemic corticosteroids/antibiotic prescription/fill was counted as an event (in other words, a single prescription/fill can serve as the qualifying prescription for multiple ED events on different dates).

Ambulatory Events: An office visit or urgent care visit or outpatient hospital visit with the presence of an ICD-10 code for COPD **AND** a dispensing of antibiotics or systemic corticosteroids +/- 5 days from the date of the visit is considered a moderate/severe exacerbation event.

- Every Ambulatory visit meeting the qualification for proximity to an antibiotics or systemic corticosteroid prescription/fill is counted as an event (in other words, a single prescription/fill can serve as the qualifying prescription for multiple Ambulatory events on different dates).

- Surgery-related antibiotics excluded: For both the ED and Ambulatory events, antibiotic prescriptions within 1 week before/after surgical procedures were excluded. Thus, these antibiotic fills were not be used qualify an ED or ambulatory visit as a COPD exacerbation.

2. EXACERBATION EPISODE LOGIC

2A. BASE CASE: SINGLE EVENT EPISODE

In the base case, a COPD exacerbation begins with the earliest occurrence of one of these three events (qualifying COPD hospitalization event, COPD ED Event, COPD Ambulatory Event). In the single event episode, the exacerbation ends 14 days from start (or hospital discharge), whichever is later.

Episode Start date = service date of a qualifying event (IP, ED, outpatient).

- If the qualifying event is a COPD hospitalization, the start date is the admission date.
- The start date is set as the service date of the medical visit, even if an ATB or OCS prescription occurred prior to this visit date.

Episode End Date = Start date + 14 days (if qualifying event is ED or Ambulatory event)
= Discharge date + 14 days (if qualifying event is an inpatient event)
If the Episode End Date is on or after the sample ID date (end of claims observation period), then an additional flag is set, **CENSORED**=1

2B. COMPLEX CASE: MULTIPLE EVENT EPISODE

If another qualifying event occurs before the episode End Date, then the episode is extended. The date of the second event is used to reset the End Date per the definition above, using information from this second event. The process continues until a 14-day period with no evidence of exacerbation (qualifying event) occurs. The End Date is set (as defined above).

Episode Length: (End date – Start date) + 1

Episode Severity: If an exacerbation episode includes a qualifying COPD inpatient event, the episode severity is set to Severe. Otherwise, set to moderate/severe.

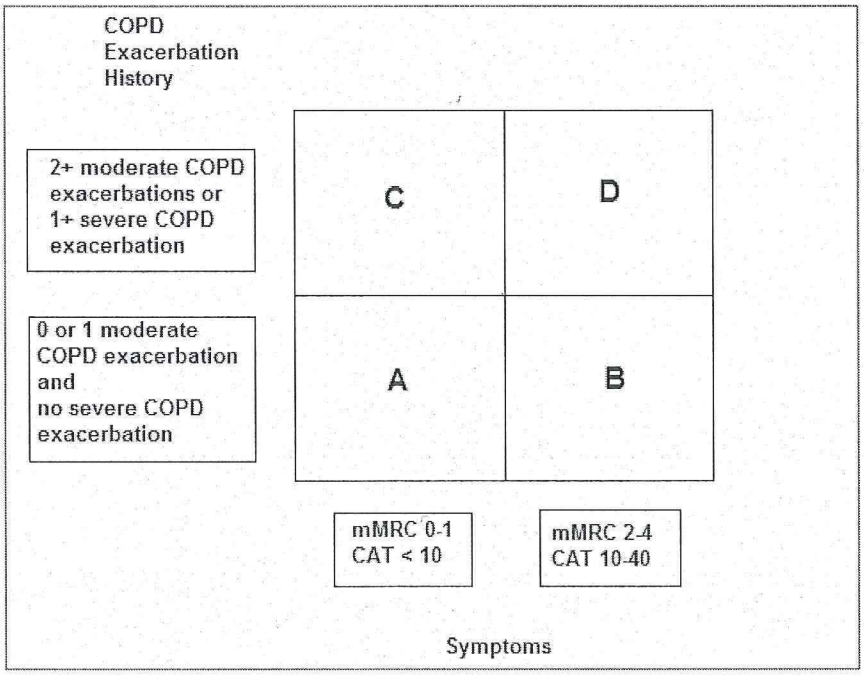
Episode Identification Strategy:

1. Identify all qualifying exacerbation events (Inpatient, ED, Ambulatory)
2. Search for the first exacerbation qualifying event
3. Set START DATE based on the first event using definition above
4. Set a provisional episode END DATE based on definitions above
5. If there is another exacerbation qualifying event between the Start Date and provisional End Date, reset the provisional END DATE based on the definition above
6. Repeat until provisional END DATE has no other qualifying events within 14 days
7. Set episode END DATE as defined above

10 APPENDIX 3. GOLD GRADE CLASSIFICATIONS

The proportion of patients in each GOLD category of exacerbation risk will be presented by treatment cohort. The calculation will use count of COPD exacerbations during the 12-month baseline combined with CAT total score and/or mMRC score to create mutually exclusive counts for each GOLD category (A, B, C, D). Since the counts of patients with CAT total score and mMRC score available may differ, the analysis may be conducted separately for patients with CAT score, mMRC score, and both PRO scores available.

Figure 8 – GOLD Grade Classification of Risk of COPD Exacerbation



11 APPENDIX 4. COVARIATES OF INTEREST FOR IPTW

The following is a list of covariates planned for inclusion in the IPTW analysis. Final covariates included in the multivariable models will be determined following review of the comparison of weighted and non-weighted descriptive results as described in Sections 4.3 - 4.4.

Clinical characteristics

- General health
- BMI
- Current smoking status
- Pack years
- Years since COPD diagnosis
- Baseline Quan-Charlson comorbidity score
- Count of COPD-related comorbid conditions
- Count of all-cause baseline medication fills
- Count of COPD-related baseline medication fills
- Count of all-cause baseline ambulatory visits (office and/or outpatient)
- Count of COPD-related baseline ambulatory visits (office and/or outpatient)
- Count of baseline COPD exacerbations

Demographic and sociodemographic characteristics

- Age
- Sex
- Race
- Ethnicity
- Geographic region
- Population density
- Marital status
- Household income
- Education

[i] Dillman DA. Mail and Internet surveys: the tailored design method, second edition. New York: John Wiley & Sons, Inc.

[ii] Edwards P, Roberts I, Clarke M. et al. Increasing response rates to postal questionnaires: systematic review. *BMJ.* 2002;324:1183

[iii] Following the Dillman tailored design method, Optum's standard data collection time period is 8 weeks.

[iv] Doherty, DE, Belfer, MH, Brunton, SA, Fromer, L, Morris, CM, Snader, TC. Chronic Obstructive Pulmonary Disease: Consensus Recommendations for Early Diagnosis and Treatment. *J Fam Prac*, 2006.

[v] Bestall C, Paul EA, Garrod R, et al. Usefulness of the Medical Research Council (MRC) dyspnea scale as a measure of disability in patients with COPD disease. *Thorax* 1999;54:581-586.

[vi] Jones PW, et al. Development and first validation of the COPD Assessment Test (CAT). *Eur Respir J* 2009;34(3):648-654.

[vii] Jones PE, et al. Properties of the COPD Assessment Test (CAT) in a cross-sectional European study. *Eur Respir J* 2012; doi: 10.1183/09021936.00177210.

[viii] Quan H, Li B, Couris CM, Fushimi K, Graham P, Hider P, Januel JM, Sundararajan V. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. *Am J Epidemiol.* 2011;173(6):676-82.

[ix] Bayliss EA, Ellis JL, Shoup JA, Zeng C, McQuillan DB, Steiner JF. Association of patient-centered outcomes with patient-reported and ICD-9-based morbidity measures. *Ann Fam Med.* 2012;10(2):126-33.

[x] HCUP Comorbidity Software. Healthcare Cost and Utilization Project (HCUP). 2009. Agency for Healthcare Research and Quality, Rockville, MD. www.hcup-us.ahrq.gov/toolssoftware/comorbidity/comorbidity.jsp.

[xi] Frytak JF, Henk JH, et al. Health Services Utilization Among Alzheimer's Patients: Evidence from Managed Care. *Alzheimer's and Dementia.* 2008;4(5):361-67.

[xii] *Ibid.*

[xiii] *Ibid.*

[xiv] Austin, PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research* 2011;46:399-424.

[xv] Mansournia MA and Altman DG. Inverse probability weighting. *BMJ* 2016;352:i189.

[xvi] Rosenbaum P and Rubin D. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70:41-55.

[xvii] Austin, PC and Stuart, EA. Moving towards best practices when using inverse probability of treatment weighting using the propensity score to estimate causal treatment effects in observational studies. *Statis Med* 2015;34:3661-3679

[xviii] Frölich, M. Finite-sample properties of propensity-score matching and weighting estimators. *The Review of Economics and Statistics* 2004;86:77-90.