Non-Interventional Study Protocol

Short title: Post-Marketing Safety Study of Roflumilast

Title: Long-Term Post-Marketing Observational Study of the Safety of Roflumilast

Study ID: D7120R00003

Sponsor: AstraZeneca Gothenburg
Pepparedsleden 1
431 83 Mölndal
Sweden
Phone: +46 31 7761000
Fax: +46 31 7763700

Study phase: Therapeutic Use

Date of Final Draft version of protocol: 7th February 2017

Protocol version no: 3.0

This protocol version includes Amendment Nos: 7

CO-ORDINATING STUDY MANAGER

[redacted]

AUTHOR(s)

[redacted]

APPROVERS

[redacted]

[redacted]

[redacted]

Lead Investigator

[redacted]

For signatures, see separate page.

Summary

Short Title of Study

Postmarketing Safety Study of Roflumilast

Study Data Sources

Multiple existing data sources in Germany, the Netherlands, UK, Denmark, Norway, Sweden, Finland and Canada.

Objectives

To evaluate the long term safety of roflumilast in the treatment of Chronic Obstructive Pulmonary Disease (COPD) with main focus on all-cause mortality.

Methodology

Based on suitable existing electronic databases covering large regions in Europe andCanada, roflumilast-exposed and matched controls (non-exposed cohorts of COPD patients) will be identified. For each region, outcomes in the exposed cohort will be compared with outcomes in their matched controls, and, subsequently, a cross-region analysis will be performed. The exposed cohorts will be identified yearly for up to 4 years and followed for at least 5 years. Yearly interim analyses will be performed starting 2 years after the last exposed patient has been identified, with first interim report being expected in 2nd quarter 2017.

Number of patients

It is estimated that 0.25% of the total population of a region will be exposed to roflumilast during the first 4 years after launch and full reimbursement. All COPD patients over 40 years of age with documented start of roflumilast exposure after granting of reimbursement in the region and after the number of roflumilast treated patients captured in that region is considered sufficient for valid results prior to the 1st quarter 2015 data cut will be included in the study and their data will be analysed.

Diagnosis and Main Criteria for Inclusion

To be included in the exposed cohort, a patient must be over 40 years of age and diagnosed with COPD, must have at least one prescription of roflumilast and must have active data in

the respective databases for at least 1 year prior to the date of the first roflumilast prescription.
The matched, non-exposed cohort is sampled to match the exposed cohort except for the roflumilast prescription, and will comprise up to 5 matched non-exposed controls per exposed patient.

Duration of Data Collection for each Patient

End of data collection is defined as 5 years after the last cohort is identified. The patients will be followed until end of data collection unless earlier censored due to death, or exit from the database, whichever occurs first.

Criteria for Evaluation

Main outcome variables

The primary outcome is 5 year all-cause mortality. Secondary outcomes are death by suicide or hospitalisation for suicide attempt, all-cause hospitalisation, major cardiovascular events, new diagnosis of malignant tumour, respiratory disease related hospitalisation, new diagnosis of depression, abnormal, unexplained weight loss, serious diarrhoea of non-infectious origin and new diagnosis of tuberculosis or hepatitis B or C.

Statistical Methods

Up to 5 matched controls will be identified for each exposed patient, based on sex, age category, years since diagnosis of COPD/chronic bronchitis (as available in the database), years of membership in database (as available) and propensity score and the two cohorts will be compared with regard to the primary and secondary outcome variables. The hazard ratios with 95% confidence intervals will be reported together with p-values. The analysis will take into account sex and age and other potential confounders. Matching will be accounted for by considering the exposed patient and his/her matched controls as one stratum. Sensitivity analyses will be done for pre-defined confounders that cannot be reliably measured, e.g. smoking.

Table of Contents

List of Tables

List of Abbreviations and Definition of Terms

ATC:	Anatomical Therapeutic Chemical Classification System
COPD:	Chronic Obstructive Pulmonary Disease
EMA:	European Medicines Agency
GCP:	Good Clinical Practice
GePaRD:	German Pharmacoepidemiological Research Database
GOLD:	Global Initiative for Chronic Obstructive Lung Disease
GP:	General Practitioner
GPP:	Good Pharmacoepidemiology Practices
ICD-9-CM	International Classification of Diseases, 9th Revision, Clinical Modification
ICD-10-GM	International Classification of Diseases, 10th Revision, German Modification
ICH:	International Conference on Harmonisation
ICS:	Inhaled corticosteroids
IEC:	Independent Ethics Committee
IRB:	Institutional Review Board
MI:	Myocardial Infarction
SAP:	Statistical Analysis Plan

1 Introduction

Chronic obstructive pulmonary disease (COPD) is a leading cause of death. Takeda A/S received central marketing authorization in the EU in July 2010 and regulatory approval in Canada in November 2010 for its selective phosphodiesterase-4 (PDE4) inhibitor roflumilast (DAXAS®) (1,2). Where launched, roflumilast is available as 500 µg tablets for maintenance treatment of severe COPD associated with chronic bronchitis in adult patients with a history of frequent exacerbations as add on to bronchodilator treatment. The recommended dose is one 500 µg tablet once daily. Roflumilast is contraindicated in patients with hypersensitivity to the tablet ingredients and in patients with moderate to severe liver impairment (Child Pugh B or C).

Roflumilast was launched in the first countries (Germany and Denmark) in September 2010. A total of 2,874 COPD patients have been treated with roflumilast in clinical trials of 12 months duration. Clinical trials of longer duration have not been conducted. Since roflumilast will be used as a permanent maintenance treatment, a long-term safety assessment exceeding 12 months was requested as a condition of approval for marketing in the EU. During the approval process, Takeda as the Marketing Authorisation Holder (MAH) for roflumilast therefore committed to the European Medicines Agency (EMA) to perform a database study, and proposed that the study be conducted in large, unselected COPD populations, reflecting the use of roflumilast in a real-life setting. Established electronic health care databases in countries where roflumilast is on the market and data on a meaningful number of roflumilast treated patients is captured will be the source on which to base this Post-Authorisation Safety Study (PASS). Databases in the following countries have been identified: Germany, UK, the Netherlands, Sweden, Denmark, Finland, Norway and Canada.

2 Study Objective(s)

The objective of this study is to evaluate the long-term safety of roflumilast in the treatment of COPD, with focus primarily on all-cause mortality. In addition, the study will evaluate potential risks, including potential safety issues identified during the clinical trials of roflumilast. Specifically this study aims to compare the incidences of all-cause mortality, death by suicide or hospitalisation for suicide attempt, all-cause hospitalisation, major cardiovascular events, new diagnosis of malignant tumour, respiratory disease related hospitalisation, new diagnosis of depression, abnormal, unexplained weight loss, serious

diarrhoea of non-infectious origin and new diagnosis of tuberculosis or hepatitis B or C in roflumilast treated (exposed) COPD patients compared with matched COPD patients not treated with roflumilast (non-exposed).

The exposed and non-exposed cohorts will be followed for five to nine years.

3 Study Administrative Structure

3.1 Conduct

The study will be conducted by independent investigators qualified in epidemiology and will be sponsored by AstraZeneca. The patient data will remain the properties of the respective database owners, and AstraZeneca will not have direct access to the data but will be co-owner of the results derived from these data, together with the respective investigators. The Lead Investigator will lead the team of investigators, each of whom is responsible for the conduct of the study in databases covering a given region, with which he/she has experience, and to which he/she can get data access. The region is a country or a geographical area within a country, where the database or databases to be employed cover the entire population or a representative sample of the population. The team of investigators will cooperate to align procedures (e.g. coding) and methods of analysis to ensure comparable analyses in spite of non-uniform datasets. If datasets from two or more Nordic countries are compatible, and pooling is feasible, the datasets from these countries may be pooled and the analysis repeated on the pooled data.

3.2 Scientific Advisory Committee

An Independent Scientific Advisory Committee, chaired by the Lead Investigator has been identified to oversee the scientific integrity of the study and review the statistical analysis plan (SAP) and results of the interim and final analyses.

3.3 Sponsor Personnel

From AstraZeneca the study will be coordinated by the Coordinating Study Manager. In addition, at least the following AstraZeneca personnel will support the study: a clinical lead for roflumilast, a patient safety physician for roflumilast, an epidemiologist for roflumilast and a statistician.

4 Study Data Sources

Electronic health care databases holding demographic data and data on health and prescribing or dispensing of medications from multiple countries in the EU and from the Saskatchewan province in Canada will be used in this COPD study. A region will be included if reimbursement has been granted in the region and the number of roflumilast treated patients captured in that region is considered sufficient for valid results.

The following sources have been selected as relevant for the study. The databases are different in structure and content, and for each region either a single database with broader input from one source, or a combination of databases will be used.

Germany

The German Pharmacoepidemiological Research Database (GePaRD) built by the Bremen Leibniz Institute for Prevention Research and Epdemiology -BIPS, is based on medical claims data from four German statutory health insurance providers (3). It includes data of more than 17 million insured from all regions in Germany. Preliminary analyses regarding age and sex distribution, the number of hospital admissions and drug use have shown that the database is representative for Germany and that the insurance population is rather stable over time. The GePaRD contains for each insured demographic information, information on hospitalizations (diagnoses and procedures), outpatient visits (diagnoses and procedures) and reimbursed out-patient prescriptions. All diagnoses are coded according to the German modification of the International Classification of Diseases (ICD-10-GM).

United Kingdom

The General Practice Research Database (GPRD), renamed the Clinical Practice Research Datalink (CPRD), in the United Kingdom, is the world's largest computerised database of anonymised longitudinal clinical records from primary care (4,5). The anonymised records are collected from a voluntary group of approximately 600 General Practitioners throughout the UK, and the database currently covers a representative population of around 4 million inhabitants.

The Netherlands

The PHARMO medical record linkage system, which is a population-based patient-centric data tracking system that includes patient demographics, mortality, drug dispensings (filled prescriptions and OTC products dispensed at the pharmacy), hospital morbidity, clinical laboratory, pathology and general practitioner (GP) information of 3.2 million community-dwelling inhabitants of 65 municipal areas in the Netherlands (6). The PHARMO databases are linked to form a network on a patient level. The General Practice Register currently comprises a representative sample of around 2 million inhabitants.

Sweden

Centre for Epidemiology, Swedish National Board of Health and Welfare, allowing linkage of the Swedish Prescribed Drug Register (filled prescriptions) with the Swedish Hospital Discharge Registry, the Swedish Death Registry, the Swedish Cancer Registry and registries holding socio-demographic data (7). These national databases hold data on all inhabitants in Sweden and thus cover the complete population of 9.4 million inhabitants.

Denmark

Statistics Denmark, allowing linkage of the community pharmacy prescription database (filled prescriptions), the Hospital Discharge Registry, the Death Registry, the Danish Cancer Registry and registries on socio-demographic data (7). As for Sweden, these databases cover all inhabitants in Denmark, a population of 5.5 million.

Finland

Kela Institute, Helsinki, allowing linkage of the Finnish Drug Register (filled, reimbursed prescriptions) with the National Registry of Hospital Care (HILMO registry, THL), the Finnish Causes of Death Registry, the Finnish Cancer Registry and registries holding socio-demographic data (7). The national databases have full coverage of the 5.3 million inhabitants in Finland.

Norway

Norwegian Institute of Public Health, allowing linkage of the Norwegian Prescription Database (filled prescriptions), the Hospital Discharge Registry, the Norwegian Causes of Death Registry, the Norwegian Cancer Registry and registries holding socio-demographic

data (7).The Hospital Registry is relatively new in Norway, but since only one year of historic data are required for this study, the study is feasible in Norway. The databases cover the complete population of 4.9 million inhabitants.

Canada

The Saskatchewan databases, which are based on data from a publicly funded health care system covering all inhabitants in the Saskatchewan province, i.e. about 1 million inhabitants, or 3.2% of the total population of Canada. The Saskatchewan databases have been widely used for pharmacoepidemiological research, including studies in COPD (8,9). The databases comprise a population registry, a database with prescription drug data, a hospital service database, a database with medical services data based on physicians claims, a cancer registry and a database with vital statistics.

5 Study Design and Plan

This study is a 'non-interventional study' as defined for drug studies in Directive 2001/20/EC (10) and will follow the guidelines for Good Pharmacoepidemiology Practices (GPP) (11). The data to be analysed have all been collected for other purposes.

An inception cohort study design, comparing roflumilast exposed COPD patients over 40 years of age with non-exposed, matched COPD controls will be used for the long-term safety surveillance of roflumilast, using several electronic databases in the EU and Canada. Analysis will be conducted and results presented separately for each region, and, subsequently, a cross-region analysis will be performed. Each exposed COPD patient, together with matched, unexposed COPD patients will be followed for at least 5 years or until death or discontinuation (loss to follow-up) from the database, whichever occurs first.
The primary outcome of interest is 5 year all-cause mortality. Secondary outcomes are: death by suicide or hospitalisation due to suicide attempt, all-cause hospitalisation, major cardiovascular events, hospitalisation related to respiratory disease, new diagnosis of cancer, depression, tuberculosis and hepatitis B or C. Cause of death, if available, will be reported as listed in the databases, categorised as cardiovascular, respiratory, cancer, suicide and other. For further details, see section 6.2.

COPD patients will be identified by a hierarchical approach, based on data availability, as follows: a) inpatient or outpatient diagnostic codes for either COPD or chronic bronchitis (in

medical records databases and claims databases with outpatient diagnostic information), b) a hospital diagnosis of COPD or chronic bronchitis, or c) chronic treatment with inhaled β_2-agonists and anticholinergic medications (in databases with no outpatient diagnostic information available and no hospital diagnosis of COPD or chronic bronchitis recorded). For further details, see section 7.1.

Matching between COPD patients exposed to roflumilast and controls will be based on age, sex, years since diagnosis of COPD/chronic bronchitis, as available in the database, years of membership in database (as available) and a propensity score (12) derived from a list of relevant variables available in the selected databases. If available with the planned matching accuracy (calliper width) (13) in the respective databases, up to 5 matched controls will be identified for each exposed patient and the data will be analysed as parallel groups, using statistical methods that take into account the matching of observations.

Once a year, for a maximum of 4 years, cohorts of patients with first exposure to roflumilast during the past year will be identified together with their matched controls. The total cohort of exposed patients and matched controls will be followed for 5 years after the last cohort is identified. End of data collection is defined as the time when the last non-censored patient in the last cohort has completed the 5 year follow-up period, therefore the first cohort identified will have up to 9 years of follow-up. See section 7.3.
Descriptive analyses of the assembled cohorts will be performed yearly and reported to the EMA. Interim analyses including primary and secondary outcomes will be performed and reported the first time 2 years after the last cohort has been assembled and yearly thereafter.

5.1 Study Schedule

Planned Start of data collection:	Date of market availability with reimbursement in region
Planned assembling of first cohort:	3rd quarter 2013
Planned completion of cohort assembling:	1st quarter 2015
First Interim Report:	31 October 2017
Last Interim Report:	30 June 2019
Planned extraction of last dataset:	31 January 2020
Planned completion of data analysis:	30 September 2020

Planned finalisation of the Study Report: 31 March 2021

The Start of Study is defined as the date of signing the agreement between Takeda and the Lead Investigator. The End of Study Follow-Up is defined as the time when the last non-censored patient in the established cohorts has completed the 5 year follow-up period. Extraction of the last dataset may be postponed if the necessary data are not yet available in the database(s) in a region at the planned time point, in which case the final study report may be delayed accordingly.

AstraZeneca will ensure that End-of-Study notification is submitted to the EMA. Each investigator will inform the relevant IEC/IRB of End-of-Study.

Based on upcoming knowledge and in agreement with the EMA, AstraZeneca may choose to terminate the study prematurely. In such case, the investigators and the Independent Scientific Advisory Committee will be informed promptly. Details of the study will be available on the European Network of Centres for Pharmacoepidemiology (ENCePP) E-register of studies.

6 Outcomes

6.1 Primary Outcome

- All-cause 5 year mortality.

6.2 Secondary Outcomes

Codes in parenthesis are International Classification of Diseases, 10th Revision (ICD10) codes.

- Death by suicide or hospitalisation for suicide attempt (Intentional Self-Harm X60-X84 or overdose, typically T39-T43)
- Hospitalisation for any cause
- Selected cardiovascular events leading to hospitalisation
    - Arrhythmia (conduction disorders and dysrhythmias) (I44, I45)
    - Myocardial infarction (MI) (I21, I22)
    - Cerebral infarction (I63) or stroke not specified as haemorrhage or infarction (I64)

- Pulmonary embolism (I26)

- Respiratory disease related hospitalisation, including hospitalisation due to COPD/bronchitis exacerbation (J00-J99)
- New diagnosis of depression (with or without hospitalisation) (F32.2-F32.3, F33.2-F33.3 (or F32 and F33 if next level code is not available), or prescription of antidepressant medication (ATC N06A)
- New diagnosis of malignant neoplasm (C00-97)
- Hospitalisation due to diarrhoea of non-infectious origin (K52.9 Unspecified, K59.1 Functional diarrhoea)
- Abnormal, unexplained weight loss (R63.4) with no new diagnosis of malignant neoplasm within 4 months after abnormal weight loss diagnosis
- New diagnosis of tuberculosis (A15) or new diagnosis of Hepatitis B or C, or other severe viral hepatitis infection (except hepatitis A) (B16-B19)

6.3 Exploratory outcomes

Descriptive analysis of these outcomes will be performed on the final dataset and reported in the final report.

Underlying cause of death, as available will be categorised as follows:

Cardiovascular related death (one of the following):

- Sudden cardiac death (I46.1)
- Death related to ischemic heart disease (I20-25)
- Death related to heart failure (I50)
- Death related to cardiomyopathy or dysrhythmia (I42 and I44/I45)
- Death related to pulmonary heart disease or diseases of pulmonary circulation (I26-28)
- Cerebrovascular disease (I60-69)

Respiratory disease related death (one of the following):

- COPD (J44)
- Lung cancer (C34)
- Influenza and/or pneumonia (J09-18)
- Other acute lower respiratory infections (J20-22)
- Tuberculosis (A15)
- Acute respiratory distress syndrome (J80)

- Pulmonary oedema (J81)
- Other interstitial pulmonary disease (J84)

Cancer related death (other than lung cancer) (C00-97 except C34 and C44)

Other cause of death

7 Selection of Study Population

The identification of patients for the cohorts to be studied will start when reimbursement has been instituted in the respective region, and a substantial number of patients are assumed to have been exposed to roflumilast, based on roflumilast sales figures from the region.

7.1 Cohort Selection

An Exposed Cohort and a Matched Non-exposed Cohort will be created for comparison of the safety outcomes. All patients with at least one prescription for roflumilast will be as assigned to the Exposed Cohort.

Exposed Cohort selection

Inclusion Criteria:

- Have one or more prescription(s) of roflumilast with the date of first prescription/dispensation defined as the Cohort Entry Date
- Have active data in the respective database(s) for at least 1 year prior to the Cohort Entry Date
- Be at least 40 years old at Cohort Entry Date
- Have a diagnosis of COPD or chronic bronchitis prior to the Cohort Entry Date (databases with outpatient diagnoses) or (databases without outpatient diagnoses) either a hospital diagnosis of COPD or chronic bronchitis prior to the Cohort Entry Date or a proxy for such indication based on prescription data.

Non-Exposed Cohort selection

Patients meeting the following inclusion criteria belong in the Non-Exposed Cohort and are eligible as Matched Non-Exposed controls. Each patient in the Non-exposed Cohort selected as matched exposed patient according to section 11.1.2. is assigned as a Cohort Entry Date the date of the first roflumilast prescription of the matched exposed patient.

Inclusion Criteria:

- Have no prescriptions of roflumilast
- Have active data in the respective database(s) for at least 1 year prior to the Cohort Entry Date
- Be at least 40 years old at Cohort Entry Date
- Have a diagnosis of COPD or chronic bronchitis prior to the Cohort Entry Date (databases with outpatient diagnoses) or (databases without outpatient diagnoses) either a hospital diagnosis of COPD or chronic bronchitis prior to Cohort Entry Date or a proxy for such indication based on prescription data

The following ICD-10 and Anatomic Therapeutic Chemical (ATC) classification system codes will be used for the identification of patients for the cohorts. For regions not using ICD-10 codes, the matching codes employed (e.g. ICD-9, ICPC, READ codes) will be described in the detailed SAP to be used for that region:

Relevant ICD-10 codes:
COPD: J44
Chronic Bronchitis: J41-42
Asthma: J45
COPD exacerbations: J44 in combination with hospitalisation or systemic corticosteroid/antibiotic prescription

Relevant ATC codes:
Roflumilast: R03DX07
B_2 stimulators (sympatomimetics): R03AC
Anticholinergics (tiotropium R03BB04, ipratropium R03BB01): R03BB
Inhaled corticosteroids: R03BA
Combination products for inhalation: R03AK
Systemic corticosteroids: H02AB
Systemic antibiotics: J01

Theophyllin: R03DA04
Oxygen: V03AN01

Exposed COPD patients in a region will be identified on a yearly basis (in 1st quarter each year), starting when data on at least 2000 exposed patients in a database are expected to be available. The expectation of data availability will be based on roflumilast sales figures for the area covered by the database(s), providing an estimate on the number of patients treated since launch/granting of reimbursement, as well as the population coverage of the database(s) and the known lag-time for data to be uploaded to the database(s). If this first evaluation based on sales indicates that 2500 patients exposed for the first time after implementation of reimbursement may be available, the region will join the study and data in the region will be accessed and the number of patients fulfilling the inclusion criteria for the exposed cohort identified. If the extracted number is 2000 or higher in the database, all exposed and non-exposed COPD patients will be identified, based on the inclusion criteria for these cohorts. If the first extracted cohort of exposed patients comprises less than 2000 patients, further data extraction will be postponed until the following year, and the next extraction will comprise both time intervals.

7.2 Controls identified by Propensity Score Matching

By matching on sex, age category and propensity score, years since diagnosis of COPD/chronic bronchitis (as available in database), years of membership in database (as available), 5 non-exposed controls are identified (if possible) for each exposed patient.

7.3 Longitudinal Follow-up

Each cohort will be longitudinally followed up on a yearly basis, with descriptive analyses of cohort characteristics, starting for each region 1 year after the first cohort identification in that region. For the primary analysis, each patient included will be followed up for a minimum of 5 years to a maximum of 9 years, depending on cohort entry date, until the first of: date of death or transfer out of database.

8 Classification of Exposure

As the first prescription of roflumilast will define entry to the exposed cohort, the effect of cumulative exposure to roflumilast with respect to primary and secondary outcomes will be assessed.

The total amount of drug dispensed can be defined as the number of defined daily doses (DDDs) or as actual amount dispensed. DDD provides a direct measure of how long the drug would last if taken as one DDD per day. Based on the dispensing date and the DDD information or prescribed dosage each prescription will be translated into a drug use period. The drug use period starts at the dispensing date or if there is ongoing exposure the start is shifted to the end of the ongoing exposure period. The shift is limited to a maximum of 1 month. The length of the prescription is given by the dosage information in the prescription or is defined based on the DDD. An additional treatment gap period (e.g. 50%), as suitable (14,15) is added to avoid unnecessary breaks between prescriptions and account for late starting dates and compliance. With this algorithm the follow-up time of each subject is divided into half closed time intervals $(t_1,t_2]$ where the left boundary t_1 is not included and the right boundary t_2 is included in the interval. Within each such time interval the exposure status (i.e. exposed or not exposed) of a given subject to a particular drug is constant. The resulting consecutive drug use periods are used to define time-dependent current exposure periods indicating treatment status at any time during the follow-up. The algorithm for the calculation of the exposure will be further detailed in the statistical analysis plan.
Based on dispensing of roflumilast and subsequent consecutive drug use periods the following time dependent variables for roflumilast exposure will be created:

Ever vs. Never:
Taking the value of ever as soon as one prescription of roflumilast has been dispensed during the follow-up period. Non-exposed patients will be categorised as never exposed.

Current exposure:
Taking the value of current to indicate current roflumilast exposure (i.e., drug use during a specific time period/at a specific time point). Specific timepoint(s) will be described further in the SAP.

Previous exposure:
Taking the value of previous to indicate periods of non-use of roflumilast if there were episodes of use previous to the current exposure timepoint.

Duration of exposure:

Time-dependent cumulative sum of duration of the previous roflumilast exposure periods at any given time (i.e. cumulative time on roflumilast). Possibly overlapping periods are only calculated ones. The duration will be categorized into four cumulative exposure classes using data specific quartiles (no exposure, 1st quartile, 2nd quartile, 3rd quartile)

Cumulative dose:
Time-dependent cumulative sum of the DDDs of roflumilast prescriptions since entry into the cohort up to the end of each period. The cumulative dose will be categorized into four cumulative exposure classes using data specific quartiles (no exposure, 1st quartile, 2nd quartile, 3rd quartile).

Time since last dose: Time to last current roflumilast exposure since entry into the cohort, classified using yearly categories (under 1 year, 1-2 years, 2-3 years, 3-4 years, 4-5 years, over 5 years) .

9 Conduct

9.1 Data collection overview

The data extracts aim to

1) collect data on COPD patients prior to and on Cohort Entry date for description of the populations and to enable the best possible propensity score matching in order to control potential confounding and especially confounding by indication to the extent possible with the available data
2) collect outcome data after the Cohort Entry Date to identify any differences in selected outcomes between the Exposed Cohort and the Non-Exposed (comparator) Cohort

The selected data sources can be categorised into regions where the databases are based entirely or partly on entries from General Practitioners (GPs) and regions with no input from GPs. For the first category, data analyses can rely on diagnosis codes for identification and sorting of co-morbidities and outcomes, whereas for the second category, such diagnostic codes will be available only if the patient has been in contact with the hospital (emergency unit, ambulatory clinic or hospital ward). In the second category, analyses will have to rely on

available diagnostic codes supplemented with proxies derived from prescription data and other sources, as available.

Table 1 below gives an overview of the data to be analysed, and the source from which they will be drawn in each region. The countries planned to be analysed in the Nordic region have similar data sources available and are presented together in one column. It may be possible to pool the data for a common analysis of these four populations.

Table 1: Type of data to be analysed and source in each region

Type of Data	Sweden, Denmark, Norway, Finland	Germany GePaRD	Netherlands Pharmo Record Linkage System	UK GPRD	Canada Saskatchewan databases
Demographics Age, gender	Sociodemographic databases	Claims data	GP, pharmacy, hospital data	GP data	Population registry
Migration or other loss from Data Source	Sociodemographic databases	Claims data	GP, pharmacy, hospital data	GP data	Population registry
Socioeconomics e.g. Education level, marital stage, employment situation, practise ID	Socioeconomic databases and socio-demographic DB, practice ID.	Claims data: Practise ID is available, as well as some information on SES	May not be available, except for practise ID	Practise ID, Index of Multiple Deprivation Score	Medical Services DB, physician and practise ID, residence information
Functional stage (capabilities for living a normal daily life)	Not available, but may be possible to find proxy in socioeconomic DB (e.g. employment situation, protected living etc.)	Not available but may be possible to find proxies	Not available, but may be possible to find proxies	To be checked if reliably available in GP data	Not reliably available, but proxy may be identified, e.g. Prescription Drug Data, incl. designation of special status
Smoking	Not available	Not available	Available, but data completeness to be checked for COPD patients	Available, but data completeness to be checked for COPD patients	Not available
COPD or chronic bronchitis	ICD10 (or ICPC) code in hospital DB and/or ATC codes in prescription DB	Claims data: ICD-10-GM code	ICD-9 codes in hospital data or ICPC codes or free text in GP data	READ codes in GP data	ICD-9 codes in Medical or Hospital Services DB
COPD severity	Ascertained by proxy information e.g. incidence of hospitalisations related to COPD,	Claims data: Ascertained by proxy information , e.g. Incidence	Available if spirometry done. If not, ascertained by proxy	Available if spirometry done. If not done, ascertained	Ascertained by proxy information.

	bronchodilator, ICS, oxygen dispensing, vaccination for pneumococci	of hospitalisations related to COPD, bronchodilator, ICS/oxygen dispensing etc., vaccination pneumococci	information	by proxy information	
History of hospitalisation for COPD exacerbation	ICD10 code in hospital DB, discharge diagnosis	ICD-10-GM code in hospital claims data	ICD-9 codes in hospital DB	READ codes in GP data	ICD-9 codes in Hospital Services Data

Table continued on next page

Table 1: Type of data to be analysed and source in each region

Type of Data	**Sweden, Denmark, Norway, Finland**	**Germany GePaRD**	**Netherlands Pharmo Record Linkage System**	**UK GPRD**	**Canada Saskatchewan databases**
Co-morbidities for Charlson Index	ICD10 codes in hospital DB and/or ATC codes in prescription DB	ICD-10-GM codes in outpatient and hospital claims data	ICD-9 codes in hospital data, ICPC codes and free text in GP data and National Medical Register	READ codes in GP data	ICD-9 codes in Hospital and Medical Services DB
Asthma	Not reliably available	ICD-10-GM codes in outpatient and inpatient claims data	ICD-9 codes in hospital data, ICPC codes and free text in GP data	READ codes in GP data	ICD-9 codes in Hospital or Medical Services DB
Vaccination against pneumococci	Prescription data	Outpatient claims data (Physician reimbursement codes)	Prescription data	Prescription data	Prescription data
Alcohol or drug abuse	Not reliably available	Alcohol abuse not reliably available (proxy information); drug abuse as ICD-10-GM codes	Not reliably available	READ codes in GP data	Not reliably available
Death	1) Hospital DB and 2) Death registry	Claims data	Death registry	GP data	Vital statistics
Cause of death	1)Hospital DB and 2)Death registry	Not available (except for patients dying in hospital). Use of algorithms	Not available	GP data linkage	Vital statistics

		in dataset will be explored			
Primary and secondary discharge diagnosis (Major cardiovascular events, respiratory disease, severe immunological disease, malignancies)	ICD10 in Hospital DB	ICD-10-GM in hospital claims data	ICD-9 in Dutch National Medical Register	GP data	Hospital services data
Abnormal, unexplained weight loss	ICD 10 in Hospital DB	ICD-10-GM in hospital claims data	ICD-9 in National Medical Register	GP data	Hospital services data
Serious diarrhoea of non-infectious origin	ICD 10 in Hospital DB	ICD-10-GM in hospital claims data	ICD-9 in National Medical Register	GP data	Hospital services data
Severe infections	ICD 10 in Hospital DB, Prescription DB	ICD-10-GM in in- and outpatient claims data	ICD-9 in National Medical Register	GP data	Hospital and Medical Services DB
Mood Disorder	ICD 10 in Hospital DB or based on Prescription DB	ICD-10-GM in in- and outpatient claims data	ICPC codes and free text in GP data	GP data	Hospital and Medical Services DB
Suicide	Death registry	ICD-10-GM in in- and outpatient claims data	ICD9 in hospital data	GP data	Vital statistics
Suicide attempt	Hospital DB	ICD-10-GM in in- and outpatient claims data	ICD9 in National Medical Register	GP data	Hospital and Medical Services data

10 Data Management

For each (set of) database(s) a SAP, tailored to the data available in the region, will be developed and approved by the relevant Investigator as well as the Scientific Advisory Committee and AstraZeneca before the first data extraction. The local SAP will be based on the analysis outlined in this protocol and will, among other things, define in detail the definitions used to identify the variables to be extracted, including diagnostic codes, proxies etc. to be used, and the hierarchy in which they will be employed.

The SAP must be amended if necessary before each additional data extraction, and the amendment approved following the same process as for the initial SAP. The first full data extract from each (set of) database(s) will be preceded by a selected extract aiming only at getting a status on the number of exposed patients for whom data are available.

10.1 Data Collection Tools and Flow

For each region, an Investigator is responsible for the conduct of the study based on the data accessible in the sources available to him/her. The Investigator appoints and trains a team of specialists who perform the data extraction and analysis according to this protocol and the local SAP.

The results of the yearly extracts will be provided to the Lead Investigator who is responsible for the analysis across regions for which a separate SAP must be developed and approved prior to the conduct of the first cross-region analysis. The Scientific Advisory Committee will review the SAP by region to ensure comparability between protocols.

10.2 Data Quality Control

Several measures will be employed to ensure the acceptable quality of the data extracts and analyses:

- All data extracts and analyses must be fully documented and logged (reproducible).
- The output must be produced by programs without further manipulation (manual handling) of data, as fully edited tables, generated by the program, for reproducibility and to avoid transcription errors.
- Each cohort must be independently extracted by a second, qualified researcher and the identity of the extracted cohorts be confirmed by comparison of patient numbers, gender distribution and mean age.
- Data analyses must be repeated by a second qualified researcher and results match, or an independent, qualified researcher must review all programs used for data analysis.
- All programs must be documented with version control and made available for audits and inspections.

- All substantial changes in programming between repeated analyses for the interim time points must be disclosed to the Lead Investigator and the AstraZeneca statistician without delay.

11 Statistical Methods and Determination of Sample Size

11.1 Statistical Analysis Plan

This study is an observational cohort study to evaluate the effect of roflumilast exposure on the risk of all cause mortality within five years after first exposure.
The study consists of four elements: a) Extraction of exposed cohort, b) identification and extraction of non-exposed cohort (see section 7.1), c) matching between exposed and non-exposed and d) descriptive analysis of cohorts and analysis of outcomes.

The statistical analyses described here will be detailed in a SAP for each region accounting for the data availability in that region. These plans will describe in detail how the data extraction, cohort identification and analyses planned in this protocol will be adapted to that region. SAP amendments will be created as needed between interim and final analyses.

11.1.1 Analysis Dataset

The main study analysis set is identified by all roflumilast exposed patients and the set of matched, non-exposed patients.. The selection of the study population is outlined in section 7.1.

When the criteria for inclusion of a region in the study described in section 4 have been met, the cohort of exposed and matched non-exposed patient cohort will be extracted on a yearly basis. The propensity score used for the matching of patients will be re-estimated each year based on the current data in order to account for possible changes in prescription patterns. The accumulated cohorts will be used for the statistical analysis.

11.1.2 Matching

The matching between exposed and non-exposed are performed by sex, age category (5 year spans above 40 years), years since diagnosis of COPD/chronic bronchitis, as available in the database, years of membership in database (as available) and propensity score.

11.1.2.1 Propensity Score

A tailored propensity score will be developed for each region. The propensity score will model the likelihood among COPD patients of receiving roflumilast.

The model for the propensity score is specific for each region, adapting to the availability of data to establish the best possible model. The model will to the extent possible include the following elements and may included additional elements identified in a region as having an impact on the decision to prescribe roflumilast:

1. Severity of COPD defined by Global Initiative for Chronic Obstructive Lung Diseases (GOLD) stage (16) (if spirometry data are available)
2. Severity based on the pattern of maintenance therapy, in particular long-term co-prescription of oxygen
3. Number of hospitalisations within the last year due to COPD exacerbation
4. Number of episodes within the last year of non-chronic prescription of systemic corticosteroid and/or systemic antibiotics
5. Charlson Comorbidity Index (17)
6. Number of hospitalisations within the last year for any cause
7. Asthma
8. Smoking status (if available)
9. Current use of immunosuppressive medication
10. Current use of theophyllin
11. Previous mood disorder defined by
    a. Diagnostic code from hospital or GP data, or
    b. Prescription of antidepressant medication
12. Functional stage (as recorded in database or by proxy, e.g. part of the working force, independent/assisted living, to be defined for each region)
13. Socioeconomic status (e.g. income category, educational level, single living, to be defined for each region)
14. Ongoing physical rehabilitation

The final details will be defined in the SAP for each region.

A logistic regression analysis will be performed on the identified COPD patients (exposed and eligible non-exposed controls) where treatment with roflumilast is the dependent,

dichotomous variable and the independent variables are listed above. Based on this analysis, the propensity score is calculated for each patient.

Smoking may be an important confounder in COPD but smoking status is usually not available in the databases at all. In regions with input from General Practice (PHARMO Linkage data and CPRD), smoking data are generally incomplete (4), but it may be more complete than on average for patients diagnosed with COPD. This assumption will be checked in the datasets from these regions. If smoking status is confirmed to be available for a high percentage (e.g. 90%) of the patients in the extracted cohorts, smoking status will be included in the propensity score, setting the missing smoking status at Cohort Entry date to non-smoking for the patients with missing status. If smoking status is missing for a higher percentage of the COPD patients in a region with GP data input, the data will be analysed with and without smoking status included in the model and the magnitude of the confounding effect assessed (sensitivity analysis).

11.1.2.2 Method of Matching

For each exposed patient up to 5 control patients will be selected who are of the same sex, within the same age group and within the same matching interval of the propensity score.

The distribution of propensity scores will be examined between exposed and eligible non-exposed controls. For propensity score matching to work well there must be a sufficient overlap in distribution of the score between the two groups. The aim is to have 5 controls for each exposed patient. The controls will be selected according to the procedure described below, based on sex, age and propensity score. A patient can only be selected as non-exposed control once in this study.

Criterion for propensity score matching in a region at time of data extraction: For at least 95% of the exposed there must be at least 1 non-exposed available that can be selected as control. If this is not the case the length of the matching interval (Δ) will be modified to achieve this percentage.
Exposed patients without matched non-exposed patients (up to 5%) will be described separately and excluded from further analysis.

If the above overlap criterion is fulfilled, propensity score matching will be used. Matching will be stratified by sex and age-group. Age will be split in 5-year age spans: 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80 and over (all years included). From the propensity model, the logit to the propensity score will be used. The length of the matching interval (Δ) will be derived as 0.2 times the standard deviation of the logit to the propensity score (Δ will be modified if necessary) (13).

With a maximum of 5, as many non-exposed patients meeting the matching criteria as possible will be selected for each exposed patient, and the exposed patient and his/her matched controls will be considered one stratum. If more than 5 matched non-exposed patients (controls) are available for an exposed patient, then controls will be sorted by a) (ascending) by absolute difference of actual age, and b) (ascending) by absolute difference in actual propensity score, and the controls on the top of this list will be selected.

If the criterion cannot be fulfilled in a region for a data extraction session, the selection of exposed patients and matched non-exposed patients (controls) will be adapted depending on the distribution of the propensity score and available data with the aim to ensure an unbiased comparison between exposed and non-exposed controls.

11.1.3 Statistical Analysis of Outcomes within each Region

Table 2 gives an overview of the variables to be included in the model per outcome category. Before analysis, the variables should be reassessed for (multi)co-linearity, and the number of variables adjusted to the number of events.

Table 2: Overview of the variables to be included in the model per outcome category

Outcome Variable to include in model	Death	All-cause hospitalis-ation	Selected major CV event	Respiratory related hospitalis-ation	New cancer	New tuberculosis or hepatitis B or C	Depression (new diagnosis)	Suicide or suicide attempt
Known coronary heart disease (other than prior MI)	X	X	X					
Hypertension	X	X	X					

Pneumococci vaccination	X	X	X	X	X	X		
Smoking status (if available in region)	X	X	X	X	X	X		
Prior mood disorder	X	X	X	X	X		X	X
Schizophrenia	X							X
Living alone	X	X	X	X			X	X
Alcohol or drug abuse	X	X	X	X	X	X	X	X
Prior intentional self harm								X

11.1.3.1 Analysis of the Primary Outcome

The primary outcome is the 5-year all cause mortality. This will be evaluated by a Cox proportional hazards regression model for counting processes, which allow the follow-up time to be divided into several periods and therefore control for baseline and time-dependent covariates, using the full observational period of available data, i.e. up to 9 years of follow up. The Cox proportional hazards regression will take into account the fact that individual matching (1:5, or 1:4 etc. according to the rule laid out in section 11.1.2.2) was performed (based on age, sex and propensity score) by considering the exposed patient and his/her matched controls as one stratum and including this as a stratum in the Cox proportional hazards model (18,19). Additional variables will be included in the model if available, as listed in Table 2.

The primary outcome will be analysed separately for each of the following roflumilast exposure variables:

1. Ever vs. never exposed,
2. Current exposure,
3. Previous exposure,
4. Duration of exposure,
5. Cumulative dose, and
6. Time since last dose.

For definitions of exposure variables see section 8.

11.1.3.1.1 Sensitivity Analyses and other Explorative Analyses for the Primary Outcome

The following sensitivity analyses and other exploratory analyses for the primary outcome will be performed.
Explorative analyses based on the Cox model will be used to investigate, whether the hazard ratio per year observed can be assumed to be constant.

A separate analysis of all cause mortality similar to the primary analysis where exposed patients dying less than 6 and 12 months, respectively, after Cohort Entry Date (and their matched controls) are excluded. This should exclude severely ill COPD patients where roflumilast is used as last resort.

The propensity score will be summarised for exposed and controls. Further, the propensity score will be categorised and within each category the 5-year mortality will be summarised for each group (exposed and controls). This will be tabulated and illustrated graphically. This illustrates the degree of consistency in risk as a function of probability of being prescribed roflumilast.

11.1.3.2 Analysis of Secondary Outcomes

See section 6.2 for detailed definitions of these outcomes. For the analysis of the secondary outcomes time to event data competing risk models (analogue to the model used for the primary analysis) will be used. While the estimate for the hazard ratio is unaffected by the presence of competing risks, the unbiased estimation of survival probabilities requires the application of a competing risk model. Matching will be accounted for in the model by stratification of the analysis and additional variables will be included in the model, if available, as listed in Table 2.

The analysis of the secondary outcomes will be performed using the following roflumilast exposure variables:

1. Ever vs. never exposure and
2. Duration of exposure.

The possible use of other roflumilast exposure variables in the analysis of secondary outcomes will be further defined in the SAP. For example, for outcomes such as serious diarrhoea, which may be associated with current use of roflumilast the exposure variables “current exposure” and “”time since last dose” can be used.

For (possibly) repeating events (such as hospitalisation) additionally a Poisson regression model will be fitted for the number of events (taking into account the matching). Time in the study will be included as an offset variable in the model.
For outcomes expected to be associated with recent rather than long-term roflumilast exposure (e.g. new diagnosis of depression) an additional sensitivity analysis will include only patients exposed during the last 6 months prior to the event, and their matching controls.

11.1.3.2.1 Death by Suicide or Hospitalisation for Suicide Attempt

For death by suicide or time to first hospitalisation for suicide attempt a competing risk model (based on the Cox proportional hazards model for the primary analysis) will be performed.
As sensitivity analysis a 'recent exposure' analysis will include only patients exposed during the last month prior to the event, and their matching controls. The model will be the same as above.

The number of hospitalisations for suicide attempt will be analyzed using a Poisson regression.

11.1.3.2.2 Hospitalisation for Any Cause

For time to first hospitalisation for any cause a competing risk model (based on the Cox proportional hazards model for the primary analysis) will be performed using the full observational period of available data.

The number of hospitalisations for any cause will be analyzed using a Poisson regression.

11.1.3.2.3 Selected Cardiovascular Events leading to Hospitalisation

For selected cardiovascular events leading to hospitalisation a competing risk model for time to first event (based on the Cox proportional hazards model for the primary analysis) will be performed using the full observational period of available data.

The number of hospitalisations due to selected cardiovascular events will be analysed using a Poisson regression.

11.1.3.2.4 Respiratory Disease-related Hospitalisation

For time to first respiratory disease related hospitalisation a competing risk model (based on the Cox proportional hazards model for the primary analysis) will be performed using the full observational period of available data.

The number of respiratory disease related hospitalisations will be analyzed using a Poisson regression.

11.1.3.2.5 New Diagnosis of Depression

For time to (first) new diagnosis of depression a competing risk model (based on the Cox proportional hazards model for the primary analysis) will be performed using the full observational period of available data.

As sensitivity analysis a ‘recent exposure’ analysis will include only patients exposed during the last 6 months prior to the event, and their matched controls. The model will be the same as above.

11.1.3.2.6 New Diagnosis of Malignant Neoplasm

For time to first new diagnosis of malignant neoplasm a competing risk model (based on the Cox proportional hazards model for the primary analysis) will be performed using the full observational period of available data.

11.1.3.2.7 Abnormal, unexplained weight loss

For time to first diagnosis of abnormal, unexplained weight loss a competing risk model (based on the Cox proportional hazards model for the primary analysis) will be performed using the full observational period of available data.

As sensitivity analysis a ‘recent exposure’ analysis will include only patients exposed during the last 3 months prior to the event, and their matched controls. The model will be the same as above.

11.1.3.2.8 Serious diarrhoea of non-infectious origin

For time to first hospitalisation with a discharge diagnosis of serious diarrhoea of non-infectious origin a competing risk model (based on the Cox proportional hazards model for the primary analysis) will be performed using the full observational period of available data.

As sensitivity analysis a 'recent exposure' analysis will include only patients exposed during the last month prior to the event, and their matched controls. The model will be the same as above.

11.1.3.2.9 New Diagnosis of Tuberculosis or New Diagnosis of Hepatitis B or C, or other Severe Viral Hepatitis Infection

For time to first new diagnosis of tuberculosis, Hepatitis B or C or other severe viral hepatitis infections a competing risk model (based on the Cox proportional hazards model for the primary analysis) will be performed using the full observational period of available data.

As sensitivity analysis a 'recent exposure' analysis will include only patients exposed during the last 6 months prior to the event, and their matched controls. The model will be the same as above.

The number of new diagnoses of tuberculosis, Hepatitis B or C or other severe viral hepatitis infections will be analyzed using a Poisson regression.

11.1.4 Analysis Across Regions

The databases within each region are locally owned. Combined analysis across regions will be performed on regional summary data only and not on the individual data level. The only possible exception is the Nordic countries where analysis of data on the individual data level may be feasible if the datasets prove compatible.

An overall estimate will not be calculated due to the structural differences between the regions. Consistency between regions will indicate robustness of results. Systematic differences between the results for each region will be examined by summary statistics and discussed in the report. Obvious differences between regions will be exploratorily examined to identify possible explanations. A separate SAP will be finalised for the analysis across regions prior to the first interim analysis of outcomes covering more than one region.

11.1.5 Consistency between Statistical Analysis Plans for all Regions

Each regional SAP and the SAP for the combined statistical analysis must be approved by the Lead Investigator and the AstraZeneca statistician to ensure the best possible comparability between regions, given the available sources.

11.1.6 Multiplicity Adjustment

The study will examine a number of secondary outcomes, raising the probability of falsely concluding that a significant association exists between roflumilast exposure and the occurrence of one or more outcomes. The following rules will therefore be adopted:

- If an outcome that was already suspected as possibly being associated with roflumilast yields a p-value < 0.05, then that association will be considered to have been independently validated
- If an outcome not suspected as possibly being associated with roflumilast consistently yields p-values < 0.05 in all regions then the result will be interpreted as evidence for a potential risk
- If an outcome that was NOT pre-specified as possibly associated yields a p-value < 0.05, then this result will be taken only as a new signal to be further investigated.

Pre-specified outcomes in this study are all outcomes listed in sections 6.1 and 6.2.

11.2 Interim Analyses

Since the study is based on existing data and the cohorts to be studied have been fully established prior to the first interim analyses, their mere conduct will not affect the results of later analyses. Hence, there will be no alpha adjustment for interim analyses. However, great caution must be exercised when drawing conclusions based on short duration of follow-up and on the large number of outcomes being examined.

The first interim analysis will be performed 2 years after selection of the final cohort. The first analysis will focus on 2-year all-cause mortality and present 2-year follow-up data on secondary outcomes. For all-cause mortality and secondary outcomes, follow-up data (up to 7 years) will be graphically presented.

Three years after selection of the final cohort the second interim analysis will be performed. The second analysis will provide data on the 3-year survival and it will update the information on the 2-year results.

Interim analyses will follow the same pattern until the final report.

11.3 Sample Size Considerations

COPD is projected to become the third leading cause of death worldwide by 2020, partly driven by increased smoking, but mainly caused by population demographics where more people live long enough to die from COPD (16). Therefore, any negative effect of roflumilast on mortality must be detected on top of a relatively high background mortality (20). Median expected survival times for patients with severe and very severe COPD are estimated to be 12 years and 7 years respectively, with 33% of patients with severe and 40% of patients with very severe COPD expected to die within a follow-up period of 5 years. Assuming a 3:1 distribution of patients with severe and very severe COPD (21,22,23), this turns into an expected 5-year all-cause mortality rate of 35%. Since frequent COPD exacerbations increase the risk of death, the 35% may be a conservative estimate for the roflumilast treated population.

An estimate of the number of roflumilast exposed patients during a 4 year time span is given in the following, using Danish prescription data as a representative for the countries to be studied. Based on 2009 data in the Danish prescription database (24) the Danish Medicines Agency has estimated that out of a total population of 5.5 million, 150,000 receive prescriptions for the treatment of symptoms of COPD. Since mild COPD is usually not presenting with clinical symptoms, and patients with moderate COPD often have not been diagnosed, it is estimated that 75,000 patients in Denmark have COPD GOLD stage III or IV, corresponding to 1.4% of the population. The ECLIPSE study (25) has revealed that roughly 22% of COPD patients with GOLD stage II, 33% of patients with GOLD stage III and 47% of patients with GOLD stage IV have frequent exacerbations in spite of ongoing therapy. Given the fact that roflumilast is the only available add-on maintenance therapy (besides symptomatic oxygen therapy) that can be offered to this population which continues to have exacerbations in spite of other COPD maintenance therapy, it seems realistic to expect that half of the patient population with GOLD stage III or IV and frequent exacerbations will be prescribed roflumilast therapy during the time span of the cohort establishment. Among these patients some will not fill their prescription, and an approximate 15% will experience initial gastrointestinal side effects leading to early treatment interruption.

During the establishment of the cohort there will be patients initially in stage II progressing into stage III and entering the roflumilast target population. This number is expected to make up for those receiving only single prescriptions.

With an estimated 75,000 Danish patients having COPD GOLD stage III or IV (in a 3:1 ratio), 33% and 47% of these, respectively, having frequent exacerbations in spite of other ongoing therapy, and 50 or 25% of these, respectively, receiving roflumilast therapy at some time point during the 4 years of cohort identification, the exposed cohort in Denmark would comprise approximately 0.24% or 0.12% of the total population. Extrapolating these percentages to the other regions, data on the following numbers of exposed patients would be available with full reimbursement in regions:

Region Patient number	DE	UK	NL	SE	DK	FI	NO	CA
With 50%	33,600	8,400	4,800	22,400	13,200	12,800	11,800	2,400
With 25%	16,800	4,200	2,400	11,200	6,600	6,400	5,900	1,200

In a region with more than 10,000 COPD patients receiving roflumilast treatment, 3 controls per exposed patient and an expected 5-year all-cause mortality rate in the control group of 35%, the power to detect at the 5% significance level an increase in the mortality rate of 2 percent points (from 35 to 37%) will be 95%, based on a comparison of rates between two groups with continuity correction. This model is simpler that the model planned for the primary analysis but the estimate gives the approximate size of the difference that can be detected based on the expected available datasets.

In regions with more exposed patients we have a high probability of detecting smaller differences in mortality rates.

In a region with only 2,000 exposed COPD patients, 3 non-exposed controls per exposed patient and an expected 5-year all-cause mortality rate in the control group of 35%, the power to detect at the 5% significance level an increase in the mortality rate of 4 percent

points (from 35 to 39%) will be 89%, based on a comparison of rates between two groups with continuity correction.

The majority of the secondary outcomes have a lower expected rate than the rate of 35% for all-cause mortality. For outcomes with lower background rate smaller absolute differences can be detected, e.g. if an outcome only occurs among 6% in the control group then an increase of 1 percent point (to 7%) in the roflumilast group can be detected with a 94% power in a population of 10,000 exposed COPD patients and 3 controls per exposed patients.

12 Discussion of Study Design

Database research and choice of databases

Electronic health care databases holding demographic data and data on health and consumption of health care system deliveries, including drug prescriptions or drug dispensations established for administrative purposes or epidemiological research have been widely used for drug safety evaluation, and recent ENCePP initiatives by the European Medicines Agency (EMA) (26) and the Food and Drug Administration Sentinel Initiative (27) aim to increase the quality and quantity of pharmacoepidemiological research, e.g. through cooperation between researchers, improved methodology and larger and more complete databases.

Main advantages of this methodology compared with prospective clinical trials or dedicated, prospective disease or drug registries are the absence of study impact on data collection, the availability of comparable (in terms of being collected the same way) historical data, the representativeness of the covered populations, the size of the populations and the relative inexpensiveness of such research, given the size of the populations. For this long-term study of roflumilast stable, established databases will be used, where data for the vast majority of patients can be followed for a minimum of 5 years. All databases inevitably have a minimum of migration, due to patients moving between countries or regions. Migration is expected to be lower than average in this population of patients with severe COPD and generally will be less when the database or group of databases to be linked cover all inhabitants in a country or larger region.

Identification of COPD cohorts

Lack of reimbursement or restricted reimbursement in a given region will limit the number of exposed patients and introduce bias in prescribing and dispensing. For these reasons sampling of patients within a given region must await granting of reimbursement in the region.

In order to avoid extraction of too small cohorts from which no solid information can be derived, the first matched cohorts in a given region will only be established, when

1. roflumilast has received reimbursement in the region, *and*/or
2. roflumilast sales figures in the region indicate that at least 2500 exposed individuals have been exposed after granting of reimbursement, *and*
3. A preliminary extract of the number of exposed individuals confirms the availability of data for at least 2000 individuals with first prescription after granting of reimbursement.

The awareness of COPD has increased over the last decade. COPD patients have often been assigned a diagnosis of chronic bronchitis and not COPD. For this reason, a diagnosis of chronic bronchitis will be used as a proxy for a COPD diagnosis in the absence of a COPD diagnosis and allow entry into analytic cohorts. In order to limit the impact of diagnostic misclassification, only patients aged 40 or older will be included. In the regions to be studied, COPD is very rare under this age. In databases without outpatient diagnostic information, patients with no hospital contact may not have a diagnosis of COPD or chronic bronchitis recorded in the databases. For these regions the COPD diagnosis will be based on data in a hierarchy where a COPD diagnosis ranks highest and is searched for first, a diagnosis of chronic bronchitis ranks second, and where chronic prescription of a typical COPD maintenance combination treatment to a patient without an asthma diagnosis is searched for only if no diagnosis of COPD or chronic bronchitis is available in the databases. When basing the diagnosis on prescription data and absence of an asthma diagnosis, there will remain a risk of including patients with asthma and no COPD and a risk of missing patients with COPD. This is the best option available for patient identification and considered acceptable in this study where the focus is on roflumilast safety and not effectiveness. Patients with a diagnosis of asthma as well as COPD or chronic bronchitis will not be excluded from the analyses, but since such patients may have a higher likelihood of exposure to roflumilast and at the same time a higher mortality risk, asthma as comorbidity will be included in the propensity score.

Confounding

Potential confounding, especially confounding by indication, is the main concern in this study. The establishment of a matched control cohort and the planned data analysis aims to control for measured, potential confounders, but balance of unknown and unmeasured confounders cannot be assumed with certainty. Therefore, residual confounding may remain which may e.g. lead to higher mortality risk estimates for the roflumilast exposed cohort, if roflumilast is channelled to more severely diseased COPD patients. Sensitivity analyses will be conducted to elucidate the potential impact of such confounders.

Mortality

COPD is projected to become the third leading cause of death in Europe, and the main driver of the disease in this geographical area is smoking. Other risk factors for COPD are exposure to other airborne particles, genetic background (e.g. alpha-1 antitrypsin deficiency), impaired lung growth and development during gestation, birth and childhood, oxidative stress, higher age, respiratory infections, previous tuberculosis, low socioeconomic status, low quality nutrition and comorbidities such as asthma (16). Most of these risk factors are also associated with the prognosis of the COPD patients, including their risk of death, and it is important that the cohorts to be compared are matched as closely as possible regarding such risk factors.

If the presence of a risk factor is likely to affect the decision to prescribe or not prescribe roflumilast to a patient, this risk factor must also be controlled in order to avoid erroneous interpretations of results.

In this study, the roflumilast exposed cohort will be compared with a cohort of non-exposed controls, matched on age, sex, years since diagnosis of COPD/chronic bronchitis, as available in the database, years of membership in database (as available), and propensity score. The propensity score models the probability of exposure to a given treatment, conditional on a patient's observed characteristics when the treatment is chosen. By matching on propensity score, exposed and non-exposed cohorts are formed with similar baseline characteristics with respect to factors influencing the administration of treatment (12). In the ideal situation where all relevant baseline characteristics are taken into account directly in the matching algorithm, all sources for confounding bias are balanced out.

Severity of COPD will enter the propensity score if this information is available, directly or by proxy. Proxy information on COPD severity will include previous hospitalisations for COPD exacerbations in the year prior to Cohort Entry, number of episodes during the last year of use of systemic corticosteroids and/or antibiotics, chronic use of oxygen and vaccination against pneumococci.

Patients with severe COPD tend to carry a heavy burden of co-morbidities (16), and the risk of death in the exposed cohort is relatively high. Since roflumilast is new in the market, it may even be higher in the first cohort to be assembled. The Charlson Comorbidity Index is shown to be a good predictor of mortality (15,28) and will be included in the propensity score. The index contains 19 categories of comorbidity, each with an associated weight: metastatic solid tumour or AIDS (weight 6), moderate or severe liver disease (weight 3), hemiplegia, moderate or severe renal disease, diabetes with end organ damage, any tumour, leukaemia, or lymphoma (weight 2), myocardial infarction, congestive heart failure, peripheral vascular disease, cerebrovascular disease, dementia, chronic pulmonary disease, connective tissue disease, ulcer disease, mild liver disease or diabetes (weight 1) . Other indexes have been developed for prediction of mortality but the Charlson Index is well established and widely used within epidemiology. ICD-9-CM and ICD-10 coding algorithms for the Charlson Index are available and will be used for standardisation of the output (29,30).

A concerning, related risk of selection bias is for ‘last resort prescribing’ of roflumilast to very ill patients with a high risk of dying in the near future. Such prescribing would transfer a non-exposed matching control into the exposed population shortly before death, with no chance for roflumilast to change the prognosis. Sensitivity analyses will therefore be performed, excluding exposed patients dying within 6 and 12 months of the second prescription, respectively.

Another concerning risk of selection bias is the risk for roflumilast being prescribed to patients with a mortality risk that is on average a bit higher than for the control group, with this increase in risk not being reflected in the elements of the propensity score. Age, sex, and years since diagnosis of COPD/chronic bronchitis, as available in the database, will also be matched between exposed and non-exposed to minimize this potential bias. Association of potential risk factors with prescribing of roflumilast

As a result of the current knowledge and of limited experience with roflumilast, the text of the approved European Summary of Product Characteristics and the Canadian Product Monograph ('labelling') (1,2), implies that the following conditions are likely to have an impact on the decision to prescribe roflumilast. At the same time they are likely to affect study outcomes:

Underweight: In the roflumilast clinical trials patients receiving roflumilast have had a small, but significant, unexplained weight loss compared with placebo. The currently approved labelling advice the prescriber that the body weight should be checked at each visit with the health care provider, and in the event of an unexplained and clinically concerning or pronounced weight decrease, roflumilast treatment should be terminated. This advice could lead prescribers to avoid prescribing of roflumilast to patients with a low body weight. In the general population, underweight has been associated with increased mortality risk, but it is debated whether or not underweight is an independent prognostic factor or the association is due to undiagnosed causes of low weight. Likewise, patients with a lower starting weight are more likely to be hospitalised with a diagnosis of unexplained weight loss than patients with a higher starting point and less urgency. Weight data is, however, not available in databases without outpatient data input and in most claims databases, and in medical records databases, weight data are not consistently available. Thus residual confounding due to less prescribing of roflumilast to underweight patients cannot be excluded. And weight may be an unmeasured confounder for hospitalisation due to abnormal, unexplained weight loss.

Diarrhoea is an identified adverse drug reaction of roflumilast. Gastrointestinal comorbidities like Morbus Chrohn and Irritable Bowel Disease are not contraindications to use of roflumilast, and there is no warning against use of roflumilast in such patients. However, the increased risk of gastrointestinal side effects related to use of roflumilast itself may limit prescribing to patients already prone to gastrointestinal symptoms, confounding the occurrence of diarrhoea of non-infectious origin severe enough to lead to hospitalisation.

Hepatic impairment: According to current labelling roflumilast should be used with caution in patients with mild hepatic impairment (Child-Pugh A) and should not be prescribed to patients with moderate or severe hepatic impairment (Child-Pugh B and C). Hepatic impairment (mild and moderate to severe liver disease) is a component of the Charlson Index and will be taken into account in the cohort matching, although severity of liver impairment is not available in ICD-codes and would need to be operationalized.

Depression: Clinical trials of roflumilast representing 3,261 person-years have revealed a small difference in the reported rate of depression for patients on roflumilast treatment versus placebo (1.2% vs. 0.8%), leading to a statement in the approved labelling that roflumilast is not recommended in patients with a history of depression associated with suicidal ideation or behaviour. The reported, very low incidences in the clinical trials points, however, to significant underreporting of depression in the clinical trials or a trial population not representative in terms of risk for depression. Reported prevalences vary (8-80% in a recent review) (31), but epidemiological studies consistently show that depression is more prevalent in COPD patients than in the general population. Furthermore, depressive symptoms are predictors of mortality in patients with COPD (32). Prior mood disorder will be included in the model for analysis of this outcome, and use of antidepressant medications will be employed as a proxy in the regions without outpatient diagnostic information.

Other conditions: Due to lack of experience, current labelling states that roflumilast treatment should not be initiated in patients with severe immunological diseases, severe acute infectious diseases or cancers (except basal cell carcinoma), in patients receiving immunosuppressive medications or theophyllin, in patients with a latent infection such as tuberculosis, viral hepatitis, herpes virus infection or herpes zoster, or in patients with congestive heart failure (NYHA grades 3 or 4). Congestive heart failure and cancer are components of the Charlson Index, and chronic prescription of immunosuppressive medications or theophyllin will be included separately in the propensity score.

Smoking

Smoking and occupational dust and chemicals are known as main drivers of COPD, and in terms of numbers smoking is by far the most important in Europe and Canada. Whereas occupational exposure to dust or chemicals will usually stop when symptomatic COPD has been diagnosed, this is not necessarily the case with smoking. Age at smoking initiation, total pack-years smoked, and current smoking status are predictive of COPD mortality (16). Continued smoking maintains the chronic inflammation in the lung tissue. Roflumilast belongs to a new class of anti-inflammatory drugs and has been shown to reduce the number of neutrophils and eosinophils in the sputum of patients with COPD (33). Prescribers may therefore anticipate that roflumilast is particularly effective in patients who continue to smoke, and a COPD patient who continues smoking may therefore have a higher likelihood of receiving a prescription on roflumilast, and at the same time a higher mortality risk. In

contrary, it may be argued that roflumilast is less likely to be prescribed to patients who continue smoking since such patients may be considered less compliant in general and thus less likely to manage the daily intake of roflumilast tablets.

No matter what direction the prescription bias may take, if any, it is important to control for this potential confounder. However, smoking status is not available in the databases without General Practice input. An attempt will be made to include smoking status in the propensity score in regions with General Practice input (PHARMO Linkage data and CPRD). If smoking status is missing for a relatively high percentage (e.g. > 10%) of the COPD patients in a region with GP data input, the outcome data will be analysed with and without smoking status included in the regression model. If high residual confounding by smoking status remains in data from regions without General Practice input, this will be a source of heterogeneity across regions.

Control of confounders – other outcomes

The secondary outcomes all-cause hospitalisation, major cardiovascular events, new diagnosis of any malignant tumour, and hospitalisations for respiratory causes are expected to occur with relatively high incidence, and the predictors are expected to be similar as for all-cause mortality. Therefore, the complex model planned for analysis of overall mortality will be used with only minor modifications for these outcomes as well.

For new diagnosis of tuberculosis or viral hepatitis, a low incidence is expected, and the number of predictors in the model will be more limited.

For new onset depression (or ongoing at Cohort Entry date) the following predictors will be included in the model: Previous episode of mood disorder, living alone and alcohol or drug abuse (if data are available) (34).

Suicide or hospitalisation for suicide attempt is expected to be rare, and for the suicide cases a high degree of misclassification with resulting underreporting is expected in death certificates (33). For this combined outcome other variables will be included in the model (prior intentional Self-Harm, alcohol or drug abuse, schizophrenia, previous mood disorder, living alone) (35).

13 Essential Documents

The following essential documents must be in the hands of AstraZeneca before the conduct phase can start in a given region:

- Written agreement between AstraZeneca, the Lead Investigator and the research institute of the Lead Investigator
- Original written agreements between AstraZeneca and the institutions holding the relevant databases
- A final protocol, signed and dated by the Sponsor and the Lead Investigator
- Signed and dated protocol agreement and amendment agreements, if any, with the original signature of the Lead Investigator
- Signed and dated protocol agreement and amendment agreements, if any, with the original signature of the Investigator who is responsible for the study in the region
- A final SAP for the given region, approved by the Investigator who is responsible for the study in the region, the Lead Investigator and AstraZeneca, and signed by the Investigator and Lead Investigator.
- Copy of written IEC / IRB approval / vote covering the study in the region, according to local regulations

14 Ethics

The study employs data that have already been collected for administrative or research purposes and implies no further contact to patients or prescribers. Therefore, patients or prescribers will not be informed about the study, and no consents to participation will be requested.

14.1 Ethical conduct of the Study

This study will be conducted in accordance with the protocol, the current version of the Declaration of Helsinki (36), and Good Pharmacoepidemiology Practices (GPP) (11). Special attention will be paid to local data protection regulations (37). The study will be registered to the ENCePP E-register of studies and an abstract of the results will also be published on the same site (http://www.encepp.eu/encepp/studiesDatabase.jsp). The ENCePP is a project led by the EMA to further stengthen the post-authorisation monitoring of medicinal products in Europe by facilitating the conduct of multi-centre, independent, post-authorisation studies focusing on benefit/risk.

14.2 Independent Ethics Committee / Institutional Review Board and other relevant Authorities

The Investigators must ensure that the protocol and any amendments are submitted to the relevant Independent Ethics Committees (IECs)/Institutional Review Boards (IRBs) according to local requirements, and that any required regular status reports to the IEC/IRB is prepared and submitted.

14.3 Patient Information and Written Informed Consent

Patient information is anonymised by the individual database holders. No personal identifiers are included in the data extracts. For this reason, it is not relevant or possible to approach patients to obtain informed consent.

15 Approval Data Protection Agency and other Authorities

The Investigators must ensure that any required study specific or general permission from the local Data Protection Agency or other authorities is available prior to data access.

16 Safety Reporting

Since patient identifiers are not revealed and no reporters are identified, no cases will meet the definitions for a Post Marketing Surveillance Adverse Drug Reaction. Only aggregate data will be collected in this study.

The progress of the study will be reported in any Periodic Safety Update Report (PSUR) to the EMA during conduct.

The results of the Interim Analyses scheduled yearly for 2nd quarter 2017 onwards will be reported in abbreviated study reports and submitted to the EMA yearly in the 3rd quarter of 2017 to 2019, respectively.

Any new safety issues revealed in the study will be reported to the EMA without delay, according to usual pharmacovigilance practices.

17 Audits and Inspections

17.1 Audit from AstraZeneca

AstraZeneca may audit the study to ensure that study procedures comply with the protocol, the SAP and any amendments. In addition, the statistical programs and other study documents may be audited. Access to patient data will not be available.

17.2 Inspection by IRB/IEC or Competent Authority

Representatives from the Data Protection Authority have the right to inspect the study at the site holding the data.

18 Reports

In cooperation with all Investigators and AstraZeneca, the Lead Investigator prepares the Abbreviated Interim Reports and the Final Study Report based on the results obtained from the relevant data extracts. The reports must be reviewed by all investigators and by the AstraZeneca study team and the Independent Scientific Advisory Committee. Comments from all reviewers must be considered for the final version. The Final Study Report must be available within 15 months from time of last data extract. Sponsor and Lead Investigator must each receive a signed original, and each Investigator a copy of the Final Study Report. The Lead Investigator and the Sponsor must receive a copy of all regional study results, either separately or as an appendix to the Final Study Report.

The study will be registered to the ENCePP's E-register of studies and the results will also be published on the same site (http://www.encepp.eu/encepp/studiesDatabase.jsp).

A summary of the main results of the study, whether positive or negative and including results from prematurely terminated studies, will always be made available to the public. An abstract of the study findings will be provided through the ENCePP E-register of studies, as soon as possible. The principal investigator may ask the ENCePP Secretariat to delay the publication of this abstract for a limited period pending response to peer-review comments.

The outcome of a study will always be presented in an objective and truthful manner providing a comprehensive and accurate description of the findings. In no way shall the interpretation and presentation of the results be aimed towards any commercial, financial or personal interests. AstraZeneca is entitled to view the final results and interpretations prior to submission for publication and to comment in advance of submission as agreed in the research contract and without unjustifiably delaying the publication.

19 Publications

The primary, final results of this study should be published in a primary publication with the Lead Investigator as first author and all other Investigators as co-authors based on guidelines set out by the International Committee of Medical Journal Editors (http://www.icmje.org/ethical_1author.html). AstraZeneca must receive any manuscript for commenting prior to submission and be given a minimum of 4 week's notice for such review. A publication committee should be established to govern additional publications. Partial or additional outcome results can only be published after the primary publication. AstraZeneca has the right to use the interim and final results for regulatory purposes and for internal presentation and distribution within the company and to partners.

20 Archiving of Study Documentation

The institutions of the Lead Investigator and other Investigators must as a minimum keep the essential documents relevant for their region, the protocol, any amendments, the SAPs for their region, the programs used for the analyses, the interim and the final study reports for 10 years after the sign-off of the final study report.

AstraZeneca shall receive a signed original of all the above documents except the statistical programs, and the approvals from the Data Protection Authorities, for which copies will be sufficient. AstraZeneca shall maintain this documentation for at least 10 years after roflumilast is no longer on the market in any country.

21 References

1. Daxas Summary of Product Characteristics, Annex 1, European Medicines Agency, May 2011.
2. Daxas Product Monograph, Health Canada, March 2011.
3. Pigeot I, Ahrens W. Establishment of a pharmacoepidemiological database in Germany: Methodological potential, scientific value and practical limitations. Pharmacoepidemiol Drug Saf. 2008;17:215-23.
4. Wood L, Martinez C. The General Practice Research Database. Role in Pharmacovigilance. Drug Saf. 2004;27:871-81.
5. Soriano JB, Maier WC, Visick G, Pride NB. Validation of general practitioner-diagnosed COPD in the UK General Practise Research Database. Eur J Epidemiol. 2001;17:1075-80.

6. Velthove KJ, Leufkens HGM, Schweizer RC, van Solinge WW, Souverein PC. Medication changes prior to hospitalisation for obstructive lung disease. A case-crossover study. Ann Pharmacother. 2010;44:267-73.
7. Furu K, Wettermark M, Martikainen JE, Almarsdottir AB, Sørensen HT. The Nordic countries as a cohort for pharmacoepidemiological research. Basic Clin Pharmacol Toxicol. 2009;106:86-94.
8. Huiart L, Ernst P, Suissa S. Cardiovascular morbidity and mortality in COPD. Chest. 2005;128:2640-6.
9. Curkendall SM, Lanes S, de Luise C, Stang MR, Jones JJ, She D, Goehring Jr E . Chronic obstructive pulmonary disease severity and cardiovascular outcomes. Eur J Epidemiol. 2006;21:803-13.
10. Directive 2001/20/EC of the European Parliament and of the Council of 4 April 2001 on the approximations of the laws, regulations and administrative provisions of the Member States relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use. Official Journal of the European Communities L/21/34 1.5.2001.
11. Epstein, M on behalf of ISPE. Guidelines for Good Pharmacoepidemiology Practices (GPP). Pharmacoepidemiol Drug Saf. 2005;14:589-95.
12. Stürmer T, Joshi M, Glynn RJ, Avorn J, Rothman KJ, Schneeweis. A review of the application of Propensity Score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods. J Clin Epidemiol. 2006;59:437-47.
13. Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. Pharm Stat. 2011;10:150-61.
14. Nielson LH, Lokkegaard E, Andreasen AH, Keilding N. Using prescription registries to define continous drug use: how to fill gaps between prescriptions. Pharmacoepidemiol Drug Saf. 2008;17:384-388.
15. Nielson LH, Lokkegaard E, Andreasen AH, Hundrup YA, Keilding N. Estimating the effect of current, previous and never use of drugs in studies based on prescription registries. Pharmacoepidemiol Drug Saf. 2009; 18:147-153.
16. From the *Global Strategy for the Diagnosis, Management and Prevention of COPD*, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2010.

17. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. J Chronic Dis. 1987;40:373-83.
18. Martens EP, de Boer A, Pestman WR, Belitser SV, Stricker BHC, Klungel OH. Comparing treatment effects after adjustment with multivariate Cox proportional hazards regression and propensity score methods. Pharmacoepidemiol Drug Saf. 2008;17:1-8.
19. Mandrekar JN, Mandrekar SJ. An Introduction to Matching and its Application Using SAS®, SUGI 29 Proceedings, Montreál, Canada [Internet] 2004, May 9-12 paper 208-29. Available from: http://www2.sas.com/proceedings/sugi29/208-29.
20. Jensen HH, Godtfredsen NS, Lange P, Vestbo J. Potential misclassification of causes of death from COPD. Eur Respir J. 2006;28:781-5.
21. Hoogendoorn M, Feenstra TL, Schermer TRJ, Hesselink AE, Rutten-van Mölken MPMH. Severity distribution of chronic obstructive pulmonary disease (COPD) in Dutch general practice. Respir Med. 2006;100:83-6.
22. Jones PW, Brusselle G, Dal Negro RW, Ferrer M, Kardos P, Levy ML, Perez T, Soler-Cataluña, van der Molen T, Adamek L, Banik N. Health-related quality of life in patients by COPD severity within primary care in Europe. Respir Med. 2011;105:57-66.
23. Agusti A, Calverley PMA, Celli B Coxson HO, Edwards LD, Lomas DA, MacNee W, Miller BE, Rennard S, Silverman EK, Tal-Singer R, Wouters E, Yates JC, Vestbo J, Characterisation of COPD heterogeneity in the ECLIPSE cohort. Respir Res. 2010;11:122.
24. Increased sales of medicine for treatment of chronic obstructive pulmonary disease (COPD). Danish Medicines Agency. 17 Nov 2010.
25. Hurst JR, Vestbo J, Anzueto A, Locantore N, Müllerova H, Tal-Singer R, Miller B, Lomas DA, Agusti A, MacNee W, Calverley P, Rennard S, Wouters EFM, Wedzicha JA. Susceptibility to Exacerbation in Chronic Obstructive Pulmonary Disease. N Engl J Med. 2010;363:1128-38.
26. ENCePP-The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. European Medicines Agency, 2011.
27. Behrman RE, Benner JS, Brown JS, McClellan M, Woodcock J, Platt R. Developing the Sentinel System – A national resource for evidence development. N Engl J Med. 2011;364:498-9.

28. De Groot V, Beckerman H, Lankhorst GJ, Bouter LM. How to measure comorbidity: a critical review of available methods. J Clin Epidemiol. 2003;56:221-9.
29. Hall WH, Ramachandran R, Narayan S, Jani AB, Vijayakumar S. An electronic application for rapidly calculating Charlson comorbidity score. BMC Cancer 2004;4:94-101.
30. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luhti J-C, Saunders LD, Beck CA, Feasby TE, Ghali WA. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care. 2005;43:1130-9.
31. Yohannes AM, Willgoss TG, Baldwin RC, Connolly MJ.Depression and anxiety in chronic heart failure and chronic obstructive pulmonary disease: Prevalence, relevance, clinical implications and management principles. Int J Geriatr Psychiatry. 2010;25:1209-21.
32. de Voogd JN, Wempe JB, Koëter GH, Postema K, van Sonderen E, Ranchor AV, Coyne JC, Sanderman R. Depressive symptoms as predictors of mortality in patients with COPD. Chest. 2009;135:619-25
33. Grootendorst DC, Gauw SA, Verhoosel RM, Sterk PJ, Hospers JJ, Bredenbröker D, Bethke TD, Hiemstra PS, Rabe KF. Reduction in sputum neutrophil and eosinophil numbers by the PDE4 inhibitor roflumilast in patients with COPD. Thorax. 2007; 62:1081-7.
34. Djernes JK. Prevalence and predictors of depression in populations of elderly: a review. Acta Psychiatr Scand. 2006;113:372-87.
35. Nordentoft M. Prevention of suicide and attempted suicide in Denmark. Epidemiological studies of suicide and intervention studies in selected risk groups. Dan Med Bull. 2007;54:306-69.
36. World Medical Association Declaration of Helsinki. Ethical principles for Medical Research Involving Human Subjects, Helsinki 1964, amended in Tokyo 1975, Venice 1983, Hong Kong 1989, South Africa 1996, Edinburgh 2000, and Seoul 2008.
37. Directive 95/46/EC of the European Parliament and of the Council of 254 October 1995 on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data. Official Journal of the European Communities L281/31 23.11.1995.

22 Supplementary Table: Summary of Main Changes to Protocol

	Change	Rationale
Section 8 added	Addition of section related to exposure definitions with more detail regarding definition of duration of exposure and cumulative exposure	This section has been added to address comments in the PRAC PSUR Assessment Report on RMP v11 and PSUR 5 dated 13 May 2013.
Sections 11.1.3.1 & 11.1.3.2	Further description of how primary and secondary analyses will be conducted	This detail has been added to address comments in the PRAC PSUR Assessment Report on RMP v11 and PSUR 5 dated 13 May 2013
Section 3.3	Sponsor personnel change	Administrative change due to change of study sponsor
Section 5.1	Milestones change	Administrative change due to change of study sponsor
Page 1	Study ID change	Administrative change due to change of study sponsor and study ID
Page 2	CO-ORDINATING STUDY MANAGER, approvers change	Administrative change due to change of study sponsor
Across the protocol	Change MAH and sponsor from Takeda to AstraZeneca	Administrative change due to change of study sponsor