Study Protocol

P3-C1-016

DARWIN EU® -Drug utilisation of salbutamol products for inhalation and therapeutic alternative inhalation products

02/12/2024

Version 3.0

Contents

LIST OF ABBREVIATIONS **4**

1. TITLE **5**

2. RESPONSIBLE PARTIES – STUDY TEAM **5**

3. ABSTRACT **6**

4. AMENDMENTS AND UPDATES **8**

5. MILESTONES **8**

6. RATIONALE AND BACKGROUND **8**

7. RESEARCH QUESTION AND OBJECTIVES **9**

8. RESEARCH METHODS **11**

8.1 Study type and study design 11
8.2 Study setting and data sources 11
8.3 Study period 17
8.4 Follow-up 17
8.5 Study population with inclusion and exclusion criteria 19
8.6 Variables 21
8.7 Study size 25
8.8 Analysis 25
8.9 Evidence synthesis 28

9. DATA MANAGEMENT **28**

9.1 Data management 28
9.2 Data storage and protection 29

10. QUALITY CONTROL **29**

11. LIMITATIONS OF THE RESEARCH METHODS **30**

12. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS **30**

13. GOVERNANCE BOARD ASPECTS **30**

14. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS **31**

15. OTHER ASPECTS **31**

16. REFERENCES **31**

17. ANNEXES **33**

Study title	DARWIN EU® - Drug utilisation of salbutamol products for inhalation and therapeutic alternative inhalation products
Protocol version	V3.0
Date	02/12/2024
EU PAS number	EUPAS1000000403

Active substance	1- Inhalation drugs as reliever therapy (inhalation form):
	• Salbutamol
	• Short acting Beta-2 agonists other than salbutamol:
	  ○ Terbutaline
	  ○ Fenoterol
	• Short acting anticholinergic drugs:
	  ○ Ipratropium bromide
	  ○ Oxitropium bromide
	• Combinations of Beta-2 Agonists and Anticholinergics:
	  ○ Fenoterol + ipratropium bromide
	  ○ Salbutamol + ipratropium bromide
	• Formoterol
	• Formoterol + inhaled corticosteroids (ICS):
	  ○ Formoterol + beclomethasone
	  ○ Formoterol + budesonide
	• Salbutamol + ICS:
	  ○ Salbutamol + beclomethasone
	  ○ Salbutamol + budesonide
	2- Salbutamol administered orally

Medicinal product	n/a
Research question and objectives	What is the real-world use of salbutamol (inhaled formulations)? 1. To describe the overall (i.e. all drug formulations combined) rates of prescribing inhaled salbutamol (irrespective of type of formulation) by calendar time (month, year). Monthly prescribing rates will be provided by database and healthcare setting (inpatient/outpatient). 2. To describe the rates of prescribing inhaled salbutamol by type of formulation and calendar time (year, month). Monthly prescribing rates will be provided by database and healthcare setting. 3. To describe the rate of prescribing other inhaled alternatives and oral salbutamol by calendar time (month, year). Monthly prescribing rates will be provided by database and healthcare setting. 4. To describe the distribution of predefined indication of use, sex and defined age group for inhaled salbutamol, by formulation.
Countries of study	United Kingdom, The Netherlands, Spain, Denmark, Germany, Croatia
Author(s)	M. Amini, m.amini@darwin-eu.org

LIST OF ABBREVIATIONS

Acronyms/term	Description
CDM	Common Data Model
CC	Coordinating centre
CIPH	Croatian Institute of Public Health
COPD	Chronic Obstructive Pulmonary Disease
COVID-19	Coronavirus Disease 2019
CPRD	Clinical Practice Research Datalink
DARWIN EU®	Data Analysis and Real-World Interrogation Network
DK-DHR	Danish Data Health Registries
DOI	Declaration of Interests
DQD	Data Quality Dashboard
DRE	Digital Research Environment
DUS	Drug Utilisation Study
EHR	Electronic Health Records
EMA	European Medicines Agency
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EU	European Union
GDPR	General Data Protection Regulation
GP	General Practitioner
ICD	International Classification of Diseases
ICS	inhaled corticosteroids
ID	Index date
IMASIS	Institut Municipal Assistència Sanitària Information System
IP	Inpatient
IPCI	Integrated Primary Care Information
IQR	Interquartile Range
IQVIA DA	IQVIA Data Analyzer Germany
NAJS	Croatian National Public Health Information System
OHDSI	Observational Health Data Sciences and Informatics
OMOP	Observational Medical Outcomes Partnership
OP	Outpatient
RxNorm	Normalized names for clinical drugs
SD	Standard Deviation
SIDIAP	The Information System for Research on Primary Care Spain
SNOMED	Systematized Nomenclature of Medicine
UK	United Kingdom
WHO	World Health Organisation
WONCA	World Organization of Family Doctors

1. TITLE

DARWIN EU® -Drug Utilisation of salbutamol products for inhalation and therapeutic alternative inhalation products

2. RESPONSIBLE PARTIES – STUDY TEAM

Study team role	Names	Organisation
Study Project Manager/Principal Investigator	Marzyeh Amini Katia Verhamme	Erasmus MC
Data Scientist(s)	Ross Williams Maarten van Kessel Cesar Barboza Ger Inberg Adam Black	Erasmus MC
Epidemiologist/ Clinical Domain Expert	Marzyeh Amini Guido van Leeuwen	Erasmus MC
Data Partner*	**Names**	**Organisation**
CPRD-GOLD	Antonella Delmestri	University of Oxford
DK-DHR	Claus Møldrup Elvira Bräuner Susanne Bruun Tine Iskov Kopp Cæcilie Brinth Christiansen	Danish Medicines Agency (DKMA)
IMASIS	Miguel-Angel Mayer Maria Angeles Leis Machin Juan Manuel Ramirez Anguita	IMIM and PSMAR
IPCI	Katia Verhamme	Erasmus MC
IQVIA DA Germany	Gargi Jadhav Isabella Kacmarczyl Akram Mendez Dina Vojinovic	IQVIA
NAJS	Jakov Vuković Ivan Pristaš Anamaria Jurčević Marko Čavlina Antea Jezidžić Pero Ivanko	The Croatian National Institute of Public Health
SIDIAP	Talita Duarte-Salles Anna Palomar Agustina Giuliodori Picco Irene López	IDIAPJGol

*Data partners' role is only to execute code at their data source, review and approve their results. They do not have an investigator role.

3. ABSTRACT

Title

DARWIN EU® – Drug utilisation of salbutamol products for inhalation and therapeutic alternative inhalation products

Rationale and background

Salbutamol is essential for managing asthma and chronic obstructive pulmonary disease (COPD) due to its rapid bronchodilation effects. The rising prevalence of these conditions in Europe, driven by aging populations and worsening air quality, has led to increased demand for salbutamol, especially in urban areas. A shortage would severely impact patient care, leading to challenges in managing acute symptoms and increased strain on alternative therapies, which are not as effective for immediate relief.

The aim of the study is to understand if salbutamol (inhaled formulation) use has been increasing over the last few years in Europe which will in turn inform a potential risk of shortage. And secondly to understand the impact of the shortage of salbutamol inhaled formulations on therapeutic alternative inhalation products. This exercise falls under preparedness and prevention activities.

Research question and objectives

Research questions

What is the real-world use of salbutamol (inhaled formulations)?

Objectives

The aim of this study is to determine the real-world use of salbutamol (inhaled formulations).

The specific objectives of this study are:

1. To describe the overall (i.e. all drug formulations combined) rate of prescribing inhaled salbutamol (irrespective of type of formulation) by calendar time (month, year). Monthly prescribing rates will be provided by database and healthcare setting (inpatient/outpatient).
2. To describe the rate of prescribing inhaled salbutamol by type of formulation and calendar time (year~~,~~ month). Monthly prescribing rates will be provided by database and healthcare setting.
3. To describe the rate of prescribing other inhaled alternatives and oral salbutamol by calendar time (month, year). Monthly prescribing rates will be provided by database and healthcare setting.
4. To describe characteristics of individuals treated with inhaled salbutamol in terms of indication of use, sex and age (in age categories) stratified by formulation and provided by database and healthcare setting.

Methods

Study design

- Population-level cohort study (Population-level drug utilisation analyses to objectives 1, 2, 3)
- Drug user cohort study (Patient-level drug utilization analyses to objective 4).

Study period

Study period will be between 01/01/2015 to the end of available date.

Population

The study population will include all individuals present in the databases for objectives 1,2,3 and all individuals who received prescriptions for inhaled salbutamol in the study period for objective 4. In these populations, rates of prescribing inhaled salbutamol and other inhaled alternatives and the indication for

use of inhaled salbutamol will be explored. Patients need to have at least 365 days of data visibility prior to index date.

Variables

Therapeutic drug class of interest as exposures (all inhaled formulations):

Drugs of interest consist of reliever therapy namely: (1) salbutamol (dry powder, pressurized metered dose inhaler, solution for inhalation), (2) short acting B2 agonists other than salbutamol (terbutaline, fenoterol), (3) short acting anticholinergic drugs (ipratropium bromide, oxitropium bromide), (4) fixed combinations of short acting B2 agonist + anticholinergic drug (fenoterol + ipratropium bromide; salbutamol+ ipratropium bromide), (5) formoterol, (6) formoterol + ICS (formoterol + beclomethasone; formoterol+ budesonide), (7) salbutamol+ ICS (salbutamol + beclomethasone; salbutamol + budesonide), (8) oral salbutamol

Conditions of interest:

Asthma, COPD and emphysema, respiratory conditions due to inhalation of chemical substances, bronchitis, lower respiratory tract infection, bronchospasm.

Other covariates:

Sex, age, calendar months and years

Data source

1. Clinical Practice Research Datalink (CPRD) GOLD, United Kingdom
2. Danish Data Health Registries (DK-DHR), Denmark
3. Institut Municipal Assistència Sanitària Information System (IMASIS), Spain
4. Integrated Primary Care Information (IPCI), Netherlands
5. IQVIA Disease Analyzer Germany (IQVIA DA Germany), Germany
6. Croatian National Public Health Information System (NAJS), Croatia
7. The Information System for Research on Primary Care (SIDIAP), Spain

Sample size

No sample size has been calculated for this drug utilisation study, as our primary focus is to determine the prescribing rate of drug utilization of inhaled salbutamol products and other inhaled alternative in each database, irrespective of the sample size. Based on a preliminary feasibility assessment, the expected number of persons counts for medication with inhaled salbutamol in the databases included in this study ranged from 36,000 (IMASIS) to 3,154,200 (CPRD GOLD).

Statistical analysis

Population-level drug utilisation

Monthly and yearly prescription rates of inhaled salbutamol and other inhaled alternatives (both treatment initiation and ongoing treatment episode) per 100,000 person-years will be estimated, stratified by database and healthcare setting and stratified by formulation for inhaled salbutamol.

Patient-level drug utilisation

Patient-level drug utilisation analyses will include describing the distribution of predefined indication of inhaled salbutamol use, sex and age (at index-date, categorized), stratified by type of formulation. Index date will be the date of each prescribing of salbutamol (irrespective whether incident or prevalent) during the study period.

The statistical analyses will be performed based on OMOP-CDM mapped data using "*DrugUtilization*" R package. A minimum cell counts of 5 will be used when reporting results, with any smaller count reported as <5.

4. AMENDMENTS AND UPDATES

None.

5. MILESTONES

Study milestones and deliverables	Planned dates
Draft Study Protocol	29 October 2024
Final Study Protocol	20 November 2024
Creation of Analytical code	20 November 2024
Execution of Analytical Code on the data	2nd December 2024
Draft Study Report	16 December 2024
Final Study Report	15 January 2025

6. RATIONALE AND BACKGROUND

Salbutamol, a short-acting beta-2 agonist, is a critical therapeutic agent used primarily in the management of asthma and chronic obstructive pulmonary disease (COPD). Inhaled salbutamol provides rapid bronchodilation and symptom relief, making it an essential medication in both chronic and acute care settings. Over recent years, the prevalence of asthma and COPD in Europe has increased, partly due to aging populations and worsening air quality. This has led to a corresponding rise in the demand for salbutamol, particularly in urban areas where environmental triggers for respiratory conditions are more common.(1, 2)

The increased utilisation of salbutamol inhalers in Europe raises concerns about the potential for supply shortages. Disruptions in the global pharmaceutical supply chain—exacerbated by events like the COVID-19 pandemic—have already highlighted vulnerabilities in the availability of essential medications.(3, 4) A shortage of salbutamol would have significant public health implications, as it is a first-line treatment for acute bronchoconstriction in both asthma and COPD patients. Without adequate supplies, patients may face difficulties managing exacerbations, increasing the risk of hospital admissions and morbidity.(5)

Additionally, a salbutamol shortage could create fluctuation effects across the respiratory care market, with increased demand for alternative inhalation therapies such as long-acting beta-agonists or corticosteroid combinations. These alternatives, while beneficial for long-term management, are not ideal for acute symptom relief and may place further strain on supply chains. Understanding current trends in salbutamol usage is therefore critical to informing preparedness and prevention strategies, allowing healthcare systems to anticipate shortages and mitigate the impact on therapeutic alternatives.(6)

The aim of the study is to understand if salbutamol (inhaled formulation) use has been increasing over the last few years in Europe which will in turn inform a potential risk of shortage. And secondly to understand the impact of the shortage of salbutamol inhaled formulations on therapeutic alternative inhalation products (alternatives for salbutamol). This exercise falls under preparedness and prevention activities.

7. RESEARCH QUESTION AND OBJECTIVES

Research question

What is the real-world use of salbutamol (inhaled formulations)?

Research objectives

The proposed objectives to be achieved in the study are described in **Table 1**.

Table 1. Primary and secondary research questions and objective.

A. Primary research question and objective.

Objectives 1, 2, and 3:	• To describe the overall (i.e. all drug formulations combined) rates of prescribing of inhaled salbutamol (irrespective of type of formulation) by calendar time (month, year). Monthly prescribing rates will be provided by database and healthcare setting (inpatient/outpatient). • To describe the rates of prescribing of inhaled salbutamol by type of formulation and calendar time (month, year). Monthly prescribing rates will be provided by database and healthcare setting. • To describe the rates of prescribing of other inhaled alternatives and oral salbutamol by calendar time (month, year). Monthly prescribing rates will be provided by database and healthcare setting.
Hypothesis:	n/a
Population *(mention key inclusion-exclusion criteria):*	The study population will include all individuals present in the databases. Within these databases, we will identify users of inhaled salbutamol, oral salbutamol, and other inhaled alternatives prescribed in the period between 01/01/2015 to the end of available data. Patients need to have at least 365 days of data visibility prior to index date.

Exposure:	1- Therapeutic drug class (all inhaled formulations):
	• Salbutamol
	• Short acting Beta-2 agonists other than salbutamol:
	  ○ Terbutaline
	  ○ Fenoterol
	• Short acting anticholinergic drugs:
	  ○ Ipratropium bromide
	  ○ Oxitropium bromide
	• Combinations of Beta-2 agonists and anticholinergics inhaled:
	  ○ Fenoterol + ipratropium bromide
	  ○ Salbutamol + ipratropium bromide
	• Formoterol:
	• Formoterol + ICS:
	  ○ Formoterol + beclomethasone
	  ○ Formoterol + budesonide
	• Salbutamol + ICS:
	  ○ Salbutamol + beclomethasone
	  ○ Salbutamol + budesonide

	2- Salbutamol administered orally
Comparator:	None
Outcome:	None
Time *(when follow up begins and ends):*	Follow-up will start on the respective date of the latest of the following: i) study start date (1st January 2015), ii) date on which individuals have 365 days of prior history. End of follow-up will be defined as the earliest of loss to follow-up, death, or end of observation period (the latest available data), whatever occurs first.
Setting:	Inpatient and outpatient setting using data from the following 7 data sources: CPRD-GOLD (United Kingdom), DK-DHR (Denmark), IMASIS (Spain), IPCI (Netherlands), IQVIA DA Germany (Germany), NAJS (Croatia), SIDIAP (Spain)
Main measure of effect:	Monthly prescribing rates of inhaled salbutamol, oral salbutamol, and other inhaled alternatives, stratified by formulation (i.e. only for inhaled salbutamol) and provided by database and healthcare setting (inpatient/outpatient).

B. Secondary research question and objective.

Objective 4:	To describe characteristics of individuals treated with inhaled salbutamol in terms of indication of use, sex and age (continuous and by age categories) stratified by formulation and provided by database and healthcare setting.
Hypothesis:	n/a
Population *(mention key inclusion-exclusion criteria):*	The study population will include all individuals prescribed inhaled salbutamol in the period between 01/01/2015 to the end of available data. To infer the treatment indications, windows of +/-7 days and -365 days of the index date will be used.
Exposure:	Inhaled salbutamol ○ Dry powder ○ Pressurized metered dose inhaler ○ Solution for inhalation ○ Formulation unknown (in case prescription is only mapped to the ingredient level)
Comparator:	None
Outcome:	None
Time *(when follow up begins and ends):*	Follow-up will start on the respective date of the latest of the following: i) study start date (1st January 2015), ii) date on which individuals have 365 days of prior history.

	End of follow-up will be defined as the earliest of loss to follow-up, death, or end of observation period (the latest available data), whatever occurs first.
Setting:	Inpatient and outpatient setting using data from the following 7 data sources: CPRD-GOLD (United Kingdom), DK-DHR (Denmark), IMASIS (Spain), IPCI (Netherlands), IQVIA DA Germany (Germany), NAJS (Croatia), SIDIAP (Spain)
Main measures:	The descriptive characteristics will include counts and proportions for binary categorical variables, as well as mean/median with interquartile range (IQR) for continuous variables.

8. RESEARCH METHODS

8.1 Study type and study design

The study types with related study designs are described in the **Table 2** below and are selected from the Draft Catalogue of Data Analytics.

A cohort study will be conducted using routinely collected health data from 7 databases. The study will comprise two consecutive parts:

- A population-level cohort study will be conducted to address objectives 1 ,2, 3, assessing the monthly rate of prescription of medication with inhaled salbutamol, oral salbutamol and other inhaled alternatives prescribing, stratified by formulation for salbutamol and provided by database and healthcare setting (inpatient/outpatient).
- Patient-level characterisation study will be used to address objective 4; to describe characteristics of individuals treated with inhaled salbutamol in terms of indication of use, sex and age (continuous and by age categories) stratified by formulation and provided by database and healthcare setting.

Table 2. Description of potential study types and related study designs.

Study type	Study design	Study classification
Population-Level DUS	Population level cohort	Off the shelf
Patient-Level DUS	New drug user cohort	Off the shelf

8.2 Study setting and data sources

This study will use routinely collected health data from 7 databases in the DARWIN EU® network of data partners from 6 European countries. All databases were previously mapped to the OMOP CDM.

Data sources

1. Clinical Practice Research Datalink (CPRD) GOLD, United Kingdom
2. Danish Data Health Registries (DK-DHR), Denmark
3. Institut Municipal Assistència Sanitària Information System (IMASIS), Spain

4. Integrated Primary Care Information (IPCI), Netherlands
5. IQVIA Disease Analyzer Germany (IQVIA DA Germany), Germany
6. Croatian National Public Health Information System (NAJS), Croatia
7. The Information System for Research on Primary Care (SIDIAP), Spain

These databases fulfil the criteria required for population-level and patient-level drug utilisation while covering different regions of Europe. Detailed information on the selected data sources is described in **Table 3**.

When it comes to assessing the reliability of data sources, the data partners are asked to describe their internal data quality process on the source data as part of the DARWIN EU® onboarding procedure. To further ensure data quality, we utilised the Achilles tool, which systematically characterises the data and generates data characteristics such as age distribution, condition prevalence per year, data density, measurement value distribution can be compared against expectations for the data. Additionally, the data quality dashboard (DQD) provides more objective checks on plausibility consistently across the data sources.

In terms of relevance, the selection of databases was based on the availability of data on the selected drugs of interest and the possibility to generate results promptly. The DARWIN EU® portal as well as information from the onboarding documents were used to assess whether databases have information on use of salbutamol or other inhaled alternatives. Data within the DARWIN EU® portal is maintained up to date by extracting the release dates for each dataset in the network and monitoring when data are out-of-date with the expected refresh cycle (typically quarterly or half-yearly). In addition, it is important to have a clear understanding of the time covered by each database release, as this can vary across different domains. To facilitate this, the *CDMOnboarding* (and *Achilles*) packages contain a 'data density' plot. This plot displays the number of records per OMOP domain monthly. This allows to get insights when data collection started, when new sources of data were added and when until when data was included. In addition, at time of inviting DPs, they were informed about study objectives and asked whether they could participate in the study.

More general-purpose diagnostic tools, CohortDiagnostics and DrugExposureDiagnostics, have been developed. The CohortDiagnostics package provides additional insights into cohort characteristics, record counts and index event misclassification. The *DrugExposureDiagnostics* package assesses ingredient specific diagnostics for drug exposure records. Upon finalisation of the study protocol and creation of the disease and drug cohorts of interest, these packages will be executed by the different DPs.

Table 3. Description of the selected data sources.

Country	Name of Database	Justification for Inclusion	Health Care setting (e.g. primary care, specialist care, hospital care)	Type of Data (EHR, claims, registries)	Number of active subjects	Feasibility count of exposure (inhaled salbutamol) *	Data lock for the last update
United Kingdom	CPRD GOLD	Adequate number of patients exposed to drug of interest (as documented in the feasibility assessment) and the possibility to generate results promptly. Use of CPRD contributes to geographical diversity of data sources included.	Primary care	EHR	17m	3,154,200	01/01/2024
Denmark	DK-DHR	Adequate number of patients exposed to drug of interest (as documented in the feasibility assessment and the possibility to generate results promptly. Use of DK-DHR contributes to geographical diversity of data sources included.	Inpatient hospital care and secondary outpatient care	EHRs, registries, others	8.5m	941,500	21/05/2024
Spain	IMASIS	Adequate number of patients exposed to drug of interest (as documented in the feasibility assessment) and the possibility to generate results promptly. Use of IMASIS contributes to geographical diversity of data sources included.	Inpatient hospital care and secondary outpatient care	EHR	1.1m	36,000	10/02/2024
Netherlands	IPCI	Adequate number of patients exposed to drug of interest (as documented in the feasibility assessment) and the possibility to generate results promptly. Use of IPCI contributes to geographical diversity of data sources included.	Primary Care	EHR	1.39m	317,000	30/04/2024

Germany	IQVIA DA Germany	Adequate number of patients exposed to drug of interest (as documented in the feasibility assessment) and the possibility to generate results promptly. Use of IQVIA DA Germany contributes to geographical diversity of data sources included.	Primary care and outpatient secondary care	Claims	43m	1,166,700	30/09/2023
Croatia	NAJS	Adequate number of patients exposed to drug of interest (as documented in the feasibility assessment) and the possibility to generate results promptly. Use of NAJS contributes to geographical diversity of data sources included.	Primary care, outpatient specialist care, and inpatient care	Registries	5.4m	558,100**	17/11/20 23
Spain	SIDIAP	Adequate number of patients exposed to drug of interest (as documented in the feasibility assessment) and the possibility to generate results promptly. Use of SIDIAP contributes to geographical diversity of data sources included.	Primary care	EHR	5.8m	1,810,800**	30/06/2023

*Based on counts from approved feasibility assessment using the drug exposure table. The count of exposure is the count of unique individuals prescribed salbutamol at least once in the entire data.

** Drug exposure person counts using drug era (as PC for drug exposure table are not available within the portal).

Clinical Practice Research Datalink GOLD, United Kingdom (University of Oxford)

The Clinical Practice Research Datalink (CPRD) GOLD is a database of anonymised electronic health records (EHR) from General Practitioner (GP) clinics in the UK that use the Vision® software system for their management.(7) The source population encompasses 98% of the UK, registered with GPs responsible for non-emergency care and referrals. Participating GPs provide CPRD EHR for all registered patients who did not specifically request to opt out of data sharing. Covering 4.6% of the current UK population, GOLD includes 4.9% of contributing GP practices, providing comprehensive information within its defined source population. GOLD contains data from all four UK constituent countries and the current regional distribution of its GP practices is 5.7% in England, 55.6% in Scotland, 28.4% in Wales, and 10.2% in Northern Ireland (May 2022).

GOLD data include patient's demographic, biological measurements, clinical symptoms and diagnoses, referrals to specialist/hospital and their outcome, laboratory tests/results, and prescribed medications. GOLD has been assessed and found broadly representative of the UK general population in terms of age, gender, and ethnicity.(7) GOLD has been widely used internationally for observational research to produce nearly 3,000 peer-reviewed publications, making GOLD the most influential UK clinical database so far(8-10).

Danish Data Health Registries (DK-DHR), Denmark

Danish health data is collected, stored and managed in national health registers at the Danish Health Data Authority and covers the entire population which makes it possible to study the development of diseases and their treatment over time. There are no gaps in terms of gender, age and geography in Danish health data due to mandatory reporting on all patients from cradle to grave, in all hospitals and medical clinics. Personal identification numbers enable linking of data across registers, so it captures data on all Danes throughout their lives, regardless of whether they have moved around the country. High data quality due to standardisation, digitisation and documentation means that Danish health data is not based on interpretation. The Danish Health Data Authority is responsible for the national health registers and for maintaining and developing standards and classifications in the Danish healthcare system. Legislation ensures balance between personal data protection and use. The current data release includes data on the entire Danish population of 5.9 million persons from 1995. It includes data from the following registries: The central Person Registry, The National Patient Registry, The Register of Pharmaceutical Sales, The National Cancer Register, The Cause of Death registry, and Coronavirus disease 2019 test and vaccination Registries.

The Institut Municipal Assistència Sanitària Information System (IMASIS), Spain

The Institut Municipal Assistència Sanitària Information System (IMASIS) is the Electronic Health Record (EHR) system of Parc de Salut Mar Barcelona (PSMar) which is a complete healthcare services organisation. Currently, this information system includes and shares the clinical information of two general hospitals (Hospital del Mar and Hospital de l'Esperança), one mental health care centre (Centre Dr. Emili Mira) and one social-healthcare centre (Centre Fòrum) including emergency room settings, that are offering specific and different services in the Barcelona city area (Spain). At present, IMASIS includes clinical information from around 1 million patients with at least one diagnosis and who have used the services of this healthcare system since 1990 and from different settings such as admissions, outpatients, emergency room and major ambulatory surgery. The diagnoses are coded using The International Classification of Diseases ICD-9-CM and ICD-10-CM. The average follow-up period per patient in years is 6.37 (SD±6.82). IMASIS-2 is the anonymized relational database of IMASIS which is used for mapping to OMOP including additional sources of information such as the Tumours Registry.

The Integrated Primary Care Information (IPCI), the Netherlands

The Integrated Primary Care Information (IPCI) database is a longitudinal observational database containing routinely collected data extracted from computer-based patient records of a selected group of general practitioners (GPs) across the Netherlands.(11) IPCI was started in 1992 by the department of Medical Informatics of the Erasmus University Medical Center in Rotterdam. The current database includes patient records from 2006 on, when the size of the database started to increase significantly. The demographic composition of the IPCI population mirrors that of the general Dutch population in terms of age and sex. Although the geographical spread is limited, GP practices are located in urban and non-urban areas.

Patient-level data includes demographic information, patient's complaints and symptoms, diagnoses, laboratory test results, lifestyle factors and correspondence with secondary care, such as referral and discharge letters. For complaints, symptoms and diagnoses, Dutch GPs use International Classification of Primary Care (ICPC-1) coding, an international standard developed and updated by the World Organization of Family Doctors' (WONCA) International Classification Committee.

IPCI data quality has been previously documented and IPCI has proved valuable for epidemiological studies.(12-16) In terms of quality control, extensive quality control steps are performed prior to each data release. These include comparison of patient characteristics between practices and checks to identify abnormal temporal data patterns in practices. Additional checks include over 200 indicators related to population characteristics (e.g. reliability of birth and mortality rates) and medical data (e.g. availability of durations of prescriptions, completeness of laboratory results, availability of hospital letters and prescriptions, proportion of patients with blood pressure measurement, etc.(11) Based on this information, two quality scores have been created. Practices with low scores have been excluded.

IQVIA Disease Analyser (DA) Germany, Germany

IQVIA Disease Analyzer (DA) Germany is a database of de-identified electronic medical records from specialized and general primary practices (GP) in Germany since 1992. This dataset encompasses approximately 3% of all outpatient practices within Germany, ensuring a substantial representation of the national healthcare landscape.(17, 18) The sampling methods used for practice selection, taking into account physician's demographics, specialty focus, community size category and federal state location, was instrumental in constructing a database that accurately mirrors the diverse spectrum of healthcare providers in the country.(17, 18) Consequently, data within IQVIA DA Germany database has been demonstrated to be representative of general and specialised practices throughout Germany.

The database contains demographics records, basic medical data, disease diagnosis according to International Classification of Diseases, 10th revision (ICD-10), and prescription records.(18) While the database partly records information on deaths and procedures, it currently does not support linkage with external data sources and therefore information on mortality is incomplete. Routine updates are conducted at regular intervals. The quality of data is assessed based on several criteria including completeness of information and correctness (e.g. linkage between diagnosis and prescriptions). IQVIA DA Germany is suitable for pharmacoepidemiologic and pharmacoeconomic studies as previously demonstrated.(18-20)

National Public Health Information System (NAJS), Croatia

The National Public Health Information System (Nacionalni javnozdravstveni informacijski sustav - NAJS) is an organised system of information services by the Croatian Institute of Public Health (CIPH). This database was established in 1998, with nationwide coverage, representing approximately 5.4 million inhabitants. Settings covered include public primary, secondary/outpatient, and inpatient care. Data is retrieved primarily from EHR and holds information on demographics, inpatient and outpatient visits, conditions and procedures, drugs (outpatient and inpatient prescriptions), measurements, and inpatient and outpatient

dates of death. NAJS provides linkage between medical and public health data collected and stored in health registries and other health data collections, including cancer registry, mortality, work injuries, occupational diseases, communicable and non-communicable diseases, health events, disabilities, psychosis and suicide, diabetes, drug abuse and others. Data is being collected from 2015 in Central electronic health information system in Croatia and from 2017 in National public health information system. The CDM population comprises all publicly insured persons residing in Croatia starting in 2017.

Information System for Research in Primary Care (SIDIAP), Spain

The Information System for Research in Primary Care (SIDIAP) is a dynamic database of pseudo-anonymized electronic health records of the primary care patient population in Catalonia, Spain.(21) It contains data of approximately 80% of the Catalan population registered in over 280 primary care practices throughout Catalonia since 2005.

The database contains data recorded in primary care centres on a daily basis. Additionally, it integrates data from external sources including biomarkers data from laboratories and records of drug prescription and dispensation. The dataset covers demographics, all-cause mortality, disease diagnoses classified under the International Classification of Diseases 10th revision (ICD-10), prescription and dispensation records of drugs, results of laboratory tests, socio-economic indicators, vaccination records, lifestyle information, parent–child linkage and various clinical parameters. Additional data from other data sources such as hospital discharges, mental health centres or specific disease registries can be obtained through diverse linkages. The demographic composition within SIDIAP closely mirrors that of the broader Catalan population, encompassing a representative spectrum of geographic distribution, age, and sex proportions. The database is updated every 6 months.

SIDIAP data quality has been previously documented and SIDIAP has proved valuable for epidemiological studies.(19, 22-29) In terms of data integrity and reliability, SIDIAP has been subject to rigorous evaluation. Quality checks have been implemented including central identification of duplicate patient ID and visual inspection for temporal patterns in the registry of a certain variable. Furthermore, the data undergoes assessment for availability (longitudinally and reliability), plausibility (range checks and unusual values) and consistency using visualization tools. Specifically, for biochemistry data, consistency for measurements taken in different laboratories is assessed, and unit conversion is undertaken when needed.

8.3 Study period

The study period will be from 1st of January 2015 to the most recent data available for each contributing data sources (see **Table 3**).

8.4 Follow-up

For both population-level and patient-level drug utilisation analyses, follow-up will start on the respective date of the latest of the following: i) study start date (1st January 2015), ii) date on which individuals have 365 days of prior history (not for children < 1 year of age).

End of follow-up will be defined as the earliest of loss to follow-up, death, or end of observation period (the latest available data), whichever occurs first.

Operational definition of index date and other primary time anchors are described in **Table 4**.

Table 4. Operational definition of time 0 (index date) and other primary time anchors.

Study population name(s)	Time Anchor Description (e.g. time 0)	Number of entries	Type of entry	Washout window	Care Setting[1]	Code Type[2]	Diagnosis position	Incident with respect to…	Measurement characteristics/ validation	Source of algorithm
All patients from database eligible for the study for objectives 1,2 and 3	Date of the start of study period with sufficient prior data availability.	Multiple entries	The first date of eligibility of patients	n/a	OP & IP	n/a	n/a	n/a	n/a	n/a
All participants from the database eligible for the study and being prescribed inhaled salbutamol, for objective 4	Each date during follow-up on which an individual was prescribed inhaled salbutamol (irrespective whether incident or prevalent use).	Multiple entries	Prescription	n/a	OP & IP	RxNorm	n/a	n/a	n/a	n/a

ID = index date,

[1] IP = inpatient, OP = outpatient, n/a = not applicable

[2] The type(s) of clinical codes that are used to define the time 0 (or another primary anchor) criterion.

Prescribing rate estimation first requires identifying an appropriate denominator population and determining the corresponding observation time. Study participants in the denominator population will begin contributing person time at risk as described above in section 8.4.

An example of entry and exit into the denominator population is shown in **Figure 1**. In this example, person ID 1 already has sufficient prior history before the study start date and the observation period ends after the study end date, so this person will contribute during the entire study period. Person IDs 2 and 4 enter the study only when they have sufficient prior history. Person ID 3 leaves when exiting the database (the end of the observation period). Lastly, person ID 5 has two observation periods in the database. The first period contributes time from study start until end of observation period, the second starts contributing time again once sufficient prior history is reached and exits at study end date.

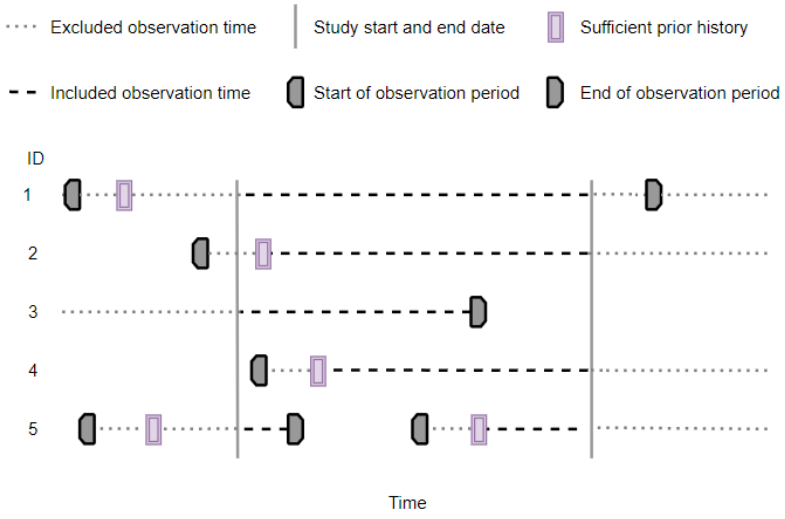

Figure 1. Included observation time for the denominator population.

8.5 Study population with inclusion and exclusion criteria

8.5.1 Population-level utilisation of selected medicinal products

The study cohort will include all individuals registered in the database between 1st of January 2015 and the most recent available data, with at least 365 days of data visibility prior to becoming eligible for study inclusion. This requirement of at least 365 days of prior data history will not hold for children <1 year of age.

Patients will be excluded if they do not have sufficient data availability prior to study start.

8.5.2 Patient-level utilisation of selected medicinal products

The study cohort will include all users of inhaled salbutamol in the period between 1st of January 2015 and with at least 365 days of data visibility prior to the first prescription of salbutamol since study start. This requirement of at least 365 days of prior data history will not hold for children <1 year of age.

The operational definitions of the inclusion and exclusion criteria are presented by means of **Table 5** and **Table 6**, respectively.

Table 5. Operational definitions of inclusion criteria.

Criterion	Details	Order of application	Assessment window	Care Settings[1]	Code Type[2]	Diagnosis position	Applied to study populations:	Measurement characteristics/ validation	Source for algorithm
Observation period in the database during the period 2015 and the most recent available data	All individuals present in the period 1st of January 2015 and the most recent available data	n/a	n/a	IP & OP	n/a	n/a	All individuals within selected databases	n/a	n/a
Prior database history	Study participants will be required to have 365 days of prior history observed before contributing observation time (except for children < 1 year of age)	After*	[-365, -1]	IP & OP	n/a	n/a	All individuals within selected databases	n/a	n/a

[1] IP = inpatient, OP = outpatient, n/a = not applicable

[2] The type(s) of clinical codes that are used to define the inclusion criteria.

*Order of application specifies whether the eligibility criterion is applied before or after selection of the study entry date. For example, selecting “before” means that all possible study entry dates are identified, and then one or more is chosen. For instance, selecting 'after' means that the first possible study entry date is chosen, followed by the application of the inclusion and/or exclusion criteria. If the patient does not meet the criterion, then the patient drops out

Table 6. Operational definitions of exclusion criteria.

Criterion	Details	Order of application	Assessment window	Care Settings[1]	Code Type	Diagnosis position	Applied to study populations:	Measurement characteristics/ validation	Source for algorithm
n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

8.6 Variables

8.6.1 Exposures

For this study, exposures of interest are prescription records of inhaled salbutamol, other inhaled alternatives, and oral salbutamol:

- Salbutamol
- Short acting Beta-2 agonists other than salbutamol:
  - Terbutaline
  - Fenoterol
- Short acting anticholinergic drugs:
  - Ipratropium bromide
  - Oxitropium bromide
- Combinations of Beta-2 agonists and anticholinergics:
  - Fenoterol + ipratropium bromide
  - Salbutamol + ipratropium bromide
- Formoterol
- Formoterol and ICS:
  - Formoterol + beclomethasone
  - Formoterol + budesonide
- Salbutamol and ICS:
  - Salbutamol + beclomethasone
  - Salbutamol + budesonide
- Salbutamol administered orally:
  - Oral salbutamol

The preliminary concept/code lists used for the identification of exposures are described in **Appendix I, Table S 1**. These will be refined during the study execution following the DARWIN EU® phenotyping standard processes, which involve the review of code lists by clinical experts, and the review of phenotypes after their execution in the participating databases.

The main salbutamol cohorts of interest consist of inhaled salbutamol monotherapy. Fixed combination of Beta-2 agonists and anticholinergics will consist of a cohort defined based on concept IDs referring to the fixed combination or use of respective Beta-2 agonist and anticholinergics prescribed on the same date. The same approach will apply to other combinations of Formoterol + ICS and Salbutamol + ICS as inhaled alternatives.

The operational definition of exposure is described by means of **Table 7.**

Table 7. Operational definitions of exposure.

Exposure group name(s)	Details	Washout window	Assessment Window	Care Setting[1]	Code Type	Diagnosis position[2]	Applied to study populations	Incident with respect to…	Measurement characteristics/ validation	Source of algorithm
Inhaled salbutamol prescribing with different formulations	Preliminary code list provided in Appendix I	n/a	Calendar month and year	IP & OP	RxNorm	n/a	All individuals present in the database during the study period and with sufficient database history (not for children < 1 year)	n/a	n/a	n/a
Other inhaled alternatives prescribing	Preliminary code list provided in Appendix I	n/a	Calendar month and year	IP & OP	RxNorm	n/a	All individuals present in the database during the study period and with sufficient database history (not for children < 1 year)	n/a	n/a	n/a

[1] ID = index date, IP = inpatient, OP = outpatient, n/a = not applicable
[2] Specify whether a diagnosis code is required to be in the primary position (main reason for encounter)

8.6.2 Outcome/s

None.

8.6.3 Other covariates, including confounders, effect modifiers and other variables

Pharmaceutical formulations of inhaled salbutamol will be assessed at time of each prescribing. The following pharmaceutical formulations will be used for stratification for population-level prescribing rates analyses (objective 2) and patient-level descriptive analyses (objective 4):

- Inhalation via solution
- Inhalation via dry powder
- Inhalation via pressured metered dose
- Formulation unknown (in case prescription is only mapped to the ingredient level)

Other covariates in the patient-level utilisation study (objective 4) will include:

- Sex: male or female
- Age: described as continuous variable and by age categories at index date. Age categories will be the following:
    - 0 to 23 months
    - 2 to 11 years
    - 12 to 17 years
    - 18 to 45 years
    - 46 to 65 years
    - 66 to 75 years
    - > 75 years
- Conditions considered as indication for inhaled salbutamol use, will be assessed within +/-7 days and -365 days of the index date (i.e. date of first prescribing of inhaled salbutamol during the study period). These indications will be identified based on the presence of disease codes. The preliminary list of concepts for prespecified conditions of interest (by standard SNOMED code) are described in **Appendix I, Table S 2**.
  Conditions of interest:
    - Asthma
    - COPD with or without emphysema
    - Respiratory conditions due to inhalation of chemical substances
    - Bronchitis
    - Lower respiratory tract infection
    - Bronchospasm

The operational definition of the covariates is described in the **Table 8**.

Table 8. Operational definitions of covariates.

Characteristic	Details	Type of variable	Assessment window	Care Settings[1]	Code Type	Diagnosis Position[2]	Applied to study populations	Measurement characteristics/ validation	Source for algorithm
Type of inhaled salbutamol formulation	Pharmaceutical formulations of inhaled salbutamol including solution, dry powder, pressured metered dose	Binary	At time of index date	IP & OP	n/a	n/a	Inhaled salbutamol users	n/a	n/a
Demographics	Distribution of sex and age (continuous and by age category) at index date	Binary, continuous	At time of index date	IP & OP	n/a	n/a	Inhaled salbutamol users	n/a	n/a
Indication of use	Prespecified indications of interest within the window of +/- 7 days and -365 days of the index date	Binary	[-7, +7] or [-365, -1] of index date	IP & OP	SNOMED	n/a	Inhaled salbutamol users	n/a	n/a

[1] IP = inpatient, OP = outpatient, n/a = not applicable

8.7 Study size

No sample size has been calculated for this drug utilization study, as our primary focus is to examine drug utilization of inhaled salbutamol, oral salbutamol and other inhaled alternatives regardless of sample size. Based on a preliminary feasibility assessment, the expected number of persons counts for medication with inhaled salbutamol in the databases included in this study range from 36,000 (IMASIS) to 3,154,200 (CPRD GOLD). These numbers are based on the overall number of inhaled salbutamol prescriptions registries in each database with no filter by study period or inclusion and exclusion criteria.

8.8 Analysis

The analysis will include calculation of population-based prescription rates, as described in section 8.8.5 as well as characterisation of individuals at time of each prescribing of inhaled salbutamol during the study period. The type of analysis by study type is presented in **Table 9**.

Table 9. Description of study types and types of analysis.

Study type	Study classification	Type of analysis
Population-Level DUS	Off-the-shelf	- Population-based prescription rates of use of a drug
Patient-Level DUS	Off-the-shelf	- Age distribution (continuous and by age category) - Sex distribution (counts and proportion of females/males) - Frequency and % of indications

8.8.1 Federated Network Analyses

All analyses will be conducted separately for each database, and will be carried out in a federated manner, allowing analyses to be run locally without sharing patient-level data.

Before sharing the study package, test runs of the analytics will be performed on a subset of the data sources and quality control checks will be performed. After all the tests are passed (see section 11 Quality Control), the final package will be released in a version-controlled study repository for execution against all the participating data sources.

The data partners will locally execute the analytics against the OMOP-CDM in R Studio and review and approve the default aggregated results. They will then be made available to the principal investigators and study team in secure online repository (Data Transfer Zone). All results will be locked and timestamped for reproducibility and transparency. The study results from all data sources will be reviewed, after which they will be shared with the team, allowing the study dissemination phase to begin.

8.8.2 Patient privacy protection

All analyses will be conducted separately for each database, and will be carried out in a federated manner, allowing analyses to be run locally without sharing patient-level data. Cell counts <5 will be suppressed when reporting results to comply with the database's privacy protection regulations.

8.8.3 Statistical model specification and assumptions of the analytical approach considered

R-packages

We will use the R packages "*IncidencePrevalence*"(30) package for the population-level estimation of drug utilization and "*CohortDiagnostics*", "*PatientProfiles*" and "*DrugUtilization*" for patient-level drug utilisation analyses including patient-level characterisation.

8.8.4 Methods to derive parameters of interest

Users of inhaled salbutamol

Users of inhaled salbutamol will be defined based on each prescription during the study period. For these individuals characteristics will be described.

Sex

Distribution of sex will be described as the proportion of male and female patients within the study population.

Age

Age will be described as a continuous variable and also categorized into predefined age groups at the index date. Summary statistics, such as mean, median, and interquartile range will be provided for the continuous age variable. For the age categories, proportions of individuals in each group will be reported.

Indication

The conditions considered indications for inhaled salbutamol use will be described in terms of their proportions within the study population.

Calendar time

Calendar time will be based on the calendar month/year at each prescription date.

Setting

Whether or not patient was in an inpatient or outpatient will be assessed at index date (date of each prescription of inhaled salbutamol and other inhaled alternatives).

Person-time from practices that cannot prescribe salbutamol will not be included in the denominator since they do not contribute to prescriptions. The results will not be stratified by primary or secondary care or any other healthcare categories, except for inpatient (hospital) and outpatient settings if possible. If this distinction cannot be made, the data will include all healthcare settings available for the data sources.

8.8.5 Methods to planned to obtain estimates with confidence intervals of measures of occurrence

8.8.5.1 Population-level drug utilization study

Prescription rates calculations will be conducted separately for inhaled salbutamol and other inhaled alternatives.

Prescription rates calculations

Monthly prescription rates (per calendar year) of use of inhaled salbutamol, other inhaled alternatives and use of oral salbutamol will be calculated as the number of prescriptions issued within the month per 100,000 person-months of the population at risk of getting exposed. Time-at-risk of subjects who die will be censored at the time of death. Similarly, time at risk of subjects who are lost to follow-up will be censored at the time of loss to follow-up (last contact). Subjects with data until the end of the study period with or without experiencing exposure will be administratively censored at the end of the study period. Prescription rates will be reported together with 95% Poisson confidence intervals.

An illustration of the calculation of prescription rates of medication use of inhaled salbutamol, oral salbutamol and other inhaled alternatives is shown below in **Figure 2**. For the numerator all prescriptions within the respective calendar month will be counted. For all patients, follow up time is time within the study period and all prescriptions will be counted.

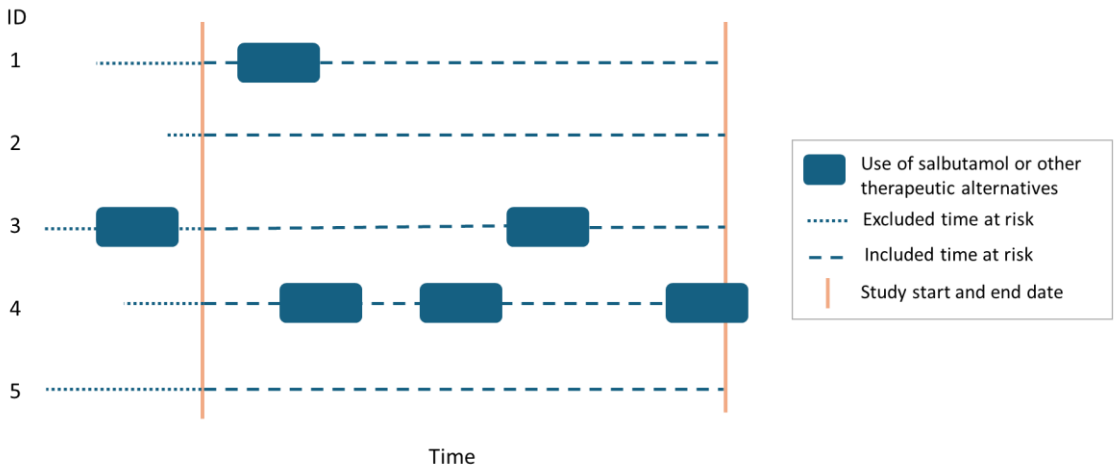

Figure 2. Person time at risk example

8.8.5.2 Patient-level drug utilization study

Characteristics at time of each prescribing (irrespective whether incident or prevalent) of inhaled salbutamol during the study period will be described.

Characteristics include indication of use, age and sex. Indication of use will be assessed within -7/+7 days and -365 days of the index date. The indication of use will not be mutually exclusive as there might be more than one indication of use, especially as assess within a year prior to the index date and the predefined indication of use not only consists of conditions but also symptoms (i.e. bronchospasm).

Characteristics will be provided by type of formulation of inhaled salbutamol (dry powder, pressurized metered dose inhaler, solution for inhalation) at index date. The index date is the date of each prescription, meaning that each individual may have multiple index dates.

8.8.6 Methods to deal with missing data

For the drug utilisation studies we assume that the absence of a prescription record means that the person does not receive the respective drug. For indications, we assume that the missingness of a record of the respective condition mean that that condition is not the indication for the drug prescription.

8.8.7 Description of sensitivity analysis

None.

8.8.8 Output

Output will include the following:

PDF report including an executive summary, and the following table(s) and figure(s).

Figure 1. Monthly prescribing rates (e.g. rates in January 2015, February 2015, March 2015, etc) of inhaled salbutamol with separate-coloured curves for each formulation, stratified by database and healthcare setting. Formulations would be dry powder, pressurized metered dose inhaler, solution for inhalation.

Figure 3. Monthly prescribing rates for i. inhaled salbutamol, ii. short acting Beta-2 agonists other than salbutamol, iii. short acting anticholinergic drugs, iv. combination of Beta-2 agonists and anticholinergics, v. Formoterol, vi. Formoterol + ICS, vii. salbutamol + ICS, viii. oral salbutamol, displaying by eight separate-coloured curves, provided by each database and healthcare setting.

Table 1. Distribution of predefined indication of use, age and sex distribution for individuals at time of each prescribing of inhaled salbutamol during study period, stratified by pharmaceutical formulations, and separately reported by database and healthcare setting.

One separate table will be provided for each data source and healthcare setting stratum. Each table includes a column for each inhaled salbutamol formulation, with cells in each column providing the following estimates: the number of prescriptions, the number of unique patients, the number (%) of males, the number (%) of females, the median age (IQR), the number (%) by age group, and the number (%) for each indication.

Interactive dashboard will be generated by incorporating all the results (tables and figures) included in the pdf report mentioned above.

8.9 Evidence synthesis

Results from analyses described in section 8.8.5 will be presented separately for each database and healthcare setting and no meta-analysis of results will be conducted to pool results over databases.

9. DATA MANAGEMENT

9.1 Data management

All databases have previously mapped their data to the OMOP common data model. This enables the use of standardised analytics and using DARWIN EU tools across the network since the structure of the data and the terminology system is harmonised. The OMOP CDM was developed and maintained by the Observational Health Data Sciences and Informatics (OHDSI) initiative and is described in detail on the wiki

page of the CDM: https://ohdsi.github.io/CommonDataModel and in The Book of OHDSI: http://book.ohdsi.org.

The analytic code for this study will be written in R and will use standardized analytics wherever possible. Each data partner will execute the study code against their database containing patient-level data, and then return the results (csv files) which will only contain aggregated data. The results from each of the contributing data sites will then be combined in tables and figures for the study report.

9.2 Data storage and protection

For this study, participants from various EU member states will process personal data from individuals which is collected in national/regional electronic health record databases. Due to the sensitive nature of this personal medical data, it is important to be fully aware of ethical and regulatory aspects and to strive to take all reasonable measures to ensure compliance with ethical and regulatory issues on privacy.

All databases used in this study are already used for pharmaco-epidemiological research and have a well-developed mechanism to ensure that European and local regulations dealing with ethical use of the data and adequate privacy control are adhered to. In agreement with these regulations, rather than combining person-level data and performing only a central analysis, local analyses will be run, which generate non-identifiable aggregate summary results.

The output files are stored in the DARWIN Remote Research Environment. These output files do not contain any data that allow identification of subjects included in the study. The RRE implements further security measures to ensure a high level of stored data protection to comply with the local implementation of the General Data Protection Regulation (GDPR) (EU) 679/20161 in the various member states.

10. QUALITY CONTROL

General database quality control

A number of open-source quality control mechanisms for the OMOP CDM have been developed (see Chapter 15 of The Book of OHDSI http://book.ohdsi.org/DataQuality.html). In particular, it is expected that data partners will have run the OHDSI Data Quality Dashboard tool (https://github.com/OHDSI/DataQualityDashboard). This tool provides numerous checks relating to the conformance, completeness and plausibility of the mapped data. Conformance focuses on checks that describe the compliance of the representation of data against internal or external formatting, relational, or computational definitions, completeness in the sense of data quality is solely focused on quantifying missingness, or the absence of data, while plausibility seeks to determine the believability or truthfulness of data values. Each of these categories has one or more subcategories and are evaluated in two contexts: validation and verification. Validation relates to how well data align with external benchmarks with expectations derived from known true standards, while verification relates to how well data conform to local knowledge, metadata descriptions, and system assumptions.

Study specific quality control

When defining drug cohorts, non-systemic products will be excluded from the list of included codes summarised on the ingredient level. A pharmacist will review the codes of the <drug(s)> of interest.
When defining cohorts for indications, a systematic search of possible codes for inclusion will be identified using CodelistGenerator R package (https://github.com/darwin-eu/CodelistGenerator). This software allows the user to define a search strategy and using this will then query the vocabulary tables of the OMOP common data model so as to find potentially relevant codes. In addition, the CohortDiagnostics R

package (https://github.com/OHDSI/CohortDiagnostics) will be run if needed to assess the use of different codes across the databases contributing to the study and identify any codes potentially omitted in error. The study code will be based on two R packages currently being developed to (1) estimate Incidence and Prevalence and (2) characterise drug utilisation using the OMOP common data model. These packages will include numerous automated unit tests to ensure the validity of the codes, alongside software peer review and user testing. The R package will be made publicly available via GitHub.

11. LIMITATIONS OF THE RESEARCH METHODS

The study will be informed by routinely collected health care data and some considerations should be mentioned. There is variability in database setting, notably the differences between primary care database and hospital databases and absence of linked data that restricts the ability to establish comprehensive connections between various patient records and outcomes.

Regarding inhaled salbutamol or alternatives use, the recording of the prescription of the drug (e.g. concept id at clinical drug level or at ingredient level), and the prescriptions themselves may vary across databases. In addition, a recording of a prescription or dispensation does not mean that the patient actually took the drug.

For this study, the indication of use is important and predefined indication of use have been identified. Not all of these indications of interest might be present within the database (i.e. in DP with less granular source coding.

In addition, the quality of condition recording used for identification of the indication might vary across the databases included in this study. Indications are not linked to prescriptions at the time they are issued or dispensed in the databases. Validation of code lists by individual patient health record review will not be performed and is not possible in view of Off the shelf studies

Additionally, the results estimated from this study will only reflect the populations from the included data sources. Also, electronic health records have certain inherent limitations because they were collected for clinical purpose rather than primarily for research use.

12. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

Adverse events/adverse reactions will not be collected or analysed as part of this evaluation. The nature of this non-interventional evaluation, through the use of secondary data, does not fulfil the criteria for reporting adverse events, according to module VI, VI.C.1.2.1.2 of the Good Pharmacovigilance Practices (https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guideline-good-pharmacovigilance-practices-gvp-module-vi-collection-management-submission-reports_en.pdf).

Only in case of prospective data collection, there is a need to describe the procedures for the collection, management and reporting of individual cases of adverse events/adverse reactions.

13. GOVERNANCE BOARD ASPECTS

All data sources require approval from their respective IRB board, with the exception of IQVIA DA Germany and DK-DHR Denmark which will not require any further specific approvals to undertake this study.

14. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

A PDF report including an executive summary, and the specified tables and/or figures will be submitted to EMA by the DARWIN EU® CC upon completion of the study.

An interactive dashboard incorporating all the results (tables and figures) will be provided alongside the PDF report. The full set of underlying aggregated data used in the dashboard will also be made available if requested.

15. OTHER ASPECTS

None.

16. REFERENCES

1. Network GA. Global asthma report. 2022.
2. Gibson G, Loddenkemper R, Sibille Y, Lundbäck B. The European Lung White Book. Respiratory Health and Disease in Europe. Sheffield, European Respiratory Society, 2013.
3. Agency EM. Medicine shortages in the EU: causes and solutions. 2021.
4. Agency EM. Shortage of salbutamol inhalation products. 2024.
5. Labaki WW, Han MK. Chronic respiratory diseases: a global view. The Lancet Respiratory Medicine. 2020;8(6):531-3.
6. Matera MG, Page CP, Calzetta L, Rogliani P, Cazzola M. Pharmacology and Therapeutics of Bronchodilators Revisited. Pharmacol Rev. 2020;72(1):218-52.
7. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). Int J Epidemiol. 2015;44(3):827-36.
8. Carey IM, Nirmalananthan N, Harris T, DeWilde S, Chaudhry UAR, Limb E, et al. Prevalence of co-morbidity and history of recent infection in patients with neuromuscular disease: A cross-sectional analysis of United Kingdom primary care data. PLoS One. 2023;18(3):e0282513.
9. Fahmi A, Wong D, Walker L, Buchan I, Pirmohamed M, Sharma A, et al. Combinations of medicines in patients with polypharmacy aged 65-100 in primary care: Large variability in risks of adverse drug related and emergency hospital admissions. PLoS One. 2023;18(2):e0281466.
10. Wigglesworth S, Neligan A, Dickson JM, Pullen A, Yelland E, Anjuman T, et al. The incidence and prevalence of epilepsy in the United Kingdom 2013-2018: A retrospective cohort study of UK primary care data. Seizure. 2023;105:37-42.
11. de Ridder MAJ, de Wilde M, de Ben C, Leyba AR, Mosseveld BMT, Verhamme KMC, et al. Data Resource Profile: The Integrated Primary Care Information (IPCI) database, The Netherlands. Int J Epidemiol. 2022;51(6):e314-e23.
12. Ali MS, Berencsi K, Marinier K, Deltour N, Perez-Guthann S, Pedersen L, et al. Comparative cardiovascular safety of strontium ranelate and bisphosphonates: a multi-database study in 5 EU countries by the EU-ADR Alliance. Osteoporos Int. 2020;31(12):2425-38.
13. Baan EJ, van den Akker ELT, Engelkes M, de Rijke YB, de Jongste JC, Sturkenboom M, et al. Hair cortisol and inhaled corticosteroid use in asthmatic children. Pediatr Pulmonol. 2020;55(2):316-21.
14. Berencsi K, Sami A, Ali MS, Marinier K, Deltour N, Perez-Gutthann S, et al. Impact of risk minimisation measures on the use of strontium ranelate in Europe: a multi-national cohort study in 5 EU countries by the EU-ADR Alliance. Osteoporos Int. 2020;31(4):721-55.

15. Engelkes M, Baan EJ, de Ridder MAJ, Svensson E, Prieto-Alhambra D, Lapi F, et al. Incidence, risk factors and re-exacerbation rate of severe asthma exacerbations in a multinational, multidatabase pediatric cohort study. Pediatr Allergy Immunol. 2020;31(5):496-505.
16. James G, Collin E, Lawrance M, Mueller A, Podhorna J, Zaremba-Pechmann L, et al. Treatment pathway analysis of newly diagnosed dementia patients in four electronic health record databases in Europe. Soc Psychiatry Psychiatr Epidemiol. 2021;56(3):409-16.
17. Rathmann W, Bongaerts B, Carius HJ, Kruppert S, Kostev K. Basic characteristics and representativeness of the German Disease Analyzer database Int J Clin Pharmacol Ther. 2018;56(10):459-66.
18. Zappacosta S, Cascarano A, Konrad M, Tanislav C, Kostev K. Association between COVID-19 and subsequent vascular events in primary care patients in Germany. Public Health. 2022;213:107-13.
19. Ly NF, Flach C, Lysen TS, Markov E, van Ballegooijen H, Rijnbeek P, et al. Impact of European Union Label Changes for Fluoroquinolone-Containing Medicinal Products for Systemic and Inhalation Use: Post-Referral Prescribing Trends. Drug Saf. 2023;46(4):405-16.
20. Tanislav C, Rosenbauer J, Zingel R, Kostev K. No increased incidence of venous thrombosis or pulmonary embolism after SARS-CoV-2 vaccination in Germany. Public Health. 2022;207:14-8.
21. Recalde M, Rodríguez C, Burn E, Far M, García D, Carrere-Molina J, et al. Data Resource Profile: The Information System for Research in Primary Care (SIDIAP). Int J Epidemiol. 2022;51(6):e324-e36.
22. Braeye T, Emborg HD, Llorente-García A, Huerta C, Martín-Merino E, Duarte-Salles T, et al. Age-specific vaccination coverage estimates for influenza, human papillomavirus and measles containing vaccines from seven population-based healthcare databases from four EU countries - The ADVANCE project. Vaccine. 2020;38(16):3243-54.
23. Burn E, Tebé C, Fernandez-Bertolin S, Aragon M, Recalde M, Roel E, et al. The natural history of symptomatic COVID-19 during the first wave in Catalonia. Nat Commun. 2021;12(1):777.
24. Lane JC, Butler KL, Poveda-Marina JL, Martinez-Laguna D, Reyes C, de Bont J, et al. Preschool Obesity Is Associated With an Increased Risk of Childhood Fracture: A Longitudinal Cohort Study of 466,997 Children and Up to 11 Years of Follow-up in Catalonia, Spain. J Bone Miner Res. 2020;35(6):1022-30.
25. Monteagudo M, Nuñez A, Solntseva I, Dhalwani N, Booth A, Barrecheguren M, et al. Treatment Pathways Before and After Triple Therapy in COPD: A Population-based Study in Primary Care in Spain. Arch Bronconeumol (Engl Ed). 2021;57(3):205-13.
26. Ortega Y, Aragonès E, Piñol JL, Basora J, Araujo A, Cabré JJ. Impact of depression and/or anxiety on the presentation of cardiovascular events in a cohort with metabolic syndrome. StreX project: Five years of follow-up. Prim Care Diabetes. 2018;12(2):163-71.
27. Ramos R, Comas-Cufí M, Martí-Lluch R, Balló E, Ponjoan A, Alves-Cabratosa L, et al. Statins for primary prevention of cardiovascular events and mortality in old and very old adults with and without type 2 diabetes: retrospective cohort study. Bmj. 2018;362:k3359.
28. Recalde M, Davila-Batista V, Díaz Y, Leitzmann M, Romieu I, Freisling H, et al. Body mass index and waist circumference in relation to the risk of 26 types of cancer: a prospective cohort study of 3.5 million adults in Spain. BMC Med. 2021;19(1):10.
29. Troncoso-Mariño A, López-Jiménez T, Roso-Llorach A, Villén N, Amado-Guirado E, Guisado-Clavero M, et al. Medication-related problems in older people in Catalonia: A real-world data study. Pharmacoepidemiol Drug Saf. 2021;30(2):220-8.
30. Eea B. IncidencePrevalence: Estimate Incidence and Prevalence using the OMOP Common Data Model 2023 [Available from: https://cran.r-project.org/web/packages/IncidencePrevalence/index.html]. 2023.

17. ANNEXES

Appendix I: List of the concept definitions of drugs and conditions are provided below. Concept IDs might be updated based on results from drug exposure diagnostics.

List of medicines concepts definition is provided in the **Table S 1**.

Table S 1. Preliminary list of medicines definitions.

Substance Name	Concept name	Concept ID including descendants	Vocabulary
Salbutamol	Albuterol	1154343	RxNorm
	Albuterol inhalation solution	1356108	RxNorm
	Albuterol inhalation powder	1356111	RxNorm
	Albuterol metered dose inhaler	40142703	RxNorm
	Albuterol inhalation suspension	2070100	RxNorm
	Albuterol dry powder inhaler	46234463	RxNorm
	Albuterol powder for oral suspension	35154239	RxNorm
Terbutaline	Terbutaline	1236744	RxNorm
	Terbutaline dry powder inhaler	42481922	RxNorm
	Terbutaline inhalation powder	36813480	RxNorm
	Terbutaline inhalation solution	36812414	RxNorm
	Terbutaline metered dose inhaler	1356244	RxNorm
Fenoterol	Fenoterol	19053979	RxNorm
	Fenoterol dry powder inhaler	40861768	RxNorm
	Fenoterol inhalant powder	40727839	RxNorm
	Fenoterol inhalation solution	35150375	RxNorm
	Fenoterol metered dose inhaler	44081619	RxNorm
Ipratropium bromide	Ipratropium	1112921	RxNorm
	Ipratropium dry powder inhaler	42483253	RxNorm
	Ipratropium gas for Inhalation	44081549	RxNorm
	Ipratropium inhalation powder	36811485	RxNorm
	Ipratropium inhalation solution	1356213	RxNorm
	Ipratropium metered dose inhaler	40143214	RxNorm

Oxitropium bromide	Oxitropium	19018882	RxNorm
	Oxitropium dry powder inhaler	40861610	RxNorm
	Oxitropium inhalation powder	36810762	RxNorm
	Oxitropium inhalation solution	35147090	RxNorm
Fenoterol + ipratropium bromide	Fenoterol/ ipratropium dry powder inhaler	43134418	RxNorm
	Fenoterol / ipratropium inhalant powder	40727834	RxNorm
	Fenoterol / ipratropium inhalation solution	36811735	RxNorm
Salbutamol + ipratropium bromide	Albuterol / ipratropium inhalant powder	40727741	RxNorm
	Albuterol / ipratropium inhalation solution	1356123	RxNorm
	Albuterol / ipratropium inhalation spray	37499261	RxNorm
	Albuterol / ipratropium metered dose inhaler	2072019	RxNorm
Formoterol	Formoterol	1196677	RxNorm
	Formoterol dry powder inhaler	42480849	RxNorm
	Formoterol inhalation powder	1356191	RxNorm
	Formoterol inhalation solution	1356187	RxNorm
	Formoterol metered dose inhaler	44107471	RxNorm
Formoterol + beclomethasone	Beclomethasone / formoterol dry powder inhaler	21158944	RxNorm
	Beclomethasone / formoterol inhalant powder	36894458	RxNorm
	Beclomethasone / formoterol inhalant solution	21090035	RxNorm
Formoterol + budesonide	Budesonide / formoterol dry powder inhaler	42479684	RxNorm
	Budesonide / formoterol inhalation Powder	35133500	RxNorm
	Budesonide / formoterol inhalation solution	783228	RxNorm
	Budesonide / formoterol inhalation suspension	2069097	RxNorm
	Budesonide / formoterol metered dose inhaler	40142910	RxNorm
Salbutamol + beclomethasone	Albuterol / beclomethasone dry powder inhaler	42483138	RxNorm

	Albuterol / beclomethasone inhalation powder	36812530	RxNorm
	Albuterol / beclomethasone Inhalation Solution	36965394	RxNorm
Beclomethasone	Beclomethasone	115572	RxNorm
Budesonide	Budesonide	939259	RxNorm
	Budesonide inhalant	43291282	RxNorm
	Budesonide metered dose inhaler	44120754	RxNorm
	Budesonide dry powder inhaler	40142920	RxNorm
	Budesonide inhalation powder	1356143	RxNorm
	Budesonide inhalation solution	35135829	RxNorm
	Budesonide inhalation suspension	1356140	RxNorm
Oral salbutamol	Albuterol extended-release oral capsule	40006960	RxNorm
	Albuterol extended-release oral tablet	40006962	RxNorm
	Albuterol oral capsule	41079955	RxNorm
	Albuterol oral solution	40104188	RxNorm
	Albuterol oral suspension	1235090	RxNorm
	Albuterol oral tablet	40006997	RxNorm
	Albuterol powder for oral suspension	35154239	RxNorm

Conditions concept name and concept IDs definition is listed in the **Table S 2**.

Table S 2. Preliminary list of conditions definitions.

Phenotype	Concept name	Concept id (including descendants)	Exclude concept id	Vocabulary
Asthma	Asthma	317009	-	SNOMED
COPD and Emphysema	Chronic obstructive lung disease	255573	-	SNOMED
Respiratory conditions due to inhalation of chemical substances	Respiratory condition caused by chemical fumes and vapors Respiratory condition caused by chemical fumes Chronic respiratory condition caused by chemical fumes	3655120 3655118 3655114	-	SNOMED

Bronchitis	Bronchitis	256451	-	SNOMED
Lower respiratory tract infection	Lower respiratory tract infection	4175297	-	SNOMED
Bronchospasm	Bronchospasm	256717	-	SNOMED

Appendix II: ENCePP checklist for study protocols

Doc.Ref. EMA/540136/2009

ENCePP Checklist for Study Protocols (Revision 4)

Adopted by the ENCePP Steering Group on 15/10/2018

Study title: DARWIN EU® - Drug Utilization of salbutamol products for inhalation and therapeutic alternative inhalation products

EU PAS Register® number: n/a
Study reference number (if applicable): n/a

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[1]	☒	☐	☐	5
1.1.2 End of data collection[2]	☒	☐	☐	
1.1.3 Progress report(s)	☐	☐	☒	
1.1.4 Interim report(s)	☐	☐	☒	
1.1.5 Registration in the EU PAS Register®	☒	☐	☐	
1.1.6 Final report of study results.	☒	☐	☐	

Comments:

[1] Date from which information on the first study is first recorded in the study dataset or, in the case of secondary use of data, the date from which data extraction starts.
[2] Date from which the analytical dataset is completely available.

Section 2: Research question		**Yes**	**No**	**N/A**	**Section Number**
2.1	Does the formulation of the research question and objectives clearly explain:	☒	☐	☐	6, 7
	2.1.1 Why the study is conducted? (e.g. to address an important public health concern, a risk identified in the risk management plan, an emerging safety issue)	☒	☐	☐	
	2.1.2 The objective(s) of the study?	☒	☐	☐	
	2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	
	2.1.4 Which hypothesis(-es) is (are) to be tested?	☐	☐	☒	
	2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☐	☒	

Comments:

Section 3: Study design		**Yes**	**No**	**N/A**	**Section Number**
3.1	Is the study design described? (e.g. cohort, case-control, cross-sectional, other design)	☒	☐	☐	8.1
3.2	Does the protocol specify whether the study is based on primary, secondary or combined data collection?	☒	☐	☐	8.2
3.3	Does the protocol specify measures of occurrence? (e.g., rate, risk, prevalence)	☒	☐	☐	8.8
3.4	Does the protocol specify measure(s) of association? (e.g. risk, odds ratio, excess risk, rate ratio, hazard ratio, risk/rate difference, number needed to harm (NNH))	☐	☐	☒	
3.5	Does the protocol describe the approach for the collection and reporting of adverse events/adverse reactions? (e.g. adverse events that will not be collected in case of primary data collection)	☐	☐	☒	

Comments:

Section 4: Source and study populations		**Yes**	**No**	**N/A**	**Section Number**
4.1	Is the source population described?	☒	☐	☐	8.2/8.5
4.2	Is the planned study population defined in terms of:				8.5
	4.2.1 Study time period	☒	☐	☐	
	4.2.2 Age and sex	☒	☐	☐	
	4.2.3 Country of origin	☒	☐	☐	
	4.2.4 Disease/indication	☒	☐	☐	
	4.2.5 Duration of follow-up	☒	☐	☐	

Section 4: Source and study populations		**Yes**	**No**	**N/A**	**Section Number**
4.3	Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	8.5

Comments:

Section 5: Exposure definition and measurement		**Yes**	**No**	**N/A**	**Section Number**
5.1	Does the protocol describe how the study exposure is defined and measured? (e.g. operational details for defining and categorising exposure, measurement of dose and duration of drug exposure)	☒	☐	☐	8.6
5.2	Does the protocol address the validity of the exposure measurement? (e.g. precision, accuracy, use of validation sub-study)	☐	☐	☒	
5.3	Is exposure categorised according to time windows?	☐	☐	☒	
5.4	Is intensity of exposure addressed? (e.g. dose, duration)	☒	☐	☐	
5.5	Is exposure categorised based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☐	☐	☒	
5.6	Is (are) (an) appropriate comparator(s) identified?	☐	☐	☒	

Comments:

Section 6: Outcome definition and measurement		**Yes**	**No**	**N/A**	**Section Number**
6.1	Does the protocol specify the primary and secondary (if applicable) outcome(s) to be investigated?	☐	☐	☒	8.6
6.2	Does the protocol describe how the outcomes are defined and measured?	☐	☐	☒	
6.3	Does the protocol address the validity of outcome measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, use of validation sub-study)	☐	☐	☒	
6.4	Does the protocol describe specific outcomes relevant for Health Technology Assessment? (e.g. HRQoL, QALYs, DALYS, health care services utilisation, burden of disease or treatment, compliance, disease management)	☐	☐	☒	

Comments:

Section 7: Bias		**Yes**	**No**	**N/A**	**Section Number**
7.1	Does the protocol address ways to measure confounding? (e.g. confounding by indication)	☐	☐	☒	
7.2	Does the protocol address selection bias? (e.g. healthy user/adherer bias)	☐	☐	☒	
7.3	Does the protocol address information bias? (e.g. misclassification of exposure and outcomes, time-related bias)	☐	☐	☒	

Comments:

Section 8: Effect measure modification		**Yes**	**No**	**N/A**	**Section Number**
8.1	Does the protocol address effect modifiers? (e.g. collection of data on known effect modifiers, sub-group analyses, anticipated direction of effect)	☐	☐	☒	

Comments:

Section 9: Data sources		**Yes**	**No**	**N/A**	**Section Number**
9.1	Does the protocol describe the data source(s) used in the study for the ascertainment of:				
	9.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview)	☒	☐	☐	8.6
	9.1.2 Outcomes? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	☐	☐	☒	8.6
	9.1.3 Covariates and other characteristics?	☒	☐	☐	8.6
9.2	Does the protocol describe the information available from the data source(s) on:				
	9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	8.6
	9.2.2 Outcomes? (e.g. date of occurrence, multiple event, severity measures related to event)	☐	☐	☒	8.6
	9.2.3 Covariates and other characteristics? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	☒	☐	☐	8.6
9.3	Is a coding system described for:				
	9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☒	☐	☐	8.6
	9.3.2 Outcomes? (e.g. International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	☐	☐	☒	8.6
	9.3.3 Covariates and other characteristics?	☒	☐	☐	8.6
9.4	Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☐	☐	☒	

Comments:

Section 10: Analysis plan		**Yes**	**No**	**N/A**	**Section Number**
10.1	Are the statistical methods and the reason for their choice described?	☒	☐	☐	8.8
10.2	Is study size and/or statistical precision estimated?	☐	☐	☒	8.7
10.3	Are descriptive analyses included?	☒	☐	☐	8.8
10.4	Are stratified analyses included?	☒	☐	☐	8.8
10.5	Does the plan describe methods for analytic control of confounding?	☐	☐	☒	
10.6	Does the plan describe methods for analytic control of outcome misclassification?	☐	☐	☒	
10.7	Does the plan describe methods for handling missing data?	☐	☐	☒	
10.8	Are relevant sensitivity analyses described?	☐	☐	☒	

Comments:

Section 11: Data management and quality control		**Yes**	**No**	**N/A**	**Section Number**
11.1	Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	☒	☐	☐	9.2
11.2	Are methods of quality assurance described?	☒	☐	☐	10.0
11.3	Is there a system in place for independent review of study results?	☐	☐	☒	

Comments:

Section 12: Limitations		**Yes**	**No**	**N/A**	**Section Number**
12.1	Does the protocol discuss the impact on the study results of:				
	12.1.1 Selection bias?	☒	☐	☐	11
	12.1.2 Information bias?	☒	☐	☐	
	12.1.3 Residual/unmeasured confounding? (e.g. anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods).	☐	☐	☒	
12.2	Does the protocol discuss study feasibility? (e.g. study size, anticipated exposure uptake, duration of follow-up in a cohort study, patient recruitment, precision of the estimates)	☒	☐	☐	8.2

Comments:

Section 13: Ethical/data protection issues	**Yes**	**No**	**N/A**	**Section Number**
13.1 Have requirements of Ethics Committee/ Institutional Review Board been described?	☒	☐	☐	13
13.2 Has any outcome of an ethical review procedure been addressed?	☒	☐	☐	
13.3 Have data protection requirements been described?	☒	☐	☐	9.2

Comments:

Section 14: Amendments and deviations	**Yes**	**No**	**N/A**	**Section Number**
14.1 Does the protocol include a section to document amendments and deviations?	☒	☐	☐	4

Comments:

Section 15: Plans for communication of study results	**Yes**	**No**	**N/A**	**Section Number**
15.1 Are plans described for communicating study results (e.g. to regulatory authorities)?	☒	☐	☐	14
15.2 Are plans described for disseminating study results externally, including publication?	☒	☐	☐	14

Comments:

Name of the main author of the protocol: Marzyeh Amini

Date: 02/12/1976

Signature: Marzyeh Amini