Study Protocol P2 C1-003

08/08/2023

Version 2.1

TABLE OF CONTENTS

DOCUMENT HISTORY

Version	Date	Description
V1.0	23rd June 2023	**Submission to EMA**
V2.0	21th July 2023	**Second version following comments from EMA**
V2.1	8 August 2023	**EUPAS register number added**

Study Title	DARWIN EU® - Co-prescribing of endothelin receptor antagonists (ERAs) and phosphodiesterate-5 inhibitors (PDE-5is) in pulmonary arterial hypertension (PAH)
Protocol version identifier	V2.1
Date of last version of protocol	8 August 2023
EU PAS register number	EUPAS106052
Active substance	

Drugs	Class	ATC code
Ambrisentan	Endothelin receptor antagonist	C02KX02
Bosentan	Endothelin receptor antagonist	C02KX01
Macitentan	Endothelin receptor antagonist	C02KX04
Sitaxentan	Endothelin receptor antagonist	C02KX03
Sildenafil	Phosphodiesterate-5 inhibitors	G04BE03
Tadalafil	Phosphodiesterate-5 inhibitors	G04BE08

Medicinal product	N/A
Research question and objectives	Research question What is the utilization pattern of endothelin receptor antagonists (ERAs) and phosphodiesterase-5 inhibitors (PDE-5is) in pulmonary arterial hypertension (PAH)? Study objectives Objective 1: To estimate proportions of patients with newly diagnosed pulmonary arterial hypertension (PAH) who initiate treatment with endothelin receptor antagonists (ERAs) or phosphodiesterase-5 inhibitors (PDE-5is), either as monotherapy or in combination, during the period from January 1, 2012, to December 31, 2022. Objective 2: To estimate the duration of prescription for ERAs and PDE-5is in patients with newly diagnosed PAH between January 1, 2012, and December 31, 2022. Objective 3: To describe the prescription patterns and sequences of ERAs and PDE-5is in patients with newly diagnosed PAH between January 1, 2012, and December 31, 2022. Objective 4: To estimate the proportion of patients with newly diagnosed PAH who experience specific events of interest, namely cardiovascular hospitalization, all-cause hospitalization, and death, after initiating treatment with ERAs and PDE-5is between January 1, 2012, and December 31, 2022.

Countries of study	Estonia, France, Germany, and the United Kingdom.
Author	Johnmary Arinze (j.arinze@darwin-eu.org) Katia Verhamme (k.verhamme@darwin-eu.org)

LIST OF ABBREVIATIONS

Acronyms/terms	Description
CDM	Common Data Model
CHUBX	Bordeaux University Hospital
CPRD GOLD	Clinical Practice Research Datalink GOLD
DA	Disease Analyzer
DARWIN EU®	Data Analysis and Real World Interrogation Network
DUS	Drug Utilization Study
EBB	Estonian Biobank
EGCUT	Estonian Genome Center at the University of Tartu
EHR	Electronic Healthcare Records
EMA	European Medicines Agency
ERAs	Endothelin receptor antagonists
GP	General Practitioner
ID	Index date
OHDSI	Observational Health Data Sciences and Informatics
OMOP	Observational Medical Outcomes Partnership
PAH	Pulmonary arterial hypertension
PDE-5is	Phosphodiesterase-5 inhibitors

1. TITLE

DARWIN EU® - Co-prescribing of endothelin receptor antagonists (ERAs) and phosphodiesterate-5 inhibitors (PDE-5is) in pulmonary arterial hypertension (PAH)

2. RESPONSIBLE PARTIES – STUDY TEAM

Table 1 shows a description of the Study team by role, name and organization.

Table 1: Description of Study Team

Study team Role	Names	Organisation
Principal Investigator(s)/ Clinical Epidemiologists	Johnmary Arinze	Erasmus MC
	Katia Verhamme	Erasmus MC
Data Partner*	**Names**	**Organization**
Local Study Coordinator/Data Analyst	Antonella Delmestri	University of Oxford – CPRD data
	James Brash	IQVIA DA Germany
	Vianney Jouhet	CHUBX France
	Raivo Kolde	University of Tartu - Estonian Biobank

*Data partners' role is only to execute code at their data source. These people do not have an investigator role.

3. ABSTRACT (STAND ALONE SUMMARY OF THE STUDY PROTOCOL)

Title

DARWIN EU® - Co-prescribing of endothelin receptor antagonists (ERAs) and phosphodiesterate-5 inhibitors (PDE-5is) in pulmonary arterial hypertension (PAH)

Rationale and Background

Pulmonary Arterial Hypertension (PAH) is a rare type of pulmonary hypertension characterized by increased mean pulmonary arterial pressure. PAH can be idiopathic, heritable, drug-induced, or secondary to other chronic diseases such as HIV infection, congenital heart diseases, portal hypertension. Median survival is about 7 years from the time of diagnostic catheterisation. The therapeutic management differs based on type and severity and usually involves a mono or combination therapy from the following agents: Endothelin receptor antagonists (ERAs), and Phosphodiesterase-5 inhibitors (PDE5-is). There is interest in understanding how these therapies are used in clinical practice to contextualise assessments of potential future development programs in this indication.

Research question and Objectives

Research question:

To understand how PAH is treated in clinical practice. Specific objectives of this study are listed below.

Study objectives:

(1) To estimate proportions of patients with newly diagnosed pulmonary arterial hypertension (PAH) who initiate treatment with endothelin receptor antagonists (ERAs) or phosphodiesterase-5 inhibitors (PDE-5is), either as monotherapy or in combination, during the period from January 1, 2012, to December 31, 2022.

(2) To estimate the duration of prescription for ERAs and PDE-5is in patients with newly diagnosed PAH between January 1, 2012, and December 31, 2022.

(3) To describe the prescription patterns and sequences of ERAs and PDE-5is in patients with newly diagnosed PAH between January 1, 2012, and December 31, 2022.

(4) To estimate the proportion of patients with newly diagnosed PAH who experience specific events of interest, namely cardiovascular hospitalization, all-cause hospitalization, and death, after initiating treatment with ERAs and PDE-5is between January 1, 2012, and December 31, 2022.

Research Methods

Study design

Retrospective cohort study in patients with newly diagnosed PAH.

Population

Patient-level characterization – Disease epidemiology study: The patient-level characterization will involve patients with newly diagnosed PAH between January 1, 2012, and December 31, 2022 (or the latest available

date if earlier), with at least 1 year of data availability prior to their diagnosis, and no record of being diagnosed with PAH in the previous year. To investigate treatment patterns, a minimum follow-up time of 30 days will be applied to capture PAH treatment (objective 1), and treatment sequences (objective 3) .

Patient-level drug utilization: New users of ERAs/PDE-5is in patients newly diagnosed with PAH in the period between January 1, 2012, and December 31, 2022 (or latest date available), with at least 1 year of data visibility prior to index date, and no use of the respective ERAs/PDE-5is in the previous 1 year, will be included for patient-level drug utilisation analyses. Therefore, children aged <1 year will be excluded.

Variables

Drug of interest: They will be identified through RxNorm drug codes

1. Endothelin receptor antagonists: Ambrisentan, Bosentan, Macitentan, and Sitaxentan.
2. Phosphodiesterate-5 inhibitors: Sildenafil and Tadalafil.

Condition of interest: PAH which will be identified through SNOMED disease codes.

Outcomes of interest: Study outcomes will be identified based on the presence of events of cardiovascular hospitalisation, all-cause hospitalisation, and death.

Data sources

1. Clinical Data Warehouse of Bordeaux University Hospital (CHUBX), France
2. Clinical Practice Research Datalink GOLD (CPRD GOLD), United Kingdom
3. Estonian Biobank (EBB), Estonia
4. IQVIA Disease Analyzer Germany (IQVIA DA Germany), Germany

Sample size

Sample size was not calculated for this Disease Epidemiology and Drug Utilization Study, as our primary focus is to examine the characteristics and treatment patterns of all incident PAH patients, irrespective of the sample size. Based on a preliminary feasibility assessment, the expected number of PAH records in the included databases for this study will be approximately 35,059.

Data analyses

The number and % of patients receiving each of a pre-specified list of PAH treatments (objective 1) and treatment combinations (objectives 1 and 3) will be described at index date, 1 to 30, 1 to 90, and 1 to 365 days post index date. A treatment pattern analysis will be conducted to describe the sequence of prescribing of the specific ERAs/PDE-5is following diagnosis. Index date will be the date of diagnosis of PAH. Sunburst plots and Sankey diagrams will be used to describe treatment patterns and sequences over time (objective 3).

Large-scale patient-level characterisation will be conducted to describe age and sex at time of PAH diagnosis. In addition, medical history will be assessed for anytime –and up to 366 days before index date and for 365 to index date. The medical history will include clinical symptoms and signs (chest pain, dyspnea, fatigue, syncope), comorbidities (obesity, congenital heart disease, heart failure, pulmonary embolism, chronic obstructive pulmonary disease, pulmonary fibrosis, obstructive sleep apnea and chronic kidney disease), and important factors possibly related to PAH diagnosis (idiopathic, heritable, connective tissue disease,

corrected congenital shunts, drug or toxin induced PAH, and others). We will also report the proportion of patients with outcomes of interest for 1, 3, and 5 years post index date (objective 4).

Patient-level ERAs/PDE-5is use: Patient-level features will be characterized, and the treatment duration of prescriptions of the respective drugs of interest will be estimated. Index date will be the date of the first prescription of the specific ERAs/ PDE-5is for each person. Treatment duration will be estimated for the first treatment era and the minimum, p25, median, p75, and maximum will be provided.

For all analyses a minimum cell count of 5 will be used when reporting results, with any smaller counts obscured.

4. AMENDMENTS AND UPDATES

Number	Date	Section of study protocol	Amendment or update	Reason
Version 2.0	21/07/2023	All	Update	Update following EMA's assessment
Version 2.1	08/08/2023	Document history	Update	EUPAS register number added

5. MILESTONES

STUDY SPECIFIC DELIVERABLE	TIMELINE
Draft Study Protocol	23rd June 2023
Final Study Protocol	21st July 2023
Creation of Analytical code	August 2023
Execution of Analytical Code on the data	September 2023
Interim Study Report (if applicable)	Not applicable
Draft Study Report	13th October 2023
Final Study Report	To be confirmed

6. RATIONALE AND BACKGROUND

Pulmonary arterial hypertension (PAH) is a rare, chronic, and severe medical condition characterized by elevated pulmonary arterial pressure and vascular resistance[1]. Its global prevalence ranges from 0.4 to 1.4 cases per 100,000 persons, with significant regional variations.[2] Idiopathic PAH is more common in North America and Europe, while acquired PAH is prevalent in Africa and Asia. PAH primarily affects females, especially in idiopathic and heritable cases, and can occur across all age groups[2]. Patients with PAH experience progressive dyspnea, fatigue, exercise intolerance, impaired functional capacity, and reduced quality of life, imposing a substantial clinical burden.[3] Complications such as right heart failure, arrhythmias, and thromboembolic events further contribute to the disease burden.[4] The median survival time from the time of diagnosis is approximately 6 – 9 years.[5,6]

The etiology of PAH is multifactorial and involves various risk factors and underlying conditions, including genetic mutations, exposure to drugs or toxins, connective tissue diseases, congenital heart defects, HIV infection, portal hypertension, and chronic liver disease.[7,8] The pathophysiology of PAH involves complex mechanisms such as vasoconstriction, vascular remodeling, inflammation, and endothelial dysfunction. Endothelial injury and dysfunction lead to impaired production of vasodilators and increased secretion of vasoconstrictors, while pulmonary vascular smooth muscle cells undergo hypertrophy and proliferation, resulting in elevated vascular resistance.[1,8]

Pharmacotherapy plays a critical role in managing PAH, aiming to improve symptoms, exercise capacity, and hemodynamics.[9] Commonly used medications include endothelin receptor antagonists (ERAs) like ambrisentan, bosentan, macitentan, sitaxentan, as well as phosphodiesterase-5 inhibitors (PDE-5Is) such as sildenafil and tadalafil. Prostacyclin analogues and calcium channel blockers have also demonstrated positive therapeutic outcomes.[8,10] The choice of therapy depends on the type and severity of PAH, and both monotherapy and combination therapy approaches are employed to achieve optimal outcomes.[11]

Despite therapeutic advancements, the prognosis of PAH remains poor, highlighting the need to understand how specific therapies are utilized in clinical practice.[11] A comprehensive evaluation of current pharmacological approaches in managing PAH could provide valuable insights for future development programs in this indication.

This study aims to provide valuable insights into the prescription patterns of ERAs/ PDE-5Is, user characteristics, and clinical outcomes among patients with pulmonary arterial hypertension in Europe. It will offer crucial context for effective post-marketing surveillance and evaluation of the real-world benefit-risk profiles of these medications in managing pulmonary arterial hypertension.

7. RESEARCH QUESTION AND OBJECTIVES

Research question:

To understand how PAH is treated in clinical practice. Specific objectives of this study are listed below.

Study objectives:

(1) To estimate proportions of patients with newly diagnosed pulmonary arterial hypertension (PAH) who initiate treatment with endothelin receptor antagonists (ERAs) or phosphodiesterase-5 inhibitors (PDE-5is), either as monotherapy or in combination, during the period from January 1, 2012, to December 31, 2022.

(2) To estimate the duration of prescription for ERAs and PDE-5is in patients with newly diagnosed PAH between January 1, 2012, and December 31, 2022.

(3) To describe the prescription patterns and sequences of ERAs and PDE-5is in patients with newly diagnosed PAH between January 1, 2012, and December 31, 2022.

(4) To estimate the proportion of patients with newly diagnosed PAH who experience specific events of interest, namely cardiovascular hospitalization, all-cause hospitalization, and death, after initiating treatment with ERAs and PDE-5is between January 1, 2012, and December 31, 2022.

Table 1: Primary and secondary research questions and objective

Objective:	To describe the patient characteristics, treatment patterns of ERAs/PDE-5is, and the proportion of patients with newly diagnosed PAH who experience the events of interest (cardiovascular hospitalization, all-cause hospitalization, and death) after initiating treatment with ERAs and PDE-5is, stratified by country/database.
Hypothesis:	Not applicable
Population ***(mention key inclusion-exclusion criteria):***	All patients with a first diagnosis of PAH identified in the database between January 1, 2012, and December 31, 2022 (or the latest available date if earlier), with at least 1 year of data availability prior to their diagnosis, and no record of being diagnosed with PAH in the previous year. Therefore, children aged <1 year will be excluded.
Exposure:	1. Endothelin receptor antagonists: Ambrisentan, Bosentan, Macitentan, and Sitaxentan. 2. Phosphodiesterate-5 inhibitors: Sildenafil and Tadalafil.
Comparator:	None
Outcome:	Cardiovascular hospitalisation, all-cause hospitalisation, and death
Time ***(when follow up begins and ends):***	Follow-up will start on the date of treatment initiation (index date) until the earliest of loss to follow-up, end of data availability or death, or end of study period (31st December 2022).
Setting:	Inpatient and outpatient setting using data from the following data sources: CPRD GOLD (UK), IQVIA DA (Germany), CHUBX (France), and Estonian Biobank (Estonia).
Main measure of effect:	Proportions, patient level drug utilisation to assess duration of use of the first exposure episode of the respective drugs of interest.

8. RESEARCH METHODS

8.1 Study Design

This will be a **patient-level characterisation** study classified as "off-the-shelf" (C1) and as described in the DARWIN EU® Complete Catalogue of Standard Data Analyses. A retrospective cohort study of all patients with incident PAH will be conducted.

A new drug user cohort will be used **(Patient Level DUS)** to characterise patient-level ERAs/ PDE-5is utilisation (treatment duration and patient characteristics of patients exposed to the drug of interest).

Table 2. Description of Potential Study Types and Related Study Designs

STUDY TYPE	STUDY DESIGN	STUDY CLASSIFICATION
Disease epidemiology - Patient-level characterisation	Cohort analysis	Off-the-shelf (C1)
Patient Level DUS	New drug/s user cohort	Off the shelf (C1)

8.2 Study Setting and data sources

This study will be conducted using routinely collected data from 4 databases in 4 European countries (3 EU countries and United Kingdom). All databases were previously mapped to the OMOP CDM .

1. Clinical Data Warehouse of Bordeaux University Hospital (CHUBX), France
2. Clinical Practice Research Datalink GOLD (CPRD GOLD), United Kingdom
3. Estonian Biobank (EBB), Estonia
4. IQVIA Disease Analyzer Germany (IQVIA DA Germany), Germany

For this study, we have carefully selected four databases from the ten databases available for DARWIN EU® in 2022. The selection process was based on two primary criteria: data reliability and relevance to the research question at hand. These selected databases demonstrate substantial record counts for both PAH and the drugs of interest. Moreover, they offer a good geographical spread, ensuring representation from diverse regions of Europe.

These chosen databases meet the requirements for conducting a patient-level drug utilization study as well as a patient-level characterization study, enabling us to explore PAH treatment and perform large-scale characterizations. Additionally, by including databases from different settings, we can effectively capture both inpatient and outpatient drug prescriptions for PAH, enhancing the comprehensiveness of our findings. We anticipate obtaining complete treatment data, which is necessary for objectives 1, 2, and 3, from all selected databases. Additionally, outcome data, required for objective 4, will also be available in most databases. However, it is important to note that the IQVIA DA Germany database lacks information regarding the date of death.

Information on the data source(s) with a justification for their choice in terms of ability to capture the relevant data is described in a **Table 3**.

Table 3: Description of data sources

Country	Name of Database	Justification for Inclusion	Health Care setting	Type of Data	Number of active subjects		Data lock for the last update
France	CHUBX	Database covers hospital care setting where PAH treatment may be initiated	Secondary care (in and outpatients)	EHR	2.1 million		05/05/2023
UK	CPRD GOLD	Database covers primary care setting where ERAs/ PDE-5is prescriptions are issued. Also research on PAH has already been conducted using CPRD data.[1]	Primary care	EHR	3 million		20/03/2023
Estonia	EBB	Database covers information from primary care and secondary care setting (insurance claims, digital prescriptions) where ERAs/ PDE-5is prescriptions are issued.	Biobank	Claims data	0.2 million		20/03/2023
Germany	IQVIA DA Germany	Database covers primary care and secondary care setting (outpatient specialist care) where ERAs/ PDE-5is prescriptions are issued. Research on PH (including PAH) has been conducted using IQVIA	Primary care and outpatient specialist care	EHR	8.5 million		13/03/2023

		DA Germany data.[1]					

PAH = Pulmonary arterial hypertension, CHUBX= Bordeaux University Hospital, UK = United Kingdom, CPRD GOLD = Clinical Practice Research Datalink GOLD, EBB = Estonian Biobank, DA = Disease Analyzer, EHR = Electronic Heath record. Exposure is based on prescription data.

[1] https://doi.org/10.1002/pul2.12000

Bordeaux University Hospital (CHUBX), France

The clinical data warehouse of the Bordeaux University Hospital comprises electronic health records on more than 2 million patients with data collection starting in 2005. The hospital complex is made up of three main sites and comprises a total of 3,041 beds (2021 figures). The database currently holds information about the person (demographics), visits (inpatient and outpatient), conditions and procedures (billing codes), drugs (outpatient prescriptions and inpatient orders and administrations), measurements (laboratory tests and vital signs) and dates of death (in or out-hospital death).[12]

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

The Clinical Practice Research Datalink (CPRD) is a governmental, not-for-profit research service, jointly funded by the National Institute for Health and Care Research and the Medicines and Healthcare products Regulatory Agency, a part of the Department of Health, United Kingdom (UK) (https://cprd.com). CPRD GOLD[13] comprises computerized records of all clinical and referral events in primary care in addition to comprehensive demographic information and medication prescription data in a sample of UK patients (predominantly from Scotland (52% of practices) and Wales (28% of practices). The prescription records include information on the type of product, date of prescription, strength, dosage, quantity, and route of administration. Data from contributing practices are collected and processed into research databases. Quality checks on patient and practice level are applied during the initial processing. Data are available for 20 million patients, including 3.2 million currently registered patients.

Access to CPRD GOLD data requires approval via the Research Data Governance Process. CPRD GOLD will be onboarded as Data Partner for DARWIN EU® in 2023.

Estonian Biobank – University of Tartu (Estonia)

The Estonian Biobank (EBB) is a population-based biobank of the Estonian Genome Center at the University of Tartu (EGCUT). Its cohort size is currently close to 200,000 participants (“gene donors” >= 18 years of age) which closely reflects the age, sex and geographical distribution of the Estonian adult population. Genomic GWAS analysis have been performed on all gene donors. The database also covers health insurance claims, digital prescriptions, discharge reports, information about incident cancer cases and causes of death from national sources for each donor.

IQVIA Disease Analyser (DA) Germany, Germany

DA Germany is collected from extracts of patient management software used by GPs and specialists practicing in ambulatory care settings.[14] Data coverage includes more than 34M distinct person records out of at total population of 80M (42.5%) in the country and collected from 2,734 providers. Patient visiting more than one provider are not cross identified for data protection reasons and therefore recorded as separate in the system. Dates of service include from 1992 through present. Observation time is defined by the first and last consultation dates. Germany has no mandatory GP system and patient have free choice of specialist. As a result, data are collected from visits to 28.8% General, 13.4% Orthopaedic Surgery, 11.8% Otolaryngology, 11.2% Dermatology, 7.7% Obstetrics/Gynaecology, 6.2% various Neurology and Psychiatry 7.0% Paediatric, 4.6% Urology, 3.7% Cardiology, 3.5% Gastroenterology, 1.5% Pulmonary and 0.7% Rheumatology practices.

Drugs are recorded as prescriptions of marketed products. No registration or approval is required for drug utilisation studies.

8.3 Study Period

The study period will be from 01/01/2012 to 31/12/2022 or the end of available data in each of the data sources if earlier (see **Table 3** for more details).

8.4 Follow-up

Study participants will be followed up from their date of first PAH diagnosis (i.e. index date) until the earliest of the following: 1) loss to follow-up, 2) 31/12/2022, 3) end of data availability, or 4) date of death.

For the Patient-level Utilisation of ERAs/ PDE-5is (i.e. to explore duration of use), participants will be followed up from the day of therapy initiation, i.e. the date of the first prescription of the drugs of interest (index date), until the earliest of loss to follow-up, end of data availability, death, or end of continuous exposure.

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 the database eligible for the study and newly diagnosed with PAH	Patient present in the database during the study period (2012-2022) and with at least 1 year of valid database history, and no record PAH diagnosis in the previous year	Single	Incident	[-365]	IP and OP	n/a	n/a	PAH	n/a	n/a
All patients from the database eligible for the study – Analysis of incident use in patients with newly diagnosed PAH	Patient present in the database during the study period (2012-2022) and with at least 1 year of valid database history	Multiple	Incident	[-365]	IP and OP	n/a	n/a	Overall, substance	n/a	n/a

[1] IP = inpatient, OP = outpatient, n/a = not applicable, PAH = Pulmonary arterial hypertension

8.5 Study population with inclusion and exclusion criteria

The study cohort will comprise all patients with newly diagnosed PAH present in the database during the study period (2012-2022) and with at least 365 days of data availability before the day they become eligible for study inclusion. The requirement of at least 365 days of data history will also hold for children. Therefore, children aged <1year will be excluded.

Given its complexity, rarity, and multifactorial nature, diagnosing PAH typically requires a comprehensive evaluation involving clinical assessment for signs/symptoms as well as various diagnostic procedures and tests. These may include electrocardiogram, chest radiography, pulmonary function tests, arterial blood gases, echocardiography, ventilation/perfusion lung scan, non-contrast and contrast-enhanced chest computed tomography examinations, digital subtraction angiography, cardiac magnetic resonance imaging, blood tests, immunology, abdominal ultrasound, cardiopulmonary exercise testing, vasoreactivity assessment, fluid challenge, and right heart catheterization. Right heart catheterization is considered the gold standard for diagnosing and classifying PAH. It provides valuable information about pulmonary arterial pressures and hemodynamic parameters.[15] In this study, cases will be identified through medical records containing SNOMED (Systematized Nomenclature of Medicine) codes that indicate a diagnosis or observation of PAH. The OMOP CDM relies on the SNOMED as the standardized vocabulary for diagnosis codes, ensuring uniform and consistent classification of various conditions.

Additional eligibility criteria were implemented, requiring patients to have a minimum follow-up time of 30 days to capture important aspects of PAH treatment (objective 1), and treatment sequences (objective 3).

Table 5. Operational Definitions of Inclusion 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
Incident PAH	Patients with newly diagnosed PAH during the study period, that is individuals without a diagnosis of PAH 1 year prior.	After	1 year	IP, OP, OT	SNO MED	First	All study participants	N/A	N/A
Prior database history of 1 year (objectives 1, 2, 3 and 4)	Study participants will be required to have a year of prior history observed before contributing observation time	After	1 year	IP, OP, OT	N/A	N/A	All patients with incident PAH	N/A	N/A
Minimum follow-up (objectives 1, 3)	Study participants will be required to have at least a minimum follow-up time of 30 days.	After	30 days	IP, OP, OT	N/A	N/A	All patients with incident PAH	N/A	N/A
Minimum potential follow-up time (objective 4)	Only patients with a diagnosis of PAH (index date) occurring one year prior to end of data availability in the database will be included to allow sufficient follow-up time to capture treatment and outcomes	After	1 year	IP, OP, OT	N/A	N/A	All patients with incident PAH	N/A	N/A

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

8.6 Variables

8.6.1 Exposure

For this study, the exposure of interest is prescription (during study period) of Ambrisentan, Bosentan, Macitentan, Sildenafil, Sitaxentan, and Tadalafil.

Exposure will be grouped at drug substance level and assessed as monotherapy or combination therapy through the treatment pattern analysis (see 8.8.3).

This list of the drugs of interests (with respective ATC code) is described in Table 6. Respective RxNorm codes (at ingredient level) are provided in appendix 1 – table 1.

Table 6: Exposure of interest

Drugs	Class	ATC code
Ambrisentan	Endothelin receptor antagonist	C02KX02
Bosentan	Endothelin receptor antagonist	C02KX01
Macitentan	Endothelin receptor antagonist	C02KX04
Sitaxentan	Endothelin receptor antagonist	C02KX03
Sildenafil	Phosphodiesterate-5 inhibitors	G04BE03
Tadalafil	Phosphodiesterate-5 inhibitors	G04BE08

Table 7: Exposure details

Exposure group name(s)	Details	Washout window	Assessment Window	Care Setting	Code Type	Diagnosis position	Applied to study populations:	Incident with respect to…	Measurement characteristics/ validation	Source of algorithm
Overall ERAs/ PDE-5is, substance, strength, route	Preliminary code lists provided in Table 6 (and table 1 of Appendix 1)	– 1 year	Calendar year	Biobank, primary and secondary care	RxNorm	N/A	All patients with newly diagnosed PAH present in the database during the study period	Previous ERAs/ PDE-5is use	N/A	N/A

8.6.2 Outcome/s

This study will examine the following four primary outcomes of interest.

- Initiation of treatment for PAH

Initiation of treatments will be assessed within a window of 30, 90, and/or 365 days following diagnosis. In relation to the first outcome, a predefined list of PAH treatments will be compiled to address Objectives 1, 2, and 3. PAH treatments will include Ambrisentan, Bosentan, Macitentan, Sitaxentan, Sildenafil and Tadalafil as listed in Table 6.

- All-cause hospitalisation

All-cause hospitalization will be determined by examining medical records with visit codes indicating inpatient visits or hospital admissions after the date of therapy initiation, i.e. the date of the first prescription of the drugs of interest (i.e. index date) following diagnosis of PAH. This will be assessed in a window of 1, 3, and 5 years following treatment initiation.

- Cardiovascular hospitalization

Cardiovascular hospitalisation as an outcome will be identified through a SNOMED code occurring within 7 days prior to admission, during hospitalization, or within 7 days following discharge will be considered as instances of cardiovascular hospitalization. Cardiovascular hospitalisation will be assessed in a window of 1, 3, and 5 years following treatment initiation after new diagnosis of PAH.

- Death

Overall death rate in patients with newly diagnosed PAH will also be identified based on the registered date of death and reported in a window of 1, 3, and 5 years following treatment initiation after new diagnosis of PAH.

Table 8. Operational Definitions of Outcome

Outcome name	Details	Primary outcome?	Type of outcome	Washout window	Care Settings[1]	Code Type	Diagnosis Position	Applied to study populations:	Measurement characteristics/validation	Source of algorithm
PAH treatments	Preliminary code lists provided in Table 6	Yes	Counts and %	365 days as incident use only	IP and OP care	RxNorm	N/A	All patients with incident PAH	N/A	N/A
Hospitalization (All-cause/ cardiovascular)	Based on visit type within OMOP-CDM	Yes	%	N/A	IP and OP care	Date of inpatient visit	N/A	All patients with incident PAH	N/A	N/A
Death rate	Based on date of death	Yes	%	N/A	IP and OP care	Date of death	N/A	All patients with incident PAH	N/A	N/A

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

8.6.3 Other covariates, including confounders, effect modifiers and other variables (where relevant)

Age at PAH diagnosis will be described. The following age grouping will be used: 1-17; 18-44; 45-64; 65 and over.

The sex (male/ female) of patients with incident PAH will also be identified.

All co-morbidities and concomitant-medications recorded prior (any time prior to the ID, 365 to 31 days prior to the ID and -30 to 1 day before ID) will be used for large-scale patient characterisation, identified as concept/code and descendants. A list of pre-specified co-morbidities and co-medications will also be provided. These will include:

- PAH-associated Relevant Medical History: chest pain, dyspnoea, fatigue, syncope, obesity, heart failure, pulmonary embolism, asthma, COPD, idiopathic pulmonary fibrosis, obstructive sleep apnoea, chronic kidney disease, connective tissue disorder, Human Immunodeficiency Virus (HIV) infection, diabetes mellitus, and systemic hypertension.
- Medications Implicated in Drug-induced PAH: selective serotonin reuptake inhibitors, aminorex, fenfluramine, benfluorex, phenylpropanolamine, dexfenfluramine, tryptophan, lithium, interferon, sofosbuvir, dasatinib, Nilotinib, ponatinib, carfilzomib, ruxolitinib, buprenorphine, tramadol, thalidomide, paclitaxel, bleomycin, cyclophosphamide, mitomycin, phentermine, mazindol.

Concept IDs of these drugs and conditions have been added to the appendix I – table 3. These codes (with descendant codes) will be further reviewed prior to the analysis.

Table 9: Operational Definitions of Covariates

Characteristic	Details	Type of variable	Assessment window	Care Settings[1]	Code Type[2]	Diagnosis Position[3]	Applied to study populations:	Measurement characteristics/ validation	Source for algorithm
Co-morbidities	Large-scale patient-level characterisation with regard to underlying comorbidities	Counts	At index date (ID), for 30 to 1 day before ID, for 365 to 31 days before ID, at any time before ID	OP, IP, OT	SNOMED	N/A	N/A	N/A	N/A
Concomitant medication#	Large-scale patient-level characterisation with regard to use of concomitant drugs	Counts	At index date (ID), for 30 to 1 day before ID, for 365 to 31 days before ID	OP, IP, OT	RxNorm	N/A	N/A	N/A	N/A

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

ERAs/ PDE-5is are the exposures of interest and are not considered as concomitant medication

8.7 Study size

No sample size has been calculated as this is a descriptive Disease Epidemiology and Drug Utilization Study where we are interested in the characteristics of all incident PAH cases and their treatments, regardless of sample size.

8.8 Analysis

Table 9. Description of Study Types and Type of analysis

STUDY TYPE	STUDY CLASSIFICATION	TYPE OF ANALYSIS
Disease epidemiology - Patient-level characterisation	Off-the-shelf (C1)	• Large-scale characterisation • Patient-level characteristics • Standard care description

8.8.1 Federated Network Analyses

Analyses will be conducted separately for each database. Before study initiation, test runs of the analytics are performed on a subset of the data sources or on a simulated set of patients and quality control checks are performed. Once all the tests are passed, the final package is released in the version-controlled Study Repository for execution against all the participating data sources.

The data partners locally execute the analytics against the OMOP-CDM in R Studio and review and approve the by default aggregated results before returning them to the Coordination Centre. Sometimes multiple execution iterations are performed, and additional fine tuning of the code base is needed. A service desk will be available during the study execution for support.

The study results of all data sources are checked after which they are made available to the team in the Digital Research Environment (DRE) and the Study Dissemination Phase can start. All results are locked and timestamped for reproducibility and transparency.

8.8.2 Patient privacy protection

Cell suppression will be applied as required by databases to protect people's privacy. Cell counts < 5 will be masked.

8.8.3 Statistical model specification and assumptions of the analytical approach considered

R-packages

We will use the R packages "CohortDiagnostics" for the patient-level characterization of demographics and clinical characteristics, as well as "TreatmentPatterns" for the patient-level characterisation of treatments.

Patient-level characterisation

Large-scale patient-level characterisation will be conducted (objective 1). Age and sex at time of PAH diagnosis will be described. Medical history and concomitant medication (excluding PAH treatments) will be assessed for anytime –and up to 366 days before index date, for 365 to 31 days before index date, for 30 to 1 day before index date, and at index date. These time windows were defined based on the options currently available in the standard analytical tools that will be used in this project. Co-variates to be presented in a summary baseline characteristics table will be pre-defined as described in section 8.6.3.

To characterise treatment of PAH, the number and % of patients receiving each of a pre-specified list of PAH treatments (objective 2, as listed in section 8.6.2) will be described at index date, 1 to 30, 1 to 90, and 1 to 365 days post index date. The index date will be the date of the PAH diagnosis for each patient. A treatment pattern analysis will be conducted to describe how use of PAH evolves over time. This analysis will produce sunburst plots and Sankey diagrams to describe treatment patterns over time (objective 3).

The number and % of patients with all-cause and cardiovascular hospitalization will be estimated from the date of first prescription of the drugs of interests to the date of first hospital admission or inpatient care due to any cause, and additionally for any cardiovascular condition (objective 4). The number of individuals who died (objective 4) will be calculated from the date of first prescription of the drugs of interests to death due to any cause and will be reported as proportion of patients who died. This analysis will be conducted only for databases with complete information on death.

For all analyses n and % will be reported. A minimum cell count of 5 will be used when reporting results, with any smaller counts reported as "<5". All analyses will be reported by country/database, overall and stratified by age and sex when possible (minimum cell count reached). Additionally, to capture treatments availability and changes over time, sunburst plots, and Sankey diagrams will be further stratified by study periods (2012-2017 and 2018-2022).

8.9 Evidence synthesis

Results from analyses described in section 8.8 will be presented separately for each database and no meta-analysis of results will be conducted.

9 DATA MANAGEMENT

9.1 Data management

All databases are mapped to the OMOP common data model. This enables the use of standardised analytics and tools across the network since the structure of the data and the terminology system is harmonised. The OMOP CDM is 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. Each data partner will execute the study code against their database containing patient-level data and will then return the results set 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 patients 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 Digital Research Environment. These output files do not contain any data that allow identification of subjects included in the study. The DRE implements further security measures in order 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 multiple myeloma, a systematic search of possible codes for inclusion was previously 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 to assess the use of different codes across the databases contributing to the study and identify any codes potentially omitted in error. This will allow for a consideration of the validity of the study cohort of patients with PAH in each of the databases, and inform decisions around whether multiple definitions are required.

The study code will be based on two R packages currently being developed to (1) characterise demographic and clinical characteristics, (2) characterise treatment patterns, and (3) estimate hospitalization and overall

survival using the OMOP CDM. 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 utilize routinely collected healthcare data, and it is important to address any potential data quality issues. One specific concern is the variation in recording comorbidities across different databases. While false positives are expected to be minimal, there may be a higher likelihood of false negatives, especially in databases that lack patient-level linkage to secondary care data. However, it is worth noting that two of the included databases, IQVIA Germany and CPRD, have previously conducted research on pulmonary arterial hypertension. This prior experience reduces the potential for misclassification. Nevertheless, acknowledging the possibility of limitations and biases in capturing outcomes related to hospitalization in a primary care database, especially in cases without linkage to secondary care data, is crucial. However, it is worth noting that CPRD GOLD as primary care database, contain valuable information on referrals to specialists or hospitals and their respective outcomes. Additionally, the included databases (Estonian Biobank, CHUBX and IQVIA Germany (for a subset of the database)) have information from secondary care settings. As a result, the potential impact of this limitation on the accuracy of capturing hospitalization outcomes would be minimal, as it is mitigated by the inclusion of databases with comprehensive data from both primary and secondary care sources. Similarly, PAH treatment is predominantly initiated in hospital settings, often depending on the severity of the presenting cases. However, the included databases effectively capture this information by linking to secondary care data or services. This linkage ensures that comprehensive data on PAH treatment initiation, even if it occurs in a hospital setting, is available within the databases but we might have misclassification of the date of first prescribing if treatment is further continued by the GP.
Additionally, we have included CHUBX, a hospital database with expertise in the treatment of pulmonary arterial hypertension. This inclusion provides reassurance that PAH cases in this database will be well phenotyped, ensuring accurate characterization of the condition.

It is important to mention that the IQVIA DA Germany database does not provide information on the date of death. Consequently, we will not be able to provide data on the proportion of patients who died during the one year following diagnosis specifically for this database. Furthermore, it is worth noting that the recording of events used for patient characterization may vary across the databases, and the documentation of treatment use may be incomplete.

12. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

In agreement with the guideline on good pharmacovigilance practice (EMA/873138/2011), there will be no requirement for expedited reporting of adverse drug reactions as only secondary data will be used in this study.

13. GOVERNANCE BOARD ASPECTS

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

14. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

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

14.1 Study Report

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 RESULTS

N/A

16. REFERENCES

1. Marc H, Christophe G, Sébastien B, Peter D, James RK, Mark RN, Andrea JO, Soni SP, Ralph TS, Kurt RS, et al. Pathology and pathobiology of pulmonary hypertension: state of the art and research perspectives. *European Respiratory Journal*. 2019;53:1801887. doi: 10.1183/13993003.01887-2018
2. Emmons-Bell S, Johnson C, Boon-Dooley A, Corris PA, Leary PJ, Rich S, Yacoub M, Roth GA. Prevalence, incidence, and survival of pulmonary arterial hypertension: A systematic review for the global burden of disease 2020 study. *Pulmonary Circulation*. 2022;12:e12020. doi: https://doi.org/10.1002/pul2.12020
3. Tobita K, Goda A, Teruya K, Nishida Y, Takeuchi K, Kikuchi H, Inami T, Kohno T, Tashiro S, Yamada S, et al. Exercise Capacity and Ventilatory Efficiency in Patients With Pulmonary Arterial Hypertension. *J Am Heart Assoc*. 2023;12:e026890.
4. Mak SM, Strickland N, Gopalan D. Complications of pulmonary hypertension: a pictorial review. *Br J Radiol*. 2017;90:20160745.
5. Hendriks PM, Staal DP, van de Groep LD, van den Toorn LM, Chandoesing PP, Kauling RM, Mager H-J, van den Bosch AE, Post MC, Boomars KA. The evolution of survival of pulmonary arterial hypertension over 15 years. *Pulmonary Circulation*. 2022;12:e12137. doi: https://doi.org/10.1002/pul2.12137
6. Quezada Loaiza CA, Velázquez Martín MT, Jiménez López-Guarch C, Ruiz Cano MJ, Navas Tejedor P, Carreira PE, Flox Camacho Á, de Pablo Gafas A, Delgado Jiménez JF, Gómez Sánchez MÁ, et al. Trends in Pulmonary Hypertension Over a Period of 30 Years: Experience From a Single Referral Centre. *Revista Española de Cardiología (English Edition)*. 2017;70:915-923. doi: 10.1016/j.rec.2016.12.044
7. Dardi F, Guarino D, Cennerazzo F, Ballerini A, Magnani I, Bertozzi R, Donato F, Martini G, Manes A, Galiè N, et al. Pulmonary Arterial Hypertension Associated with Portal Hypertension and HIV Infection: Comparative Characteristics and Prognostic Predictors. *J Clin Med*. 2023;12.
8. Lan NSH, Massam BD, Kulkarni SS, Lang CC. Pulmonary Arterial Hypertension: Pathophysiology and Treatment. *Diseases*. 2018;6.
9. McLaughlin VV, Gaine SP, Howard LS, Leuchte HH, Mathier MA, Mehta S, Palazzini M, Park MH, Tapson VF, Sitbon O. Treatment goals of pulmonary hypertension. *J Am Coll Cardiol*. 2013;62:D73-81.
10. McLaughlin VV, Shah SJ, Souza R, Humbert M. Management of pulmonary arterial hypertension. *J Am Coll Cardiol*. 2015;65:1976-1997.
11. McLaughlin VV, Archer SL, Badesch DB, Barst RJ, Farber HW, Lindner JR, Mathier MA, McGoon MD, Park MH, Rosenson RS, et al. ACCF/AHA 2009 expert consensus document on pulmonary hypertension: a report of the American College of Cardiology Foundation Task Force on Expert Consensus Documents and the American Heart Association: developed in collaboration with the American College of Chest Physicians, American Thoracic Society, Inc., and the Pulmonary Hypertension Association. *Circulation*. 2009;119:2250-2294.
12. Brat GA, Weber GM, Gehlenborg N, Avillach P, Palmer NP, Chiovato L, Cimino J, Waitman LR, Omenn GS, Malovini A, et al. International electronic health record-derived COVID-19 clinical course profiles: the 4CE consortium. *NPJ Digit Med*. 2020;3:109. doi: 10.1038/s41746-020-00308-0
13. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, Smeeth L. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol*. 2015;44:827-836. doi: 10.1093/ije/dyv098

14. Rathmann W, Bongaerts B, Carius H, Kruppert S, Kostev K. Basic characteristics and representativeness of the German Disease Analyzer database. *Int J Clin Pharmacol Ther*. 2018;56:459-466.
15. Humbert M, Kovacs G, Hoeper MM, Badagliacca R, Berger RMF, Brida M, Carlsen J, Coats AJS, Escribano-Subias P, Ferrari P, et al. 2022 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension. *Eur Respir J*. 2023;61.

17. ANNEXES

Appendix I: List of Stand-Alone documents (e.g., lists with concept codes (conditions & drugs etc.)

Appendix II: ENCePP checklist for study protocols

APPENDIX I – TABLE 1: LISTS WITH CONCEPT DEFINITIONS FOR EXPOSURE

Prescriptions will be identified based on the relevant ingredient. Non-systemic products will be excluded from the code list.

Drugs	Class	ATC code	Concept id
Ambrisentan	Endothelin receptor antagonist	C02KX02	1337068
Bosentan	Endothelin receptor antagonist	C02KX01	1321636, 19113007
Macitentan	Endothelin receptor antagonist	C02KX04	44507580
Sitaxentan	Endothelin receptor antagonist	C02KX03	36878846, 44012794
Sildenafil	Phosphodiesterate-5 inhibitors	G04BE03	1316262
Tadalafil	Phosphodiesterate-5 inhibitors	G04BE08	1336926, 36849344

Appendix I – Table 3: Characteristics of the study population

APPENDIX I – TABLE 2 CHARACTERISTICS

Characteristics	CHUBX	CPRD GOLD	EBB	IQVIA DA Germany
Number of patients				
Treated, %				
Follow up time after index (days), Mean (SD)				
Age at index (years), Mean (SD)				
Female sex, %				
Clinical symptoms & signs				
• Chest pain				
• Dyspnoea				
• Fatigue				
• Syncope				
Comorbidities (%)				
• Obesity				
• Heart failure				
• Pulmonary embolism				
• Asthma				
• COPD				
• Idiopathic pulmonary fibrosis				
• Obstructive sleep apnoea				
• Chronic kidney disease				
• Connective tissue disorder				
• Human Immunodeficiency Virus (HIV) infection				
• Diabetes mellitus				
• Systemic hypertension				
• Ischemic heart disease				
Drugs associated with PAH (%)				
• Selective serotonin reuptake inhibitors				
• Aminorex				
• Fenfluramine				
• Benfluorex				
• Phenylpropanolamine				
• Dexfenfluramine				
• Tryptophan				
• Lithium				
• Interferon				
• Sofosbuvir				
• Dasatinib				
• Nilotinib				

Characteristics	CHUBX	CPRD GOLD	EBB	IQVIA DA Germany
• Ponatinib				
• Carfilzomib				
• Ruxolitinib				
• Buprenorphine				
• Tramadol				
• Thalidomide				
• Paclitaxel				
• Bleomycin				
• Cyclophosphamide				
• Mitomycin				
• Phentermine				
• Mazindol				

APPENDIX I – TABLE 3: PRELIMINARY CODE LISTS FOR BASELINE CHARACTERIZATION

Clinical observations/ conditions	Concept ID	Code
Chest pain	77670	29857009
Dyspnea	312437	267036007
Fatigue	4223659	84229001
Syncope	135360	271594007
Obesity	433736	414916001
Heart failure	316139	84114007
Pulmonary embolism	440417	59282003
Asthma	317009	195967001
COPD	255573	13645005
Idiopathic pulmonary fibrosis	45763750	700250006
Obstructive sleep apnea	442588	78275009
Chronic kidney disease	46271022	709044004
Connective tissue disorder	253549	105969002
Human Immunodeficiency Virus (HIV) infection	439727	86406008
Diabetes mellitus	201820	73211009
Systemic hypertension	320128	59621000
Drugs associated with PAH	4185932	414545008
Selective serotonin reuptake inhibitors	21604709	N06AB
Aminorex	36850446	OMOP5168063
Fenfluramine	753860	4328
Benfluorex	19035533	18880
Phenylpropanolamine	1139993	8175
Dexfenfluramine	719057	3268
Tryptophan	19006186	10898
Lithium	19124477	6448
Interferon	35884376	OMOP5031280
Sofosbuvir	44785094	1484911
Dasatinib	1358436	475342
Nilotinib	1394023	662281
Ponatinib	43013182	1364347
Carfilzomib	42873638	1302966
Ruxolitinib	40244464	1193326
Buprenorphine	1133201	1819
Tramadol	1103314	10689
Thalidomide	19137042	10432
Paclitaxel	1378382	56946
Bleomycin	1329241	1622
Cyclophosphamide	1310317	3002
Mitomycin	1389036	632
Phentermine	735340	8152
Mazindol	794229	6664

APPENDIX II – ENCEPP CHECKLIST FOR STUDY PROTOCOLS

Study title:
DARWIN EU® - Co-prescribing of endothelin receptor antagonists and phosphodiesterate-5 inhibitors in pulmonary arterial hypertension (PAH).

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[2]	☐	☒	☐	5
1.1.2 End of data collection[3]	☒	☐	☐	
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:

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 generalized)	☒	☐	☐	
2.1.4 Which hypothesis(-es) is (are) to be tested?	☐	☐	☒	
2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☐	☒	

[2] 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.
[3] Date from which the analytical dataset is completely available.

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	☒	☐	☐	
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 categorizing 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 categorized according to time windows?	☒	☐	☐	
5.4	Is intensity of exposure addressed? (e.g., dose, duration)	☒	☐	☐	
5.5	Is exposure categorized 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 utilization, 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)	☐	☐	☒	

Section 7: Bias	**Yes**	**No**	**N/A**	**Section Number**
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

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
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:				

Section 12: Limitations	**Yes**	**No**	**N/A**	**Section Number**
12.1.1 Selection bias?	☒	☐	☐	
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).	☐	☐	☒	11
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: Johnmary T. Arinze

Date: 13/07/2023

Signature: J.T. Arinze