



Study Protocol P3-C1-008

10/07/2024

Version 3.0

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
	Dissemination level: Public	

Table of Contents

Table of Contents	2
Document History	3
LIST OF ABBREVIATIONS	5
1 TITLE	6
2 Responsible parties – study team	6
3 ABSTRACT (Standalone summary of the study protocol)	6
4 AMENDMENTS AND UPDATES	8
5 MILESTONES	9
6 RATIONALE AND BACKGROUND	9
7 RESEARCH QUESTION AND OBJECTIVES	10
8 RESEARCH METHODS	12
8.1 Study type and Study Design	12
8.2 Study Setting and Data Sources	12
8.2.1 Data sources selected	12
8.2.2 Rationale for the data sources selected	13
8.3 Study Period	15
8.4 Follow-up	15
8.5 Study Population with inclusion and exclusion criteria	18
8.6 Variables	20
8.6.1 Exposure/s	20
8.6.2 Outcome/s	22
8.6.3 Other covariates, including confounders, effect modifiers and other variables (where relevant)	22
8.7 Study size	23
8.8 Analysis	23
8.8.1 Federated Network Analyses	23
8.8.2 Patient privacy protection	23
8.8.3 Methods to derive parameters of interest	23
8.8.4 Population-level incidence calculations	24
8.8.5 Population-level prevalence calculations	25
8.8.6 Patient-level characteristics on/before index date	26
8.8.7 Sensitivity analysis	26
8.8.8 R packages	27
8.9 Evidence synthesis	27
9 DATA MANAGEMENT	27
10 QUALITY CONTROL	27
11 LIMITATIONS OF THE RESEARCH METHODS	28
12 PROTECTION OF HUMAN SUBJECTS	29
13 MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS	29
14 GOVERNANCE BOARD ASPECTS	29

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
		Dissemination level: Public

15	PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS	29
16	REFERENCES.....	30
17	ANNEXES	31

DOCUMENT HISTORY

VERSION	DATE	DESCRIPTION
V1.0	29.05.2024	First Draft submitted to EMA
V2.0	21.06.2024	Accepted version for archive
V3.0	10.07.2024	Adding EUPAS and clarification of text

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
	Dissemination level: Public	

Study Title	DARWIN EU® - Drug Utilisation Study on GLP-1 Receptor Agonists		
Protocol version identifier	V3.0		
Date of last version of protocol	10/07/2024		
EU PAS register number	EUPAS1000000223		
Active substance	Use of GLP-1 RA (objective 1) and use of medicines used for diabetes and/or weight loss that help contextualise exposure to GLP-1 RA (objective 2):		
	Objective	Substance name	ATC code
	1	Exenatide	A10BJ01
	1	Liraglutide	A10BJ02
	1	Lixisenatide	A10BJ03
	1	Albiglutide	A10BJ04
	1	Dulaglutide	A10BJ05
	1	Semaglutide	A10BJ06
	1	Beinaglutide	A10BJ07
	1	Tirzepatide	A10BX16
	2	Orlistat	A08AB01
	2	Metformin	A10BA02
	2	Naltrexone/bupropion (Mysimba)	A08AA62
	2	Phentermine/topiramate	A08AA51
	2	Phentermine	A08AA01
Medicinal product	N/A		
Research question and objectives	This study aims to provide an overview of the characteristics of patients prescribed with a GLP-1 RA medicinal product and how these have changed over the past ten years: 1) To characterise the incidence and prevalence of use (prescription/dispensation) of GLP-1 RA. Additionally, new drug users of GLP-1 RA stratified by presence or absence of diagnosis of diabetes mellitus type 2 and/or diagnosis of obesity will be characterised by age, sex, body mass index, initial dose, cumulative dose and a list of prespecified indications/comorbidities. 2) To characterise the incidence and prevalence of use		

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
		Dissemination level: Public

	(prescription/dispensation) of medicines used for weight loss that help contextualise exposure to GLP-1 RA.
Country(ies) of study	Belgium, Germany, Netherlands, Spain and the UK
Author	Marta Pineda Moncusí, Y. Guo

LIST OF ABBREVIATIONS

Acronyms/terms	Description
ATC	Anatomical Therapeutic Chemical
CDM	Common Data Model
CPRD GOLD	Clinical Practice Research Datalink GOLD
DA	Disease Analyzer
DARWIN EU®	Data Analysis and Real World Interrogation Network
DUS	Drug Utilization Study
EEA	European Economic Area
EHR	Electronic Health Records
EMA	European Medicines Agency
EU	European Union
GLP-1 RA	Glucagon-like peptide-1 receptor agonists
GP	General Practitioner
IPCI	Integrated Primary Care Information
LPD	Longitudinal Patient Data
OMOP	Observational Medical Outcomes Partnership
SIDIAP	Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
		Dissemination level: Public

1 TITLE

DARWIN EU® - Drug Utilisation Study on GLP-1 receptor agonists

2 RESPONSIBLE PARTIES – STUDY TEAM

Study team Role	Names	Organisation
Study Project Manager/Principal Investigator	Marta Pineda Moncusi	University of Oxford
Data Scientist	Yuchen Guo Edward Burn Martí Català Sabaté	University of Oxford University of Oxford University of Oxford
Epidemiologist	Annika Jödicke Daniel Prieto Alhambra	University of Oxford University of Oxford and Erasmus MC University
Clinical Domain Expert	Annika Jödicke	University of Oxford
Data Partner*	Names	Organisation
Local Study Coordinator/Data Analyst		CPRD Gold
	Talita Duarte	SIDIAP
	Fèlix Barenys	
	Irene López Sánchez	
	Katia Verhamme	IPCI
	Dina Vojinovic	IQVIA DA Germany, IQVIA LPD Belgium
	Isabella Kaczmarczyk	
	James Brash	

*Data partners' role is only to execute code at their data source, review and approve their results. These people do not have an investigator role. Data analysts/programmers do not have an investigator role and thus declaration of interests (DOI) for these people is not needed.

3 ABSTRACT (STANDALONE SUMMARY OF THE STUDY PROTOCOL)

Title

DARWIN EU® - Drug Utilisation Study on GLP-1 receptor agonists

Rationale and Background

A shortage of medicines containing Glucagon-Like Peptide-1 receptor agonists (GLP-1 RA) is affecting EU Member States since 2022 and will continue throughout 2024. The medicines belong to the class of GLP-1 RA are either authorised for the treatment of diabetes or authorised for weight management, with the exception of Mounjaro (tirzepatide), a glucose-dependent insulinotropic polypeptide (GIP) and GLP-1 RA that is authorised for both indications.

The shortage is based on an increased demand for these medicines in addition to other causes, e.g. capacity constraints. Of concern has been the off-label use of GLP-1 RAs for weight management which has been mentioned frequently in the news and social media and is exacerbating existing shortages with serious consequences for public health.

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
		Dissemination level: Public

This study aims to provide an overview of the characteristics of patients prescribed with a GLP-1 RA medicinal product and how these have changed over the past ten years. This will help contextualise what determinants might be driving the demand for GLP-1 RA vis-à-vis the observed shortage of medicines, including exploring comparative trends of prescription of other medicinal products used in diabetes and for weight management as well as patterns of off-label use.

Research question and objectives

This study aims to provide an overview of the characteristics of patients prescribed with a GLP-1 RA medicinal product and how these have changed over the past ten years.

The specific objectives of this study are:

- 1) To determine the incidence and prevalence of prescriptions of GLP-1 RA medicinal products (overall and stratified by age, sex, database, indications (only for incidence), and calendar time (per month), for the past 10 years. Additionally, new drug users of GLP-1 RA (stratified by presence or absence of diagnosis of diabetes mellitus type 2 and diagnosis of obesity indications) will be characterised by age, sex, body mass index, initial dose, cumulative dose and a list of prespecified indications.
- 2) To determine the incidence and prevalence of prescriptions of medicines used for weight loss that help contextualise exposure to GLP-1 RA, including orlistat, naltrexone/bupropion, phentermine/topiramate, phentermine and metformin (overall and stratified by age, sex, database, and calendar time (per month), for the past 10 years.

Research Methods

Study design

- Population level cohort study (Objective 1.1, Population-level drug utilisation study on GLP-1 RA)
- Patient level cohort study (Objective 1.2, Patient-level characterisation on GLP-1 RA users stratified by presence or absence of diabetes mellitus type 2 and/or obesity)
- Population level cohort study (Objective 2, Population-level drug utilisation study on medicines used for weight loss that help contextualise exposure to GLP-1 RA)

Population

Population-level drug utilisation of GLP-1 RA and medicines used for diabetes and/or weight loss that help contextualise exposure to GLP-1 RA: All individuals present in the database in the last 10 years of available data will be included in the analysis. For this population, incidence and prevalence of use of GLP-1 RA and medicines used for weight loss that help contextualise exposure to GLP-1 RA will be explored.

Patient level cohort study on GLP-1 RA users stratified by presence or absence of diabetes mellitus type 2 and/or obesity: new drug users of GLP-1 RA stratified by presence or absence of diagnosis of diabetes mellitus type 2 and/or diagnosis of obesity will be characterised by age, sex, body mass index, initial dose, cumulative dose and a list of prespecified indications/comorbidities.

Variables

Drugs of interest: list of 8 GLP-1 RA medicines (Objective 1) and 5 other medicines used for diabetes and/or weight loss that help contextualise exposure to GLP-1 RA (Objective 2).

Calendar month, age, sex and indications (diabetes mellitus type 2 and obesity, and only for incidence) will be used for stratification.

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
		Dissemination level: Public

Age, sex, body mass index, initial dose, cumulative dose and a list of prespecified indications/comorbidities will be used for the characterisation.

Data sources

- 1) CPRD Gold (UK, Primary Care Database)
- 2) SIDIAP (Spain, Primary Care Database)
- 3) IPCI (Netherlands, Primary Care Database)
- 4) IQVIA DA Germany (combination of primary and secondary care (outpatient visits) database)
- 5) IQVIA LPD Belgium (combination of primary and secondary care (outpatient visits) database)

Sample size

No sample size has been calculated. Based on a preliminary study feasibility assessment the expected number of prescriptions in the period investigated is expected to be between <1,000 and 760,000 across the five data sources considered.

Data analyses

Population-level drug utilisation study on GLP-1 RA and medicines used for weight loss that help contextualise exposure to GLP-1 RA: monthly incidence rates of use of medicines of interest per 100,000 person-years, and monthly point prevalence of medicines of interest. Incidence will be stratified by age, sex and indications cohorts in each database of the study. Prevalence will be stratified by age and sex in each database of the study.

Patient-level drug utilisation study: new drug users of GLP-1 RA stratified by presence or absence of diagnosis of diabetes mellitus type 2 and/or diagnosis of obesity will be characterised by age, sex, body mass index, initial dose, cumulative dose, and a list of prespecified indications/comorbidities.

4 AMENDMENTS AND UPDATES

NUMBER	DATE	SECTION OF STUDY PROTOCOL	AMENDMENT OR UPDATE	REASON

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
		Dissemination level: Public

5 MILESTONES

STUDY SPECIFIC DELIVERABLE	TIMELINE
Draft Study Protocol	29/05/2024
Final Study Protocol	June 2024
Creation of Analytical code	July 2024
Execution of Analytical Code on the data	July-August 2024
Interim Study Report (if applicable)	Not applicable
Draft Study Report	August 2024
Final Study Report	To be confirmed

6 RATIONALE AND BACKGROUND

A shortage of medicines containing GLP-1 Receptor Agonists is affecting EU Member States since 2022 and will continue throughout 2024. The shortage involves the medicinal products Ozempic (semaglutide), Victoza (liraglutide), Trulicity (dulaglutide) and Saxenda (liraglutide).[1]

The shortage is based on an increased demand for these medicines in addition to other causes, e.g. capacity constraints.

Medicinal products belonging to the class of Glucagon-Like Peptide-1 receptor agonists (GLP-1 RA) are either authorised for the treatment of diabetes (Ozempic, Victoza, Trulicity, Bydureon, Rybelsus) or authorised for weight management (Saxenda, Wegovy); with the exception of Mounjaro (tirzepatide), a glucose-dependent insulinotropic polypeptide (GIP) and GLP-1 RA that is authorised for both indications. Of concern has been the off-label use of GLP-1 RAs for weight management which has been mentioned frequently in the news and social media and is exacerbating existing shortages with serious consequences for public health.

This study aims to provide an overview of the characteristics of patients prescribed with a GLP-1 RA medicinal product and how these have changed over the past ten years. This will help contextualise what determinants might be driving the demand for GLP-1 RA vis-à-vis the observed shortage of medicines, including exploring comparative trends of prescription of other medicinal products used in diabetes and for weight management as well as patterns of off-label use.

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
		Dissemination level: Public

7 RESEARCH QUESTION AND OBJECTIVES

Table 1. Primary and secondary research questions and objective.

A. Primary research question and objective.

Objective:	To determine the incidence and prevalence of prescriptions of the GLP-1 RA medicines during the last 10 years of available data, stratified by age, sex, indications and speciality (only for incidence), and calendar month in each of the databases. Additionally, new drug users of GLP-1 RA (stratified by presence or absence of diagnosis of diabetes mellitus type 2 and/or diagnosis of obesity indications) will be characterised by age, sex, initial dose, cumulative dose and a list of prespecified indications/comorbidities.																				
Hypothesis:	Not applicable																				
Population (<i>mention key inclusion-exclusion criteria</i>):	The study cohort will comprise all individuals present in the database in a 10-year period from the most recent data available. Additional eligibility criteria for calculation of incidence rates will be applied, where a minimum follow-up of 365 days of data availability and no prior use of the respective drug of interest will be required.																				
Exposure:		<table><tr><th>Substance name</th><th>ATC code</th></tr><tr><td>Exenatide</td><td>A10BJ01</td></tr><tr><td>Liraglutide</td><td>A10BJ02</td></tr><tr><td>Lixisenatide</td><td>A10BJ03</td></tr><tr><td>Albiglutide</td><td>A10BJ04</td></tr><tr><td>Dulaglutide</td><td>A10BJ05</td></tr><tr><td>Semaglutide</td><td>A10BJ06</td></tr><tr><td>Beinaglutide</td><td>A10BJ07</td></tr><tr><td>Tirzepatide</td><td>A10BX16</td></tr></table>	Substance name	ATC code	Exenatide	A10BJ01	Liraglutide	A10BJ02	Lixisenatide	A10BJ03	Albiglutide	A10BJ04	Dulaglutide	A10BJ05	Semaglutide	A10BJ06	Beinaglutide	A10BJ07	Tirzepatide	A10BX16	
Substance name	ATC code																				
Exenatide	A10BJ01																				
Liraglutide	A10BJ02																				
Lixisenatide	A10BJ03																				
Albiglutide	A10BJ04																				
Dulaglutide	A10BJ05																				
Semaglutide	A10BJ06																				
Beinaglutide	A10BJ07																				
Tirzepatide	A10BX16																				
Comparator:	None																				
Outcome:	None																				
Time (<i>when follow up begins and ends</i>):	The study will comprise the last 10 years of available data counting backwards from the respective latest date of data lock of the respective databases (i.e., 2014 to the most recent data lock).																				

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
	Dissemination level: Public	

Setting:	Primary care and outpatient setting using data from the following data sources: 1) SIDIAP (Spain, Primary Care Database) 2) IPCI (Netherlands, Primary Care Database) 3) CPRD Gold (UK, Primary Care Database) 4) IQVIA DA Germany (Outpatient care) 5) IQVIA LPD Belgium (Primary Care Database)
Main measure of effect:	Incidence and prevalence rates of GLP-1 RA medicines

B. Secondary research question 1 and objective.

Objective:	To determine the incidence and prevalence of prescriptions of 5 medicines used for weight loss that help contextualise exposure to GLP-1 RA during the last 10 years of available data, stratified by age, sex, and calendar month in each of the databases.														
Hypothesis:	Not applicable														
Population (<i>mention key inclusion-exclusion criteria</i>):	The study cohort will comprise all individuals present in the database in a 10-year period from the most recent data available. Additional eligibility criteria for calculation of incidence rates will be applied, where a minimum follow-up of 365 days of data availability and no prior use of the respective drug of interest will be required.														
Exposure:	<table><tr><th>Drug name</th><th>ATC code</th></tr><tr><td>Orlistat</td><td>A08AB01</td></tr><tr><td>Metformin</td><td>A10BA02</td></tr><tr><td>Naltrexone/bupropion (Mysimba)</td><td>A08AA62</td></tr><tr><td>Phentermine/topiramate</td><td>A08AA51</td></tr><tr><td>Phentermine</td><td>A08AA01</td></tr></table>			Drug name	ATC code	Orlistat	A08AB01	Metformin	A10BA02	Naltrexone/bupropion (Mysimba)	A08AA62	Phentermine/topiramate	A08AA51	Phentermine	A08AA01
Drug name	ATC code														
Orlistat	A08AB01														
Metformin	A10BA02														
Naltrexone/bupropion (Mysimba)	A08AA62														
Phentermine/topiramate	A08AA51														
Phentermine	A08AA01														
Comparator:	None														
Outcome:	None														
Time (<i>when follow up begins and ends</i>):	The study will comprise the last 10 years of available data counting backwards from the respective latest date of data lock of the respective databases (i.e., 2014 to the most recent data lock).														

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
	Dissemination level: Public	

Setting:	Primary care and outpatient setting using data from the following data sources: 1) SIDIAP (Spain, Primary Care Database) 2) IPCI (Netherlands, Primary Care Database) 3) CPRD Gold (UK, Primary Care Database) 4) IQVIA DA Germany (Outpatient care) 5) IQVIA LPD Belgium (Primary Care Database)
Main measure of effect:	Incidence and prevalence rates of medicines used for weight loss that help contextualise exposure to GLP-1 RA medicines

8 RESEARCH METHODS

8.1 Study type and Study Design

Retrospective cohort studies will be conducted using routinely collected health data from five databases. The study contains two study types classified as “off-the-shelf” (C1) as described in the DARWIN EU® Complete Catalogue of Standard Data Analyses ([Table 2](#)):

- 1) A population-based cohort study will be conducted to address objective 1.1 and objective 2, assessing the incidence and prevalence of GLP-1 RA medicines and the selected medicines used for weight loss that help contextualise exposure to GLP-1 RA medicines.
- 2) A patient-based cohort study will be conducted to address objective 1.2, characterising first incidence users of GLP-1 RA with no record of diabetes mellitus type II nor obesity.

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 (C1)
Patient Level DUS	New drug/s user cohort	Off the shelf (C1)

8.2 Study Setting and Data Sources

8.2.1 Data sources selected

Primary care and outpatient setting using data from the following data sources:

- 1) Sistema d’Informació per al Desenvolupament de la Investigació en Atenció Primària (SIDIAP, Spain)
- 2) Integrated Primary Care Information (IPCI, Netherlands)
- 3) Clinical Practice Research Datalink GOLD (CPRD Gold, United Kingdom)
- 4) IQVIA Disease Analyzer Germany (IQVIA DA Germany, Germany)
- 5) IQVIA Longitudinal Patient Data Belgium (IQVIA LPD Belgium, Belgium)

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
		Dissemination level: Public

All databases were previously mapped to the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM). Information on data source(s) planned to be used is described in [Table 3](#).

8.2.2 Rationale for the data sources selected

This study will be conducted using 5 of the 24 databases onboarded for DARWIN EU® in 2024, representing 5 European countries. The selection of databases for this study has been based on data quality and relevance for the proposed research question.

GLP-1 RA medicines are commonly used/prescribed in primary care. Despite the possibility that its initiation may be given by a care specialist, GLP-1 RA medicines are often prescribed and used by primary care practitioner, and therefore appear to be prescribed often. With this in mind, we selected the following data sources, which covers the health care settings of interest: SIDIAP, IPCI, CPRD Gold, IQVIA DA Germany and IQVIA LPD Belgium.

The proposed data sources capture the relevant data necessary to answer the research questions of this study, cover the healthcare settings where the drugs of interest may be used (i.e., primary care), and they have recent data covering the 10-year study period. Previous checks evaluating the reliability of the data sources performed at the time of their onboarding showed that drug records represented between the 11% and 37% of the total number of records in the selected databases. Use of medicines (prescription or dispensation) was previously captured accurately in all databases in previous DARWIN EU studies (such as EUPAS105644, EUPAS105641 or EUPAS50789).

The selected databases fulfil the criteria required for a population-level drug utilisation study, as all these have counts for most of the GLP1-a treatments in the order of thousands of users, a consistent denominator (except for the last few months before data extraction for 4 and 5) allowing for accurate incidence estimation, and quick approvals/governance enabling rapid analyses as requested for a July deadline.

Table 3. Description of the selected Data Sources.

Country	Name of Database	Justification for Inclusion	Health Care setting	Type of Data	Number of active subjects	Feasibility count of exposure (if relevant)	Data lock for the last update
Spain	SIDIAP	Covers primary care setting where GLP-1 RA medicines are commonly used/prescribed	primary care records	EHR	8.5M	From <1000 to 347,000	2023-06-30
Netherlands	IPCI	Covers primary care setting where GLP-1 RA medicines are commonly used/prescribed	primary care records	EHR	2.9M	From <1,000 to 109,000	2024-04-30
United Kingdom	CPRD Gold	Covers primary care setting where GLP-1 RA medicines are	primary care records	EHR	17.4M	From 1,400 to	2024-01-01

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
	Dissemination level: Public	

Country	Name of Database	Justification for Inclusion	Health Care setting	Type of Data	Number of active subjects	Feasibility count of exposure (if relevant)	Data lock for the last update
		commonly used/prescribed				2.44 million	
Germany	IQVIA DA Germany	Covers primary care setting where GLP-1 RA medicines are commonly used/prescribed	outpatient: primary care and specialist records	EHR	43M	From 8,200 to 1.86 million	2023-09-30
Belgium	IQVIA LPD Belgium	Covers primary care setting where GLP-1 RA medicines are commonly used/prescribed	outpatient: primary care and specialist records	EHR	1.1M	From <1,000 to 188,000	2023-12-31

Feasibility count of exposure is based record counts on prescription data of the different drugs of interest. Abbreviations: GLP-1 RA = Glucagon-like peptide-1 receptor agonists, SIDIAP = Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària, IPCI = Integrated Primary Care Information, CPRD GOLD = Clinical Practice Research Datalink GOLD, DA = Disease Analyzer, LPD = Longitudinal Patient Data, EHR = Electronic Health record.

Information on data source(s) used is described below:

1. Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària [SIDIAP]
(Spain, Primary Care Database)

SIDIAP is collected from EHR records of patients receiving primary care delivered through Primary Care Teams (PCT), consisting of GPs, nurses and non-clinical staff.[4, 5] The Catalan Health Institute manages 286 out of 370 such PCT with a coverage of 5.6M patients, out of 7.8M people in the Catalan population (74%). The database started to collect data in 2006. The mean follow-up is 10 years. The observation period for a patient can be the start of the database (2006), or when a person is assigned to a Catalan Health Institute primary care centre. Date of exit can be when a person is transferred-out to a primary care centre that does not pertain to the Catalan Health Institute, or date of death, or date of end of follow-up in the database. Drug information is available from prescriptions and from dispensing records in pharmacies. Drugs not prescribed in the GP setting might be underreported; and disease diagnoses made at specialist care settings are not included. Studies using SIDIAP data require previous approval by both a Scientific and an Ethics Committee.

2. Integrated Primary Care Information [IPCI] (Netherlands, Primary Care Database)

The Integrated Primary Care Information (IPCI) database is a longitudinal observational database containing routinely collected data from computer-based patient records of a selected group of GPs throughout the Netherlands (N=723). IPCI was started in 1992 by the department of Medical Informatics of the Erasmus University Medical Center in Rotterdam with the objective to enable better post marketing surveillance of drugs. The current database includes patient records from 2006 on, when the size of the database started to increase significantly. In 2016, IPCI was certified as Regional Data Center. Since 2019 the data is also standardized to the OMOP CDM, enabling collaborative research in a large network of databases within the Observational Health Data Sciences and Informatics (OHDSI) community. The primary goal of IPCI is to enable medical research. In addition, reports are generated to inform GPs and their organizations about the

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
		Dissemination level: Public

provided care. Contributing GPs are encouraged to use this information for their internal quality evaluation. The IPCI database is registered on the European Medicines Agency (EMA) ENCePP resources database (<http://www.encepp.eu>).

3. Clinical Practice Research Datalink GOLD [CPRD] (UK, Primary Care Database)

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 [6] 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 data requires approval via the Research Data Governance Process.

4. IQVIA Disease Analyzer Germany [IQVIA Germany] (Primary and Secondary Care Database)

IQVIA Germany is collected from extracts of patient management software used by general practitioners (GPs) and specialists practicing in ambulatory care settings.[7] Data coverage includes more than 34M distinct person records out of a 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.

5. IQVIA Longitudinal Patient Data Belgium [IQVIA Belgium] (Primary Care Database)

IQVIA Belgium is a computerised network of GPs who contribute to a centralised database of anonymised data of patients with ambulatory visits. Currently, around 300 GPs from 234 practices are contributing to the database covering 1.1M patients from a total of 11.5M Belgians (10.0%). The database covers time from 2005 through the present. Observation time is defined by the first and last consultation dates. Drug information is derived from GP prescriptions. Drugs obtained over the counter by the patient outside the prescription system are not reported. No explicit registration or approval is necessary for drug utilisation studies.

8.3 Study Period

The study will comprise the last 10 years of available data counting backwards from the respective latest date of data lock of the respective databases (i.e., 2014 to the most recent data lock reported in [Table 3](#)).

8.4 Follow-up

To determine the incidence and prevalence estimates, we require the appropriate population and their contributed observation time to first be identified. The Operational Definition of follow-up is reported in [Table 4](#).

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
	Dissemination level: Public	

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	Diagnosi s position	Incident with respect to...	Measurement characteristics/ validation	Source of algorithm
All patients with prevalent use of medicines of interest	Patient present in the database during the study period (2014 to most recent data lock)	Multiple	Prevalent	None	PC, OP	n/a	n/a	n/a	n/a	n/a
All patients with incident use of medicines of interest	Patient present in the database during the study period (2014 to most recent data lock) and with at least 365 days of valid database history.	Multiple	Incident	[-Inf, -1]	PC, OP	n/a	n/a	Specific medicine of interest	n/a	n/a

¹ PC = Primary Care, OP = outpatient, n/a = not applicable

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
		Dissemination level: Public

For incidence rates, an additional criterion is required to ensure we are capturing first incident use (i.e., new drug users). Thus, follow-up will start from the date they have reached at least 365 days of data availability. Thus, the denominator population will begin contributing person time on the respective date of the latest of the following: 1) study start date, 2) date at which the observation period starts, 3) date at which the observation period has already sufficient prior history. Participants will stop contributing person time at the earliest date of the following: 1) end of available data in each of the data sources or 2) date at which the observation period of the specific person ends.

An example of entry and exit into the denominator population is shown in **Figure 2**. In this example, person ID 1 has already sufficient prior history before the study start date and observation period ends after the study end date, so will contribute during the complete study period. Person ID 2 and 4 enter the study only when they have sufficient prior history. Person ID 3 leaves when exiting the database (the end of 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.

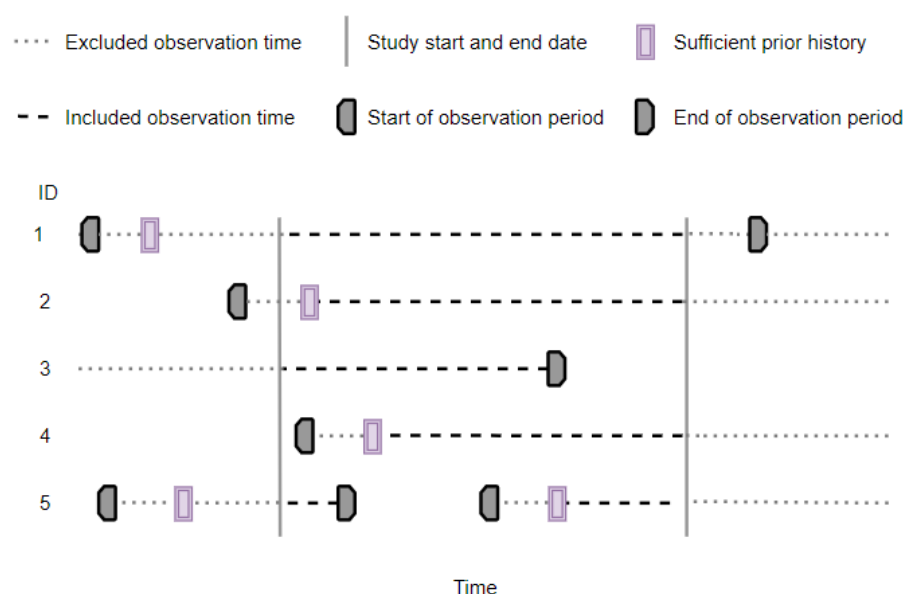


Figure 2. Included observation time for the denominator population in incidence rates.

8.5 Study Population with inclusion and exclusion criteria

The study cohort will comprise all individuals present in the database during the study period (i.e., 2014 to the most recent data lock reported).

Additional eligibility criteria for calculation of incidence rates will be applied, where a minimum follow-up of 365 days of data availability will be required to exclude individuals with a prior use of the respective drug of interest.

The operational definitions of the inclusion criteria are presented in **Table 5**.

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
	Dissemination level: Public	

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
Observation period in the database during the 10-year period from the latest available data	See under inclusion criterion	After	n/a	PC, OP	n/a	n/a	All individuals within the selected databases	n/a	n/a
Prior database history of 365 days for incidence rates calculation	Study participants will be required to have 365 days of prior history observed before contributing observation time in incidence rate calculations.	After	[-365, -1]	PC, OP	n/a	n/a	All individuals within the selected databases	n/a	n/a
Washout period for incidence rates calculation	Study participants will be required to have not used the same drug of interest before in incidence rate calculations.		[-Inf, -1]	PC, OP	n/a	n/a	All individuals within the selected databases	n/a	n/a

¹ PC = Primary Care, OP = outpatient, n/a = not applicable

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
		Dissemination level: Public

8.6 Variables

8.6.1 Exposure/s

For this study, the exposure of interest is use (during study period) of GLP-1 RA and use of medicines used for weight loss that help contextualise exposure to GLP-1 RA. This list (with respective ATC code) is described in [Table 6](#). Further details are described in the [Appendix I](#).

The list contains all GLP-1 RA medicines. However, albiglutide is no longer authorised, and therefore it won't be included in the study. Additionally, beinaglutide use has not been approved for its use in EU yet whilst tirzepatide was recently approved for its marketing in the EU for treating not only diabetes mellitus type 2 but also for weight management. [8, 9] Thus, these two drugs present no counts in the data sources and will be added in the analysis in case of future routine repeated analysis.

Details of exposure are described in [Table 7](#).

Table 6. Exposure of interest.

Objective	Substance name	ATC code
1) GLP-1 RA	Exenatide	A10BJ01
1) GLP-1 RA	Liraglutide	A10BJ02
1) GLP-1 RA	Lixisenatide	A10BJ03
1) GLP-1 RA	Albiglutide	A10BJ04
1) GLP-1 RA	Dulaglutide	A10BJ05
1) GLP-1 RA	Semaglutide	A10BJ06
1) GLP-1 RA	Beinaglutide	A10BJ07
1) GLP-1 RA	Tirzepatide	A10BX16
2) Medicines used for weight loss that help contextualise exposure to GLP-1 RA	Orlistat	A08AB01
2) Medicines used for weight loss that help contextualise exposure to GLP-1 RA	Metformin	A10BA02
2) Medicines used for weight loss that help contextualise exposure to GLP-1 RA	Naltrexone/bupropion (Mysimba)	A08AA62
2) Medicines used for weight loss that help contextualise exposure to GLP-1 RA	Phentermine/topiramate	A08AA51
2) Medicines used for weight loss that help contextualise exposure to GLP-1 RA	Phentermine	A08AA01

ATC = Anatomical Therapeutic Chemical, GLP-1 RA = Glucagon-like peptide-1 receptor agonists.

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
	Dissemination level: Public	

Table 7. Operational Definitions of Exposure.

Exposure group name(s)	Details	Washout window	Assessment Window	Care Setting ¹	Code Type	Diagnosis position	Applied to study populations	Incident with respect to...	Measurement characteristics/ validation	Source of algorithm
GLP-1 RA	Preliminary code lists provided in Appendix I	[-Inf, -1] for incident use only	Monthly	PC, OP	RxNorm	n/a	All individuals present in the database during the study period who have had a prescription / dispensation of the medicine of interest	Previous medicine use (of the specific drug of interest)	n/a	n/a
Medicines that help contextualise exposure to GLP-1 RA	Preliminary code lists provided in Appendix I	[-Inf, -1] for incident use only	Monthly	PC, OP	RxNorm	n/a	All individuals present in the database during the study period who have had a prescription / dispensation of the medicine of interest	Previous medicine use (of the specific drug of interest)	n/a	n/a

¹ PC = primary care, OP = outpatient, n/a = not applicable

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
		Dissemination level: Public

8.6.2 Outcome/s

NA.

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

Covariates for stratification in population-level drug utilisation study (objective 1.1 and 2):

- Age: age will be stratified in three groups (<18, 18-64, 65+).
- Sex (men/women)
- Time period: Monthly
- Speciality (when/if available, and only for incident use of GLP-1 RA)
- Indications cohorts (only for incident use of GLP-1 RA):
 1. Users of GLP-1 RA with diagnosis of DM2¹ and obesity²
 2. Users of GLP-1 RA with diagnosis of DM2¹ and excluding obesity²
 3. Users of GLP-1 RA with diagnosis of obesity² and excluding DM2¹
 4. Users of GLP-1 RA excluding obesity² and DM2¹

¹ DM2: diagnosis of diabetes mellitus type II before first use of GLP-1 RA.

² Obesity: diagnosed or measured anytime during the observation period.

Covariates for stratification in population-level drug utilisation study (objective 1.2):

- Age: age will be stratified in three groups (<18, 18-64, 65+).
- Sex (men/women)
- Body mass index: body mass index may be stratified in groups (such as <27, 27-29, 30- 34, 35-39, ≥40, or among larger groups such as <27, 27-39, ≥40) if enough values are reported.
- List of prespecified indications/comorbidities based on concept_ID.
 - Myocardial infarction
 - Stroke
 - Major adverse cardiac events (MACE): defined as a composite outcome of myocardial infarction and stroke
 - Heart failure
 - Chronic kidney disease
- Initial dose
- Cumulative dose

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
		Dissemination level: Public

8.7 Study size

No sample size has been calculated for this study. The expected number of prescriptions across data sources is expected to be roughly between <1,000 and 760,000.

8.8 Analysis

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 and the Dissemination Phase can start. All results are locked and timestamped for reproducibility and transparency.

8.8.2 Patient privacy protection

To prevent confidentiality issues, cell counts lower than 5 will be reported as “<5”.

8.8.3 Methods to derive parameters of interest

Incidence and prevalence calculations will be reported overall and stratified. Characterisation of new users will only be reported among new drug users of GLP-1 RA with no diagnosis of diabetes mellitus type 2 and no diagnosis of obesity.

Calendar time

Calendar time for incidence calculations will be based on monthly periods of the index prescription/dispensation in the population level analysis (e.g., 01/01/2014 – 31/01/2014, 01/02/2014 – 30/02/2014, etc.).

Calendar time for point prevalence calculations will be based on the first day of each calendar month in the population level analysis (e.g., 01/01/2024, 01/02/2024, etc.).

Age

Age at index date will be calculated using January 1st of the year of birth as proxy for the actual date of birth. The following age groups will be used for stratification: <18, 18-64, 65+.

Sex

Results will be presented stratified by sex (men/women).

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
		Dissemination level: Public

Speciality (when/if available, and only for incident use of GLP-1 RA)

When/If recorded, prescriber specialty such as family medicine, paediatrics, etc. will be used to stratify incident users of each of the GLP-1 RA medicines.

Indications cohorts (only for incident use of GLP-1 RA)

- Diabetes mellitus type 2 (DM2): diagnosis of DM2 before first use of GLP-1 RA.
- Obesity: diagnosed or measured anytime during the observation period.

Incidence calculations of each of the GLP-1RA will be also stratified by the combination of DM2 and obesity, creating the following four cohorts:

1. Users of GLP-1 RA with diagnosis of DM2 and obesity
2. Users of GLP-1 RA with diagnosis of DM2 and excluding obesity
3. Users of GLP-1 RA with diagnosis of obesity and excluding DM2
4. Users of GLP-1 RA excluding obesity and DM2

List of prespecified indications/comorbidities

A list of prespecified indications/comorbidities will be used to characterise the four indication cohorts:

- Myocardial infarction
- Stroke
- Major adverse cardiac events (MACE): defined as a composite outcome of myocardial infarction and stroke
- Heart failure
- Chronic kidney disease

Body mass index:

Body mass index will be used to characterise the four indication cohorts.

Initial dose and cumulative dose:

Initial dose and cumulative dose will be used to characterise the four indication cohorts.

8.8.4 Population-level incidence calculations

New drug users for the different medicines of interest will be captured.

Monthly incidence rates of the medicines of interest will be calculated as the number of users per 100,000 person-years of the population at risk of getting exposed during the period for each calendar year. Those study participants who enter the denominator population will then contribute time at risk up to their first use (prescription or dispensation) during the study period. Or if they do not have a drug exposure, they will contribute time at risk up as described above in section 8.4 (Follow-up). Incidence rates will be given together with 95% Poisson exact confidence intervals.

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
		Dissemination level: Public

An illustration of the calculation of incidence use for each of the medicines of interest is shown below in **Figure 3**.

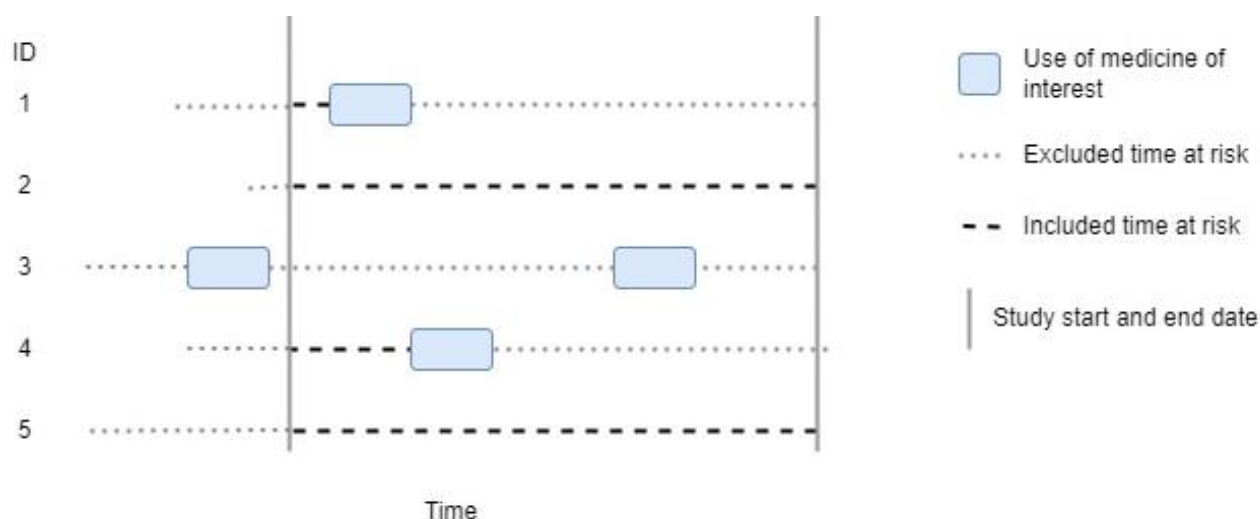


Figure 3. Example for capturing first incidence use of medicines of interest.

Patient ID 1 and 4 contribute time at risk between the study start until when they become incident users of the medicine of interest. Patient ID 2 and 5 contribute time at risk between the study start and end date as no use of medicines of interest is observed between this period nor before the study start date. Patient ID 3 was excluded from the analysis and do not contribute time at risk since a previous use of the same medicine of interest was observed before the study start date.

8.8.5 Population-level prevalence calculations

Monthly point prevalence of the medicines of interest will be calculated as the total number of individuals who use the medicine of interest during at the start of each calendar month, divided by the population at risk of getting exposed during that point. Therefore, point prevalence gives the proportion of individuals exposed at the first of each calendar month.

Point prevalence will be given together with 95% confidence intervals. Binomial 95% confidence intervals will be calculated using the Wilson Score method for binomial distribution.

An illustration of the calculation of point prevalent use for each of the medicines of interest is shown below in **Figure 4**.

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
	Dissemination level: Public	

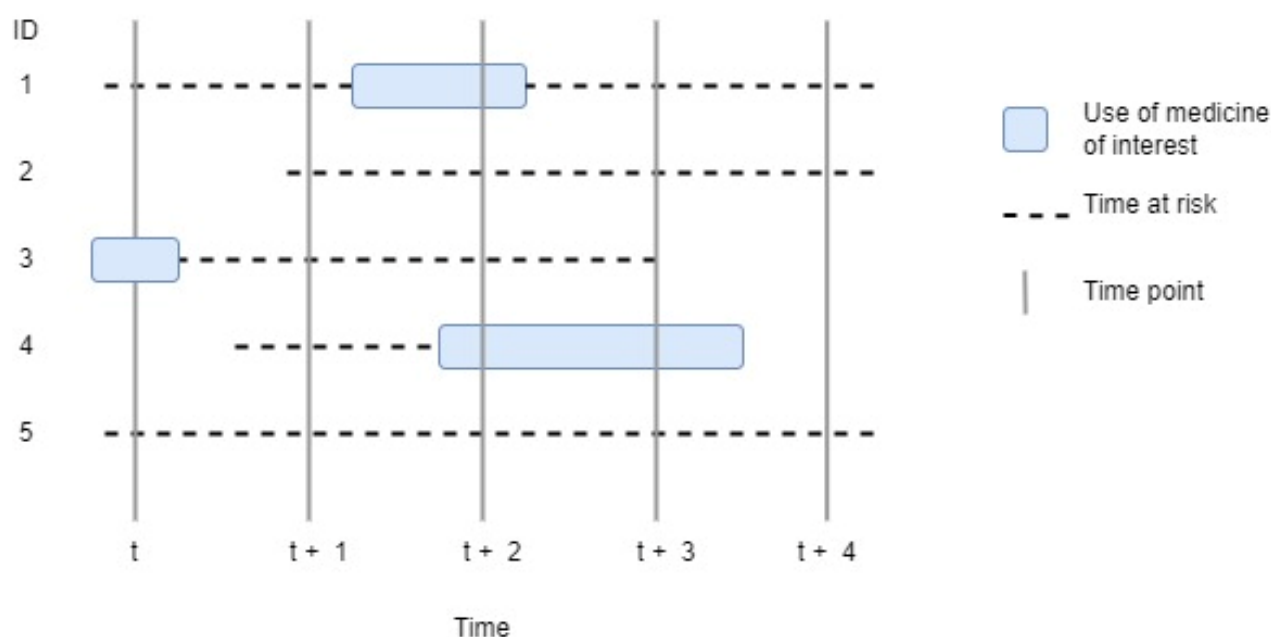


Figure 4. Example for capturing point prevalence use of medicines.

Point prevalence use is captured at the different time points, represented as “t” or “t + n”. In each of the time points, the number of users of the medicine of interest contributes to the numerator. For instance, from the 5 participants, there is only one user for each time point at time “t” and at time “t + 3”, whilst “t + 1” and “t + 4” have no users, and “t + 2” have two users of the medicine of interest. Patient ID 1 and 3 only contribute one time point each: “t” and “t + 2”, respectively. Conversely, Patient ID 4 contributes to two-time points, “t + 2” and “t + 3”, as their use of the medicine of interest continued.

8.8.6 Patient-level characteristics on/before index date

New drug users of GLP-1 RA stratified by presence or absence of diagnosis of diabetes mellitus type 2 and/or diagnosis of obesity will be characterised by age, sex, body mass index, initial dose, cumulative dose, and a list of prespecified indications/comorbidities.

Average age and number of men and women (N, %) at index date will be reported. The list of prespecified indications/comorbidities will be reported using the number of persons and percentage (N, %) with a record within the prespecified time windows. Body mass index will be reported as average and/or as number of persons and percentage (N, %) among the body mass index categories described in 8.6.3. We will consider the following time windows (and report the most relevant/s): during the prior month, during the previous year, and anytime before index date.

8.8.7 Sensitivity analysis

Overall monthly point prevalence estimates stratified for the different concepts_id (except ingredient concept_id) included in the GLP-1 RA drugs semaglutide, exenatide, liraglutide, lixisenatide and dulaglutide will be reported as sensitivity analysis to characterise the use of individual strengths.

An example of different concepts_id for semaglutide is presented in [Table 8](#). Concepts with less than 30 counts will not be run.

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
		Dissemination level: Public

Table 8. Example of concepts_id for semaglutide that are not ingredient concept_id.

concepts_id	Ingredient name	Name
36809756	semaglutide	3 ML semaglutide 1.34 MG/ML Injectable Solution [Ozempic] by Novo Nordisk
36809755	semaglutide	3 ML semaglutide 1.34 MG/ML Injectable Solution
36809147	semaglutide	1.5 ML semaglutide 1.34 MG/ML Injectable Solution [Ozempic] by Novo Nordisk
36809148	semaglutide	1.5 ML semaglutide 1.34 MG/ML Injectable Solution

8.8.8 R packages

We will use the R package “IncidencePrevalence” for the population-level estimation of drug utilisation. Additionally, we will use the R package “DrugUtilization” for patient-level characterisation.

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

All databases will have been 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>.

This 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.

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

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
		Dissemination level: Public

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 antibiotics 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 so data quality issues must be considered. In particular, a recording of a prescription or dispensation does not mean that the patient actually took the drug. In addition, assumptions around the duration of drug use will be unavoidable.

The actual reason for prescription of the drug is not recorded in any of the databases. We will assess indication via a proxy based on pre-defined conditions recorded on the date of therapy initiation. Therefore, recording of potential indication may be incomplete. In addition, the completeness of recordings of co-morbidities used for patient characterisation may vary across databases.

Additionally, prescriber speciality may also not be available in the selected data bases. We will include it in the analysis when possible.

Database-specific limitations:

In IQVIA LPD Belgium and IQVIA DA Germany, the observation period of the patients in these databases is calculated based on the last visit, observation or interaction of the patient with the health care system. This methodology impacts the individuals considered “at risk” for the different medicines of interest of the study (i.e., the individuals included in the denominator populations) during the latest months of available data from the latest data lock, where healthy and/or non-frequent users of the health care system will not be considered active. Consequently, the denominators that will be used to calculate the incident use of drugs in the population may present an artefactual decrease whilst the incident users will remain, incrementing the incidence and prevalence ratios. The presence of these artefacts will be considered when interpreting the results.

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
		Dissemination level: Public

12 PROTECTION OF HUMAN SUBJECTS

For this study, the data partners 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.

For each collaborator, the output files will be stored according to the specific legislation of each database, such as the use of Remote Research Environments. These output files do not contain any data that allow identification of subjects included in the study. The RRE 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) 679/2016 in the various member states.

13 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.

14 GOVERNANCE BOARD ASPECTS

SIDIAP, IPCI and CPRD will require to undergo their respective ethical approvals.

15 PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

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.

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
		Dissemination level: Public

16 REFERENCES

1. EMA. *Public information on medicine shortages*. 2024; Available from: <https://www.ema.europa.eu/en/human-regulatory-overview/post-authorisation/medicine-shortages-availability-issues/public-information-medicine-shortages#ema-shortages-catalogue-11001>.
2. Au, K. and W. Yang, *GLP-1 Receptor Agonists in Weight Loss and Bariatric Surgery: Balancing Efficacy and Gastrointestinal Adverse Events*. *Obesity Surgery*, 2024. **34**(4): p. 1382-1383.
3. UK, D. *our response to serious supply issues of drugs for people with type 2 diabetes*. 2024 20-03-2024; Available from: <https://www.diabetes.org.uk/about-us/news-and-views/our-response-serious-supply-issues-drugs-people-living-type-2-diabetes>.
4. Recalde, M., et al., *Data Resource Profile: The Information System for Research in Primary Care (SIDIAP)*. *Int J Epidemiol*, 2022. **51**(6): p. e324-e336.
5. Raventos, B., et al., *Transforming the Information System for Research in Primary Care (SIDIAP) in Catalonia to the OMOP Common Data Model and Its Use for COVID-19 Research*. *Clin Epidemiol*, 2023. **15**: p. 969-986.
6. Herrett, E., et al., *Data Resource Profile: Clinical Practice Research Datalink (CPRD)*. *Int J Epidemiol*, 2015. **44**(3): p. 827-36.
7. Rathmann, W., et al., *Basic characteristics and representativeness of the German Disease Analyzer database*. *Int J Clin Pharmacol Ther*, 2018. **56**(10): p. 459-466.
8. ChiCTR. *Chinese Clinical Trial Registry: A Multicenter, double-blinded, randomized, placebo-controlled, parallel-group study to evaluate the efficacy and safety of Beinaglutide therapy adjunctive with life style intervention in overweight/obesity adult. (ChiCTR1900023428)*. 2019; Available from: <https://www.chictr.org.cn/showprojEN.html?proj=38105>.
9. EMA. *Mounjaro - opinion on variation to marketing authorisation (tirzepatide)*. 2023 10/11/2023; Available from: <https://www.ema.europa.eu/en/medicines/human/variation/mounjaro>.

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
		Dissemination level: Public

17 ANNEXES

Appendix I: Preliminary code lists for GLP-1 RA and for medicines that help contextualise exposure to GLP-1 RA.

Objective	Substance Name	Class	ATC code	Ingredient ConceptID
1) GLP-1 RA	Exenatide	GLP-1 analogues	A10BJ01	1583722
1) GLP-1 RA	Liraglutide	GLP-1 analogues	A10BJ02	40170911
1) GLP-1 RA	Lixisenatide	GLP-1 analogues	A10BJ03	44506754
1) GLP-1 RA	Albiglutide	GLP-1 analogues	A10BJ04	44816332
1) GLP-1 RA	Dulaglutide	GLP-1 analogues	A10BJ05	45774435
1) GLP-1 RA	Semaglutide	GLP-1 analogues	A10BJ06	793143
1) GLP-1 RA	Beinaglutide	GLP-1 analogues	A10BJ07	36850844
1) GLP-1 RA	Tirzepatide	Other blood glucose lowering drugs, excl. insulins	A10BX16	779705
2) Medicines used for weight loss that help contextualise exposure to GLP-1 RA	Orlistat	Peripherally acting antiobesity products	A08AB01	741530
2) Medicines used for weight loss that help contextualise exposure to GLP-1 RA	Metformin	Biguanides	A10BA02	1503297
2) Medicines used for weight loss that help contextualise exposure to GLP-1 RA	Naltrexone/Bupropion (Mysimba)	Centrally acting antiobesity products	A08AA62	1714319 + 750982
2) Medicines used for weight loss that help contextualise exposure to GLP-1 RA	Phentermine/topiramate	Centrally acting antiobesity products	A08AA51	735340 + 742267
2) Medicines used for weight loss that help contextualise exposure to GLP-1 RA	Phentermine	Centrally acting antiobesity products	A08AA01	735340

GLP-1 = Glucagon-like peptide-1 receptor agonists

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
	Dissemination level: Public	

Appendix II: ENCePP checklist for study protocols.

ENCePP Checklist for Study Protocols (Revision 4)

Study title: Drug Utilisation Studies on GLP-1 Agonists

EU PAS Register® number: TBC Study reference number (if applicable): P3-C1-008

Section 1: Milestones	Yes	No	N/A	Section Number
1.1 Does the protocol specify timelines for				6 - milestones
1.1.1 Start of data collection ¹	☒	☐	☐	
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:

--

Section 2: Research question	Yes	No	N/A	Section Number
2.1 Does the formulation of the research question and objectives clearly explain:	☒	☐	☐	7 – Research questions and objectives
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 <i>a priori</i> hypothesis?	☐	☐	☒	

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.

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
	Dissemination level: Public	

<u>Section 3: Study design</u>	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)	☒	☐	☐	8.8.4-8.8.6
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:

--

<u>Section 4: Source and study populations</u>	Yes	No	N/A	Section Number
4.1 Is the source population described?	☒	☐	☐	8.2
4.2 Is the planned study population defined in terms of:				8.3-8.5
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:

--

<u>Section 5: Exposure definition and measurement</u>	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)	☒	☐	☐	8.6.1
5.2 Does the protocol address the validity of the exposure measurement? (e.g. precision, accuracy, use of validation sub-study)	☐	☐	☒	

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
	Dissemination level: Public	

<u>Section 5: Exposure definition and measurement</u>	Yes	No	N/A	Section Number
5.3 Is exposure categorised according to time windows?	☒	☐	☐	8.6.1
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:

<u>Section 6: Outcome definition and measurement</u>	Yes	No	N/A	Section Number
6.1 Does the protocol specify the primary and secondary (if applicable) outcome(s) to be investigated?	☐	☐	☒	
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:

<u>Section 7: Bias</u>	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:

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
	Dissemination level: Public	

<u>Section 8: Effect measure modification</u>	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:

--

<u>Section 9: Data sources</u>	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)	☒	☐	☐	8.6.1
9.1.2 Outcomes? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	☐	☐	☒	
9.1.3 Covariates and other characteristics?	☒	☐	☐	8.6.3
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)	☐	☐	☒	
9.2.2 Outcomes? (e.g. date of occurrence, multiple event, severity measures related to event)	☐	☐	☒	
9.2.3 Covariates and other characteristics? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	☒	☐	☐	8.6.3
9.3 Is a coding system described for:				
9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☒	☐	☐	8.6.1, and Appendix I
9.3.2 Outcomes? (e.g. International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	☐	☐	☒	
9.3.3 Covariates and other characteristics?	☐	☒	☐	
9.4 Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☐	☐	☒	

Comments:

--

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
	Dissemination level: Public	

<u>Section 10: Analysis plan</u>	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?	☒	☐	☐	8.8.6
10.4 Are stratified analyses included?	☒	☐	☐	8.6.3, 8.8.3
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:

--

<u>Section 11: Data management and quality control</u>	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
11.2 Are methods of quality assurance described?	☒	☐	☐	10
11.3 Is there a system in place for independent review of study results?	☐	☐	☒	

Comments:

--

<u>Section 12: Limitations</u>	Yes	No	N/A	Section Number
12.1 Does the protocol discuss the impact on the study results of:				11
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).	☐	☐	☒	
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)	☐	☐	☐	

	D2.2.3 - Study Protocol for P3-C1-008	
	Author(s): M. Pineda-Moncusí, Y. Guo	Version: V3.0
	Dissemination level: Public	

Comments:

<u>Section 13: Ethical/data protection issues</u>	Yes	No	N/A	Section Number
13.1 Have requirements of Ethics Committee/ Institutional Review Board been described?	☒	☐	☐	12
13.2 Has any outcome of an ethical review procedure been addressed?	☒	☐	☐	12
13.3 Have data protection requirements been described?	☒	☐	☐	12

Comments:

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

Comments:

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

Comments:

Name of the main author of the protocol: Marta Pineda Moncusí

Date: 27/May/2024

Signature: MPM