Study Protocol

P4-C1-001

DARWIN EU® - Clozapine and the incidence of agranulocytosis over time

10/04/2025

Version 2.0

CONTENTS

Study title	Clozapine and the incidence of agranulocytosis over time
Protocol version	V2.0
Date	10/04/2025
EU PAS number	EUPAS1000000549
Active substance	Clozapine, WHO ATC code N05AH02
Medicinal product	Not applicable
Research question and objectives	Research question: What is the incidence of agranulocytosis and neutropenia over time in new users of clozapine? Study objectives: 1. To estimate the incidence rates of agranulocytosis and neutropenia in consecutive weekly and monthly intervals following the initiation of clozapine treatment, overall and stratified by age and sex. 2. To characterise the timing of agranulocytosis and neutropenia events during clozapine treatment using Kaplan-Meier curves, overall and stratified by age and sex. 3. To characterise individuals initiating clozapine treatment in terms of demographics and pre-specified conditions related to the indication for clozapine use. 4. To determine the treatment duration for clozapine use.
Country(ies) of study	Finland, Denmark, Croatia, Germany and Spain
Author(s)	Ellen Gerritsen, e.gerritsen@darwin-eu.org Dina Vojinovic, d.vojinovic@darwin-eu.org

LIST OF ABBREVIATIONS

Acronyms/term	Description
ATC	Anatomical Therapeutic Chemical classification system
CDM	Common Data Model
CHI	Catalan Health Institute
CIPH	Croatian Institute of Public Health
CM	Clinical Modification
DA	Disease Analyzer
DARWIN EU®	Data Analysis and Real World Interrogation Network
DK-DHR	Danish Data Health registries
DRE	Digital Research Environment
DOI	Declaration of interests
DQD	Data Quality Dashboard
DRE	Digital Research Environment
DUS	Drug Utilisation Study
ED	Emergency Department
EEA	European Economic Area
EHR	Electronic Health Records
EMA	European Medicines Agency
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EU	European Union
EUDA	European Union Drug Agency
FinOMOP-HILMO	Finnish Care Register for Health Care
GP	General Practitioner
GDPR	General Data Protection Regulation
ICD	International Classification of Diseases
ICU	Intensive Care Unit
ID	Index date
IP	Inpatient
MA	Marketing Authorisation
NA	Not applicable
NAJS	Croatian National Public Health Information System
OHDSI	Observational Health Data Sciences and Informatics
OMOP	Observational Medical Outcomes Partnership
OP	Outpatient
OT	Other
PAS	Post-Authorization Studies
PDP	Parkinson's disease psychosis
RCT	Randomised Controlled Trial
SD	Standard deviation
SIDIAP	Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària
SNOMED	Systematized Nomenclature of Medicine
WHO	World Health Organisation

1. TITLE

DARWIN EU® - Clozapine and the incidence of agranulocytosis over time

2. RESPONSIBLE PARTIES – STUDY TEAM

Study team Role	Names	Organisation
Principal Investigator	Ellen Gerritsen Dina Vojinovic	IQVIA
Data Scientist	Akram Mendez Isabella Kaczmarczyk	IQVIA
Study Manager	Natasha Yefimenko	Erasmus MC
Data Partner*	**Names**	**Organisation**
FinOMOP-HILMO	Tuomo Nieminen Tiina Wahlfors	Finnish Care Register for Health Care
DK-DHR	Claus Møldrup Elvira Bräuner Susanne Bruun	Danish Medicines Agency
NAJS	Ivan Pristaš Marko Čavlina Antea Jezidžić Jakov Vuković Anamaria Jurčević	Croatian Institute for Public Health
IQVIA DA Germany	Gargi Jadhav James Brash	IQVIA
SIDIAP	Talita Duarte Salles Elena Roel Agustina Giuliodori Picco	Institute for Primary Health Care Research Jordi Gol i Gurina

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

3. ABSTRACT

Title

DARWIN EU® – Clozapine and the incidence of agranulocytosis over time

Rationale and background

Clozapine is an effective treatment for treatment-resistant schizophrenia and Parkinson's disease psychosis, but it carries a risk of severe haematological complications, including neutropenia and agranulocytosis. Emerging evidence suggests that the risk is highest in the initial months of treatment, yet stringent haematological monitoring requirements remain in place throughout long-term use. These requirements may hinder clinical practice, leading to underuse, early treatment discontinuation, or reluctance to initiate therapy. This study aims to provide epidemiological evidence on the incidence and timing of clozapine-associated neutropenia and agranulocytosis across Europe.

Research question and objectives

Research question

What is the incidence of agranulocytosis and neutropenia over time in new users of clozapine?

Study objectives

1. To estimate the incidence rates of agranulocytosis and neutropenia in consecutive weekly and monthly intervals following the initiation of clozapine treatment, overall and stratified by age and sex.
2. To characterise the timing of agranulocytosis and neutropenia events during clozapine treatment using Kaplan-Meier curves, overall and stratified by age and sex.
3. To characterise individuals initiating clozapine treatment in terms of demographics and pre-specified conditions related to the indication for clozapine use.
4. To determine the treatment duration for clozapine use.

Methods

Study design

This retrospective cohort study aims to describe the epidemiology of agranulocytosis and neutropenia in new clozapine users at a population level (objective 1), characterise the time to onset of these conditions during clozapine treatment (objective 2) and analyse drug utilisation patterns, including demographics, pre-specified conditions related to clozapine indication and treatment duration in new users (objective 3 and 4).

Study period

1st of January 2010 to 31st of December 2024 (or latest available data).

Population

Population-level descriptive epidemiology: Population-level descriptive epidemiology will include all new users of clozapine registered in the respective databases between 1st of January 2010 and 31st of December 2024 (or latest data available). Eligible individuals must have at least 1 year of data visibility prior to becoming eligible for study inclusion and no history of clozapine use. Additionally, to ensure sufficient

follow-up, only individuals who initiated clozapine treatment at least 1 year before the end of the available data in the respective data source will be included. Children <1 year of age will be excluded.

Patient-level characterisation: Patient-level characterisation will include all new users of clozapine registered in the respective databases between 1st of January 2010 and 31st of December 2024. Eligible individuals must have at least 1 year of data visibility prior to becoming eligible for study inclusion and no history of clozapine use. Additionally, to ensure sufficient follow-up, only individuals who initiated clozapine treatment at least 1 year before the end of the available data will be included. Children <1 year of age will be excluded.

Patient-level utilisation of clozapine: Patient-level drug utilisation analysis will include all new users of clozapine in the period between 1st of January 2010 and 31st of December 2024, with at least 1 year of data visibility prior to becoming eligible for study inclusion and no history of clozapine use prior to the index date. To ensure sufficient follow-up, only individuals who initiated clozapine treatment at least one year before the end of the available data will be included.

Variables

Drug of interest: Clozapine.

Outcomes of interest: Agranulocytosis (narrow definition), neutropenia (narrow definition) and a combined outcome of neutropenia and agranulocytosis (broad definition) following the initiation of clozapine treatment. To ensure that only incident cases are captured, individuals with a prior history of agranulocytosis or neutropenia will be excluded.

Data source

1. Finnish Care Register for Health Care (FinOMOP-HILMO), Finland
2. Danish Data Health Registries (DK-DHR), Denmark
3. Croatian National Public Health Information System (NAJS), Croatia
4. IQVIA Disease Analyzer Germany (IQVIA DA Germany), Germany
5. The Information System for Research in Primary Care (SIDIAP), Spain

Sample size

No sample size has been calculated as this is a descriptive study.

Statistical analysis

Population-level descriptive epidemiology: Incidence rates of newly diagnosed agranulocytosis and neutropenia will be estimated following clozapine treatment initiation (objective 1). These rates will be expressed as the number of individuals with the newly diagnosed outcome of interest following clozapine initiation per 1,000 person-years. Incidence rates will be calculated for consecutive weekly (0-7 days, 8-14 days, 15-21 days etc.) and monthly intervals (0-30 days, 31-60 days, 61-90 days etc.) since the initiation of clozapine treatment (index date), with a maximum follow-up period of 24 months for reporting both weekly and monthly estimates. The statistical analysis will be performed based on OMOP-CDM mapped data using the "*IncidencePrevalence*" R package. The results will be reported overall and stratified by age and sex.

Patient-level characterisation: The timing of agranulocytosis and neutropenia events during clozapine treatment will be characterised using Kaplan-Meier curves (objective 2). This analysis will be conducted

using the "*CohortSurvival*" R package based on OMOP-CDM mapped data. The results will be stratified by age and sex.

Patient-level utilisation of clozapine: Characterisation including age and sex will be assessed at the date of new (incident) prescription of clozapine (index date) (objective 3). The frequency of pre-specified conditions related to clozapine initiation will be assessed at any time prior to 1 day before index date, 365 days prior to 1 day before index date and at the index date (objective 3). Duration of treatment will be calculated and summarised providing the minimum, quartiles and maximum, where available (objective 4). The FinOMOP-HILMO database will not participate in objective 4 as its recorded medication data are standardised to a fixed duration, making it impossible to accurately determine the actual treatment length. Statistical analyses will be conducted using the "*CohortCharacteristics*" and "*DrugUtilisation*" R package based on OMOP-CDM mapped data.

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

4. AMENDMENTS AND UPDATES

None.

5. MILESTONES

Study milestones and deliverables	Planned dates*
Draft Study Protocol	19th March 2025
Final Study Protocol	15th April 2025
Creation of Analytical code	March/April 2025
Execution of Analytical Code on the data	April 2025
Draft Study Report	13th May 2025
Final Study Report	TBC

*Planned dates are dependent on obtaining approvals from the internal review boards of the data sources.

6. RATIONALE AND BACKGROUND

Clozapine is an atypical antipsychotic widely regarded as the most effective treatment for individuals with treatment-resistant schizophrenia and it is also indicated for Parkinson's disease psychosis (PDP), with occasional off-label use for other neuropsychiatric conditions. Despite its efficacy, clozapine is associated with severe haematological complications such as neutropenia and agranulocytosis.[1] Neutropenia, characterized by an abnormally low neutrophil count, increases susceptibility to life-threatening infections.[2, 3] Agranulocytosis is a more severe form of neutropenia often necessitating immediate medical intervention. To mitigate these risks, stringent haematological monitoring protocols have been established, requiring frequent blood testing.[1]

Emerging evidence suggests that the highest risk for these adverse events occurs in the initial months of treatment, yet stringent haematological monitoring remains required throughout long-term use. This stringent monitoring potentially limits clinical practice, leading to under-prescription, early discontinuation, or reluctance to initiate clozapine treatment.[2, 3]

Given these considerations, a comprehensive understanding of the incidence and timing of clozapine-associated neutropenia and agranulocytosis is essential for informing clinical guidelines, optimising monitoring strategies and improving patient outcomes. This study aims to generate epidemiological evidence on the temporal patterns of these haematological adverse events in individuals initiating clozapine across Europe.

7. RESEARCH QUESTION AND OBJECTIVES

Research question

What is the incidence of agranulocytosis and neutropenia over time in new users of clozapine?

Study objectives

1. To estimate the incidence rates of agranulocytosis and neutropenia in consecutive weekly and monthly intervals following the initiation of clozapine treatment, overall and stratified by age, and sex.
2. To characterise the timing of agranulocytosis and neutropenia events during clozapine treatment using Kaplan-Meier curves, overall and stratified by age and sex.
3. To characterise individuals initiating clozapine treatment in terms of demographics and pre-specified conditions related to the indication for clozapine use.
4. To determine the treatment duration for clozapine use.

Description of the proposed objectives to be achieved in the study is displayed in **Table 1**.

Table 1. Primary and secondary research questions and objectives.

A. Study objectives 1 and 2.

Objective:	**Objective 1:** To estimate the incidence rates of agranulocytosis and neutropenia in consecutive weekly and monthly intervals following the initiation of clozapine treatment, overall and stratified by age and sex. **Objective 2:** To characterise the timing of agranulocytosis and neutropenia events during clozapine treatment using Kaplan-Meier curves, overall and stratified by age and sex.
Hypothesis:	Not applicable
Population *(mention key inclusion-exclusion criteria):*	All new users of clozapine registered in the respective databases between 1st of January 2010 and 31st of December 2024 (or latest date available). Eligible individuals must have at least 1 year of data visibility prior to becoming eligible for study inclusion and no history of clozapine use. Additionally, to ensure sufficient follow-up, the individuals initiating clozapine treatment between 1st of January 2010 and one year before the end of available data in the respective database will be included. Children <1 year of age will be excluded.
Exposure:	Clozapine
Comparator:	None

Outcome:	Agranulocytosis (narrow definition), neutropenia (narrow definition) and agranulocytosis and neutropenia (broad definition) following the initiation of clozapine treatment.
Time *(when follow up begins and ends):*	Follow-up will start when study participants fulfil inclusion criteria (i.e. first prescription of clozapine between 1st of January 2010 and 31st of December 2024, with at least 1 year of data visibility prior to becoming eligible for study inclusion and no history of clozapine use). End of follow-up will be defined as earliest of following: 1) end of clozapine treatment, 2) first outcome of interest, 3) loss to follow-up, 4) end of data availability, 5) date of death or 6) end of study period (31st of December 2024).
Setting:	Primary care, registry, outpatient specialist care and inpatient care setting using data from the following 5 data sources: FinOMOP-HILMO, DK-DHR, NAJS, IQVIA DA Germany and SIDIAP.
Main measure of effect:	Weekly and monthly incidence rates of newly diagnosed agranulocytosis and neutropenia following clozapine treatment initiation (expressed as the number of individuals with the newly diagnosed outcome of interest following clozapine initiation per 1,000 person-years), overall and stratified by age and sex. Kaplan-Meier curves for the timing of agranulocytosis and neutropenia events during clozapine treatment, overall and stratified by age and sex.

B. Study objectives 3 and 4.

Objective:	**Objective 3**: To characterise individuals initiating clozapine treatment in terms of demographics and pre-specified conditions related to the indication for clozapine use. **Objective 4:** To determine the treatment duration for clozapine use.
Hypothesis:	Not applicable
Population *(mention key inclusion-exclusion criteria):*	All new users of clozapine registered in the respective databases between 1st of January 2010 and 31st of December 2024. Eligible individuals must have at least 1 year of data visibility and no history of clozapine use. To ensure sufficient follow-up, only individuals who initiated clozapine treatment at least 1 year before the end of the available data will be included. Children <1 year of age will be excluded.
Exposure:	Not applicable
Comparator:	None
Outcome:	Clozapine
Time *(when follow up begins and ends):*	Follow-up will start when study participants fulfil inclusion criteria (i.e. first prescription of clozapine between 1st of January 2010 and 31st of December 2024, with at least 1 year of data visibility prior to becoming eligible for study inclusion and no history of clozapine use). End of follow-up will be defined as earliest of following: 1) end of clozapine treatment, 2) loss to follow-up, 3) end of data availability, 4) date of death or 5) end of study period (31st of December 2024).
Setting:	Primary care, registry, outpatient specialist care and inpatient care setting using data from the following 5 data sources: FinOMOP-HILMO, DK-DHR, NAJS, IQVIA DA Germany and SIDIAP.
Main measure of effect:	Age and sex for new (incident) users of clozapine. Frequency of pre-specified conditions related to clozapine initiation. Treatment duration for clozapine use expressed as minimum, quartiles and maximum.

8. RESEARCH METHODS

8.1 Study type and study design

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

A cohort study will be conducted using routinely collected health data from 5 data sources. The study will comprise three consecutive parts:

- Population-level cohort study (Objective 1, Population-level descriptive epidemiology of agranulocytosis and neutropenia in new users of clozapine).
- Cohort analysis (Objective 2, Patient-level characterisation to the time of onset of agranulocytosis and neutropenia during clozapine treatment).
- New drug user cohort (Objective 3 and 4, Patient-level drug utilisation regarding demographics, pre-specified conditions related to clozapine indication and treatment duration).

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

Study type	Study design	Study classification
Population-level descriptive epidemiology	Population-level cohort	Off the shelf
Patient-level characterisation	Cohort analysis	Off the shelf
Patient Level DUS	New drug/s user cohort	Off the shelf

8.2 Study setting and data sources

The study will be conducted using routinely collected data from 5 data sources in 5 European Union (EU) countries. All databases were previously mapped to the OMOP Common Data Model (CDM).

1. Finnish Care Register for Health Care (FinOMOP-HILMO), Finland
2. Danish Data Health Registries (DK-DHR), Denmark
3. Croatian National Public Health Information System (NAJS), Croatia
4. IQVIA Disease Analyzer Germany (IQVIA DA Germany), Germany
5. The Information System for Research in Primary Care (SIDIAP), Spain

For this study, we have selected 5 databases that were considered fit for purpose from the databases available in the DARWIN EU® Database Catalogue. The selection process was based on several key criteria. Firstly, the number of individuals prescribed clozapine and number of individuals diagnosed with agranulocytosis and neutropenia in the suggested databases are sufficient. Based on a preliminary feasibility assessment, the expected number of person counts for agranulocytosis differs across databases and ranges from 300 in FinOMOP-HILMO to 42,900 person count in IQVIA DA Germany. For neutropenia, the person count varies from 11,300 in FinOMOP-HILMO to 13,500 person count in SIDIAP. Secondly, the geographical distribution of the data sources was considered to ensure a diverse and representative sample. Additionally, we selected databases which cover the relevant setting for this particular outcome

and exposure (hospital, primary care linked with hospital data, registry and primary care including specialist data). The experience gained from databases that had previously participated in similar DARWIN EU® studies was considered, leveraging their proven reliability and data quality. Lastly, only databases that could deliver timely (within 4 months) were considered.

Information on data sources planned to be used with a justification for their choice in terms of ability to capture the relevant data is described in **Table 3**.

When it comes to assessing the reliability of data sources, the data partners are asked to describe their internal data quality process on the source data as part of the DARWIN EU® onboarding procedure. To further ensure data quality, we utilise the Achilles tool, which systematically characterises the data and presents it in a dashboard format that is inspected. The generated data characteristics such as age distribution, condition prevalence per year, data density and measurement value distribution can be compared against expectations for the data. Additionally, the data quality dashboard (DQD) provides more objective checks on plausibility consistently across the data sources. In terms of relevance, more general-purpose diagnostic tools, "*CohortDiagnostics"* (https://github.com/darwin-eu-dev/CohortDiagnostics) and "*DrugExposureDiagnostics"* (https://darwin-eu.github.io/DrugExposureDiagnostics/), were developed. "*CohortDiagnostics"* R package evaluates phenotype algorithms for OMOP CDM datasets, offering a standard set of analytics for understanding patient capture including data generation. It provides additional insights into cohort characteristics, record counts and index event misclassification. "*DrugExposureDiagnostics"* R package assesses ingredient specific diagnostics for drug exposure records. Furthermore, timeliness is guarded by extracting the release dates for each dataset in the network and monitoring when data are out-of-date with the expected refresh cycle (typically quarterly or half-yearly). In addition, it is important to have clear understanding of the time period covered by each released database, as this can vary across different domains. To facilitate this, the CdmOnboarding (and Achilles) packages contain a 'data density' plot. This plot displays the number of records per OMOP domain on a monthly basis. This allows to get insights into when data collection started, when new sources of data were added and until when data was included.

Table 3. Description of the selected data sources.

Country	Name of Database	Justification for Inclusion	Health Care setting	Type of Data	Number of individuals	Data lock for the last update
Finland	FinOMOP-HILMO	The database includes healthcare settings where prescriptions for clozapine and diagnoses of agranulocytosis and/or neutropenia may be recorded	Primary care, outpatient specialist care and inpatient care	EHR and registries	7.1 million	01/10/2024
Denmark	DK-DHR	Database covers healthcare setting where prescriptions for clozapine and diagnoses of agranulocytosis and/or neutropenia may be recorded	Registry	Registry	8.6 million	18/01/2025
Croatia	NAJS	The database includes healthcare settings where prescriptions for clozapine and diagnoses of agranulocytosis and/or neutropenia may be recorded	Primary care, outpatient specialist care and inpatient care	EHR and registries	5.4 million	17/11/2023
Germany	IQVIA DA Germany	The database includes healthcare settings where prescriptions for clozapine and diagnoses of agranulocytosis and/or neutropenia may be recorded	Primary care and outpatient specialist care	EHR	43 million	30/09/2024
Spain	SIDIAP	The database includes healthcare settings where prescriptions for clozapine and diagnoses of agranulocytosis and/or neutropenia may be recorded	Primary care with link to hospital data	EHR	8.6 million	30/06/2023

FinOMOP-HILMO = Finnish Care Register for Health Care; DK-DHR = Danish Data Health Registries; NAJS = Croatian National Public Health Information System; DA = Disease Analyzer; SIDIAP = Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària; EHR = Electronic Heath record;

FinOMOP-HILMO, Finland

This database covers both public and private, primary and specialised inpatient and outpatient health care encounters in Finland starting from 2011. The entire public sector and private inpatient encounters have been included since 2011, while private outpatient encounters, including occupational care, are included since 2020. The main content of the THL CDM is The Finnish Care Register for Health Care, which is a continuation of the former Hospital Discharge Register, which that originally gathered data on patients discharged from hospitals. The Care Register has comprehensive data on the use of services and service users from Finnish public inpatient and outpatient primary and specialised care nationwide. Since 1998, the register has covered both public outpatient and inpatient specialised care and private inpatient care (TerveysHilmo). From 2011 the register has covered public primary care (AvoHilmo). From 2020 the register has covered private outpatient care and occupational care. In addition, the CDM also contains the vaccination data from the Finnish National Vaccination Register, the vaccination data from the Finnish National Vaccination Register, and COVID-19 test results from the Finnish National Infectious Diseases Register, which is maintained by THL. The CDM includes all the above-mentioned data sources and is limited to observation periods commencing after 1.1.2011. The National Population is used to form the base population. This ensures up-to-date location (municipality of residence) of patients and complete death occurrences (although not the cause of death). Using the complete population as a basis for the person table also facilitates calculations on a population level, e.g. incidence rates. The current CDM population comprises all persons having been alive and residing in Finland since the beginning of 2011.

The FinOMOP-HILMO database will not participate in the analysis estimating treatment duration (objective 4) as all recorded medication data are standardised to a fixed duration making it impossible to accurately determine the actual length of treatment for individual patients.

Danish Data Health Registries (DK-DHR), Denmark

Denmark Danish health data is collected, stored and managed in national health registers at the Danish Health Data Authority and covers the entire population which makes it possible to study the development of diseases and their treatment over time. There are no gaps in terms of gender, age and geography in Danish health data due to mandatory reporting on all patients from birth to death, in all hospitals and medical clinics. Personal identification numbers enable linking of data across registers, so it captures data on all Danes throughout their lives, regardless of whether they have moved around the country. The high quality of Danish health data is attributed to standardisation, digitisation and comprehensive documentation, which together enhance accuracy, consistency and reliability, minimising potential for interpretation errors. The Danish Health Data Authority is responsible for the national health registers and for maintaining and developing standards and classifications in the Danish healthcare system. Legislation ensures balance between personal data protection and use. The current data release includes data on the entire Danish population of 5.9 million persons from 1995. It includes data from the following registries: The central Person Registry, The National Patient Registry, The Register of Pharmaceutical Sales, The National Cancer Register, The Cause of Death registry, the Laboratory Database (including coronavirus disease 2019 test results) and the Vaccination Registry (including COVID-19 vaccinations).

National Public Health Information System (NAJS), Croatia

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

primarily from EHR and holds information on demographics, inpatient and outpatient visits, conditions and procedures, drugs (outpatient and inpatient prescriptions), measurements, and inpatient and outpatient dates of death. NAJS provides linkage between medical and public health data collected and stored in health registries and other health data collections, including cancer registry, mortality, work injuries, occupational diseases, communicable and non-communicable diseases, health events, disabilities, psychosis and suicide, diabetes, drug abuse and others. The CDM population comprises all publicly insured persons residing in Croatia starting in 2015. NAJS will provide data from 2017 onwards only, as prior data might include information on duplicated patients.

IQVIA Disease Analyser (DA) Germany, Germany

IQVIA Disease Analyzer (DA) Germany is a database of de-identified electronic medical records from specialised and general primary practices (GP) in Germany since 1992.[4] This dataset encompasses approximately 3% of all outpatient practices within Germany, ensuring a substantial representation of the national healthcare landscape. The sampling methods used for practice selection, taking into account physician's demographics, specialty focus, community size category and federal state location, were instrumental in constructing a database that accurately mirrors the diverse spectrum of healthcare providers in the country. Consequently, data within IQVIA DA Germany database has been demonstrated to be representative of general and specialised practices throughout Germany.

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

No registration or approval is required for drug utilisation studies. As previously demonstrated, IQVIA DA Germany is suitable for pharmacoepidemiologic and pharmacoeconomic studies.[5, 6]

Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària (SIDIAP), Spain

The Information System for Research in Primary Care (SIDIAP) is a clinical database of anonymized patient records in Catalonia, Spain. The Spanish public healthcare system covers more than 98% of the population, and more than two thirds of the Catalan population see their GP at least once a year. The computerisation of the primary care patient records of the Catalan Health Institute (CHI) was complete in 2005. SIDIAP was designed to provide a valid and reliable database of information from clinical records of patients registered in primary care centres for use in biomedical research. SIDIAP contains data of anonymized patients' healthcare records for nearly six million people (approximately 80% of the Catalan population) registered in 287 primary care practices throughout Catalonia since 2005. It includes data collected by health professionals during routine visits in primary care, including anthropometric measurements, clinical diagnoses (International Classification of Diseases 10th revision ICD-10), laboratory tests, prescribed and dispensed drugs, hospital referrals, demographic and lifestyle information. It was previously shown that SIDIAP population is highly representative of the entire Catalan region in terms of geographic, age, and sex distributions. The high quality of these data has been previously documented, and SIDIAP has been successfully applied to epidemiological studies of key exposures and outcomes. Quality checks to identify duplicate patient IDs are performed centrally at each SIDIAP database update. Checks for logical values and data harmonisation are performed. For biochemistry data, consistency for measurements taken in different laboratories is assessed, and unit conversion is undertaken when needed.

8.3 Study period

The study period will be from 1st of January 2010 until the earliest of 31st of December 2024 or the date of the last database update for each respective database (please see **Table 3** for more details on the last update for each database).

8.4 Follow-up

For population-level descriptive epidemiology of agranulocytosis and neutropenia following initiation of clozapine treatment, as well as for patient-level characterisation up to the onset of outcomes of interest, follow-up will start when study participants fulfil inclusion criteria. Eligibility requirements are a first prescription of clozapine between 1st of January 2010 and 31st of December 2024, with at least 1 year of data visibility prior to becoming eligible for study inclusion and no history of clozapine use. End of follow-up will be defined as earliest of following: 1) end of clozapine treatment, 2) first outcome of interest, 3) loss to follow-up, 4) end of data availability, 5) date of death or 6) end of study period (31st December 2024). Additionally, incidence rates will be estimated in the pre-defined consecutive weekly (0-7 days, 8-14 days, 15-21 days etc.) and monthly intervals (0-30 days, 31-60 days, 61-90 days etc.) since the initiation of clozapine treatment (post index date), with a maximum follow-up period of 24 months for reporting both weekly and monthly estimates. Therefore, the study participants will be censored at the end of each time window if they have not experienced the outcome of interest.

For patient-level utilisation, follow-up will start when study participants fulfil inclusion (i.e. first prescription of clozapine between 1st of January 2010 and 31st of December 2024, with at least 1 year of data visibility prior to becoming eligible for study inclusion and no history of clozapine use). End of follow-up up will be defined as earliest of following: 1) end of clozapine treatment, 2) loss to follow-up, 3) end of data availability, 4) date of death or 5) end of study period (31st of December 2024).

For all analyses, additional eligibility criteria will be added to ensure sufficient follow-up (the individuals initiating clozapine treatment between 1st of January 2010 and one year before the end of available data in the respective database will be included).

The operational definition of the index date and other primary time anchors are presented by means of **Table 4**.

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

Study population name(s)	Time Anchor Description (e.g. time 0)	Number of entries	Type of entry	Washout window	Care Setting[1]	Code Type[2]	Diagnosis position	Incident with respect to…	Measurement characteristics/validation	Source of algorithm
New users of clozapine	Date of first prescription of clozapine	Single entry	Incident	[-Inf, -1]	IP, OP, OT	RxNorm	n/a	Use of clozapine	n/a	n/a

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

8.5 Study population with inclusion and exclusion criteria

The study population will include all new users of clozapine registered in the respective data sources between 1st of January 2010 and 31st of December 2024 (or latest date available). “New use” refers to a first prescription of clozapine within the study period with no clozapine prescription in the medical history.

Eligibility Criteria: At least 1 year of data visibility before starting clozapine treatment.

Additional eligibility criteria: To ensure sufficient follow-up, only individuals who initiated clozapine treatment at least one year before the end of the available data in the respective data source will be included.

The operational definitions of inclusion criteria are presented by means of **Table 5.**

Table 5. Operational definitions of inclusion criteria.

Criterion	Details	Order of application*	Assessment window	Care Settings[1]	Code Type	Diagnosis position[2]	Applied to study populations:	Measurement characteristics/ validation	Source for algorithm
New clozapine users	Individuals initiating clozapine treatment	n/a	[-Inf, -1]	IP, OP, OT	RxNorm	n/a	All study populations	n/a	n/a
Prior database history	Study participants will be required to have 365 days of prior history observed before contributing observation time	Prior	[-365, 0]	IP, OP, OT	n/a	n/a	Individuals initiating clozapine treatment	n/a	n/a
Washout period	Individuals who initiated clozapine treatment will be required to have not used clozapine ever in the past.	Prior	[-Inf, -1]	IP, OP, OT	RxNorm	n/a	Individuals initiating clozapine treatment	n/a	n/a
Observational period in the data source during the period 01/01/2010-31/12/2024 (or the latest date available)	All individuals present in the data source in the period 2010-2024 (or the latest date available)	After	n/a	IP, OP, OT	n/a	n/a	Individuals initiating clozapine treatment	n/a	n/a
Minimum potential follow-up (objective 1, 2, 3 and 4)	Potential follow-up time	After	n/a	IP, OP, OT	n/a	n/a	Individuals initiating clozapine treatment	n/a	n/a

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

[2] Specify whether a diagnosis code is required to be in the primary position (main reason for encounter); *Order of application specifies whether the eligibility criterion is applied before or after selection of the study entry date. For example, selecting “before” means that all possible study entry dates are identified, and then one or more is chosen. For instance, selecting 'after' means that the first possible study entry date is chosen, followed by the application of the inclusion and/or exclusion criteria. If the patient does not meet the criterion, then the patient drops out.

8.6 Variables

8.6.1 Exposure

For this study, exposure of interest is use (during the study period) of clozapine. Preliminary code list is provided in **Appendix I**.

The operational definition of exposure is described by means of **Table 6.**

8.6.2 Outcomes

For this study, the outcomes of interest are neutropenia (narrow definition), agranulocytosis (narrow definition) and a combined outcome of neutropenia and agranulocytosis (broad definition). The narrow definitions will be specified through a set of concepts codes as outlined in **Appendix I**. The broad definition will use a combined set of condition concept codes covering both agranulocytosis and neutropenia. Phenotype of agranulocytosis and neutropenia will be determined following input from EMA.

To ensure that only incident cases are captured, individuals with a prior history of agranulocytosis or neutropenia will be excluded.

The operational definition of the outcomes is presented in **Table 7**.

Table 6. Operational definition of exposure.

Exposure group name(s)	Details	Washout window	Assessment Window	Care Setting[1]	Code Type	Diagnosis position[2]	Applied to study populations	Incident with respect to…	Measurement characteristics/ validation	Source of algorithm
Clozapine	Preliminary code list provided in Appendix I	[-Inf, -1]	Weekly and monthly intervals post index date	IP, OP, OT	RxNorm	n/a	All individuals present in the database	Previous use of clozapine	n/a	n/a

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

[2] Specify whether a diagnosis code is required to be in the primary position (main reason for encounter).

Table 7. Operational definitions of outcome.

Outcome name	Details	Primary outcome?	Type of outcome	Washout window	Care Settings[1]	Code Type	Diagnosis Position[2]	Applied to study populations	Measurement characteristics/ validation	Source of algorithm
Agranulocytosis (narrow definition)	Preliminary code list provided in Appendix I	Yes	Binary	[-Inf, 0]	IP, OP, OT	SNOMED	n/a	All eligible individuals within the data source (objective 1 and 2)	n/a	n/a
Neutropenia (narrow definition)	Preliminary code list provided in Appendix I	Yes	Binary	[-Inf, 0]	IP, OP, OT	SNOMED	n/a	All eligible individuals within the data source (objective 1 and 2)	n/a	n/a
Agranulocytosis and neutropenia (broad definition)	Preliminary code list provided in Appendix I	Yes	Binary	[-Inf, 0]	IP, OP, OT	SNOMED	n/a	All eligible individuals within the data source (objective 1 and 2)	n/a	n/a

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

[2] Specify whether a diagnosis code is required to be in the primary position (main reason for encounter).

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

Population-level descriptive epidemiology (incidence rates of agranulocytosis and neutropenia following initiation of clozapine treatment)

Covariates for stratification in population-level descriptive epidemiology (objective 1) will include:

- Weekly intervals post-index
- Monthly intervals post-index
- Age groups: overall, young (≤ 34 years), middle-age (35 - 64 years) and older adults (≥ 65 years)
- Sex: both, males and females

Patient-level characterisation (time to onset of agranulocytosis and neutropenia during clozapine treatment)

Covariates for stratification of patient-level characterisation (objective 2) will include:

- Age groups: overall, young (≤ 34 years), middle-age (35 - 64 years) and older adults (≥ 65 years)
- Sex: both, males and females

Patient-level utilisation of new users of clozapine (characterisation of new (incident) clozapine users)

Additional variables for the characterisation of new clozapine users (objective 3) will include age, sex and a pre-specified list of conditions related to clozapine indication. These conditions will encompass both authorized and non-authorized indications for clozapine use including:

- Treatment-resistant schizophrenia
- Psychotic disorders in Parkinson's disease
- Delirium in Parkinson's disease
- Suicidal or aggressive behaviour in patients with schizophrenia
- Unknown
- None

The “unknown" indication category includes individuals that are not present in indication cohort but have records of other conditions in the condition occurrence table. The “none" category will be assigned to individuals that are neither in an indication cohort or the condition occurrence table, they will be considered as having no observed indication.

The frequency of these conditions will be assessed at three time points: 1) any time prior to 1 day before index date, 2) 365 days prior to 1 day before index date and 3) at the index date.

The operational definition of the covariates is described in **Table 8**. The preliminary list of concepts for the prespecified conditions of interest is provided in **Appendix I**.

Table 8. Operational definitions of covariates.

Characteristic	Details	Type of variable	Assessment window	Care Settings[1]	Code Type	Diagnosis Position[2]	Applied to study populations	Measurement characteristics/ validation	Source for algorithm
Demographics (age, sex)	Characterisation in terms of age and sex	Counts	At ID	IP, OP, OT	SNOMED	n/a	All new users of clozapine eligible for the study	n/a	n/a
Indication for prescribing	Check for pre-specified conditions of interest related to use of clozapine	Counts	At ID, in window around ID [-365, -1] and any time prior to ID [-Inf, -1]	IP, OP, OT	SNOMED	n/a	All new users of clozapine eligible for the study	n/a	n/a

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

8.7 Study size

No formal sample size calculation was conducted for this descriptive study, as the primary objective is to describe the incidence of agranulocytosis and neutropenia over time in individuals following initiation of clozapine, irrespective of sample size.

8.8 Analysis

The type of analysis by study type is fixed, as can be observed in **Table 9**.

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

Study type	Study classification	Type of analysis
Population-level descriptive epidemiology	Off-the-shelf	- Incidence rates of agranulocytosis and neutropenia following treatment initiation of clozapine
Patient-level characterisation	Off-the-shelf	- Prognosis / progression to a pre-specified outcome (diagnosis of agranulocytosis or neutropenia)
Patient Level DUS	Off-the-shelf	- Characterisation of patient-level features (age, sex) - Frequency of pre-specified conditions related to clozapine treatment initiation - Estimation of minimum, p25, median, p75 and maximum treatment duration for clozapine.

8.8.1 Federated network analysis

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.

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 package "*IncidencePrevalence*" (https://github.com/darwin-eu/IncidencePrevalence) for population-level descriptive epidemiology to estimate incidence rates of agranulocytosis and neutropenia following new initiation clozapine treatment. For patient-level characterisation, we will use "*CohortSurvival*" (https://github.com/darwin-eu/CohortSurvival) to generate the Kaplan-Meier's curves for the timing of agranulocytosis and neutropenia events during clozapine treatment. Additionally, "*CohortCharacteristics*" (https://github.com/darwin-eu/CohortCharacteristics) and "*DrugUtilisation*" (https://github.com/darwin-eu/DrugUtilisation) will be used to characterise new users of clozapine and to calculate treatment duration of clozapine.

Population-level descriptive epidemiology – incidence calculation

Incidence rates of newly diagnosed agranulocytosis and neutropenia following clozapine treatment initiation will be estimated. These rates will be expressed as the number of individuals with the outcome of interest following clozapine initiation divided by the sum of person-years contributed by the population at risk of the outcome during the follow-up period.

Study participants in the denominator population will begin contributing person time from the point they fulfil inclusion criteria, defined as the first prescription of clozapine between 1st of January 2010 and 31st of December 2024, with at least 1 year of data visibility prior to becoming eligible for study inclusion and no history of clozapine use. Follow-up will be censored as earliest occurrence of following: end of clozapine treatment, diagnosis of the first outcome of interest, loss to follow-up, death or end of study period (31st of December 2024) or data availability.

Incidence rates will be calculated for consecutive weekly (0-7 days, 8-14 days, 15-21 days etc.) and monthly intervals (0-30 days, 31-60 days, 61-90 days etc.) since the initiation of clozapine treatment (index date), with a maximum follow-up period of 24 months for reporting both weekly and monthly estimates. The incidence rates will be expressed per 1,000 person-years and reported alongside 95% Poisson confidence intervals.

Incidence rates will be stratified by age and sex. Age-specific cohorts will be defined based on age-boundary eligibility criteria and sex-specific cohorts will be defined based on sex eligibility criteria.

Patient-level characterisation: the time to onset of agranulocytosis and neutropenia in clozapine initiators

The timing of agranulocytosis and neutropenia events during clozapine treatment will be characterised using Kaplan-Meier survival analysis. Results will be presented as Kaplan-Meier curves, alongside estimated probabilities of agranulocytosis and neutropenia following treatment initiation. Individuals will be censored at the earliest occurrence of any of the following: end of clozapine treatment, diagnosis of the first outcome of interest, loss to follow-up, death or end of study period (31st of December 2024) or end of data availability.

Patient-level drug utilisation

New users will be selected based on their recorded prescription of clozapine within the study period. For each patient, at least 1 year of data visibility will be required prior to a prescription. To ensure incident use, individuals must have at least one year of data availability before treatment initiation and no prior exposure to clozapine. If the start date of a prescription does not fulfil the exposure washout criteria, the whole exposure will be eliminated.

New drug user cohort study will be used to characterise patient-level drug utilisation in terms of demographics, pre-specified conditions related to the indication for clozapine use and treatment duration. Demographics (age and sex) will be described at the index date while frequency of pre-specified conditions will be assessed at any time prior to 1 day before index date, 365 days prior to 1 day before index date and at the index date.

Drug exposure calculations

Drug eras will be defined as follows: exposure starts at the date of the first prescription after an indefinite washout period. For each prescription, the estimated duration of use is retrieved from the drug exposure table in the CDM, using the start and end date of the exposure. Subsequent prescriptions will be combined into continuous exposed episodes (drug eras) using the following specifications: two drug eras will be merged into one continuous drug era if the distance in days between end of the first era and start of the second era is ≤ 30 days. The time between the two joined eras will be considered as exposed by the first era as shown in **Figure 1**, first row.

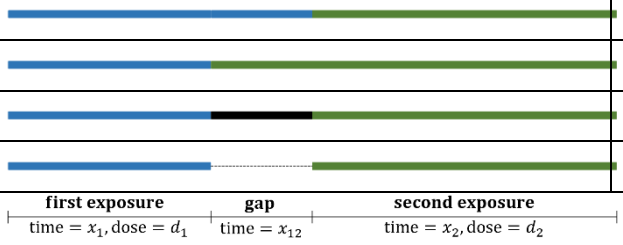
Gap era joint mode	Schematics	Dose in between	Cumulative dose	Cumulative time
"first"		d_1	$d_1 \cdot (x_1 + x_{12}) + d_2 \cdot x_2$	$x_1 + x_{12} + x_2$
"second"		d_2	$d_1 \cdot x_1 + d_2 \cdot (x_2 + x_{12})$	$x_1 + x_{12} + x_2$
"zero"		0	$d_1 \cdot x_1 + d_2 \cdot x_2$	$x_1 + x_{12} + x_2$
"join"		NA	$d_1 \cdot x_1 + d_2 \cdot x_2$	$x_1 + x_2$

Figure 1. Gap era joint mode.

Treatment duration

Treatment duration will be summarised providing the minimum, quartiles, maximum duration of treatment era. For databases where duration cannot be calculated, due to e.g. missing information on quantity or dosing, treatment duration will not be provided.

8.8.4 Methods to deal with missing data

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

8.8.5 Sensitivity analysis

Not applicable.

8.9 Evidence synthesis

Results from analyses described in section 8.8 Analysis 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 CDM. 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.

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 cohorts for medicinal products, 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,

"*DrugExposureDiagnostics*" will be run if needed to assess the use of different codes across the databases contributing to the study.

The study code will be based on four R packages namely the "*CohortCharacteristics*", "*CohortSurvival*", "*IncidencePrevalence*", and "*DrugUtilisation*" packages. 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 healthcare data, and it is important to consider several factors that may influence the interpretation of the results.

General limitations:

Data sources/setting: this study utilises data from five sources: FinOMOP-HILMO, DK-DHR, NAJS, IQVIA DA Germany and SIDIAP. The results derived from these databases may not be representative of prescriptions in other countries or databases. Variations in results are expected across different countries and healthcare settings. Additionally, discrepancies may arise due to differences in how observation periods are handled across data sources. For instance, IQVIA DA Germany uses the last interaction with the healthcare system to define the end of the observation period. As a result, infrequent users may have shorter follow-up periods, decreasing the time at risk (i.e., the denominator) for incidence rate calculations. This could lead to an overestimation of incidence rates in the final months of the study period, as users are fully captured by the end of the study.

Drug prescriptions: a recorded prescription does not necessarily indicate that the patient actually took the drug. Therefore, assumptions of actual use are made.

Characterisation/indication: the accuracy and consistency of pre-defined conditions, which is crucial for patient characterisation and identification of the potential indication may vary across the data sources included in the study. The actual indication for prescribing the drug of interest is not directly recorded as such in the databases. Instead, we assessed indication through proxies by analysing pre-defined conditions recorded around the date of therapy initiation. Consequently, the estimation of potential indications may be incomplete, given that the actual indications were not directly recorded in the data.

Treatment duration: The FinOMOP-HILMO database does not allow for a reliable assessment of treatment duration. As all recorded medication data are standardised to a fixed duration, it is not possible to accurately determine the actual length of treatment for individual patients.

Study-specific limitations:

Phenotype of agranulocytosis and neutropenia: Outcomes of interest are defined based on standard concept ids without incorporating laboratory measurements. Diagnostic codes may not capture subclinical or transient cases, lack granularity in disease severity, and vary across healthcare settings. Consequently, this may lead to an underestimation of the true incidence.

In DK-DHR, there is lack of granularity in neutropenia and agranulocytosis classification. The available codes do not distinguish between neutropenia and agranulocytosis. Additionally, there is incomplete capture of specific conditions such as secondary agranulocytosis and chronic idiopathic neutropenia, coding gap for rare disorders (e.g. Kostmann syndrome) and lack of unique codes for combination conditions such as agranulocytosis with AIDS or HIV-infection. These gaps may potentially lead to misclassification of certain cases.

12. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

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

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

13. GOVERNANCE BOARD ASPECTS

Some of the data sources require approval from their respective IRB board, except for IQVIA DA Germany which will not require any further specific approvals to undertake this study.

14. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

14.1 Study report

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

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

15. REFERENCES

1. Verdoux H, et al., *The time has come for revising the rules of clozapine blood monitoring in Europe. A joint expert statement from the European Clozapine Task Force.* Eur Psychiatry, 2025.
2. Northwood K, et al., *Evaluating the epidemiology of clozapine-associated neutropenia among people on clozapine across Australia and Aotearoa New Zealand: a retrospective cohort study.* Lancet Psychiatry, 2024.
3. Rubio, J.M., et al., *Long-term persistence of the risk of agranulocytosis with clozapine compared with other antipsychotics: a nationwide cohort and case-control study in Finland.* Lancet Psychiatry, 2024. **11**(6): p. 443-450.
4. 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.
5. Tanislav, C., et al., *No increased incidence of venous thrombosis or pulmonary embolism after SARS-CoV-2 vaccination in Germany.* Public Health, 2022. **207**: p. 14-18.
6. Ly, N.F., et al., *Impact of European Union Label Changes for Fluoroquinolone-Containing Medicinal Products for Systemic and Inhalation Use: Post-Referral Prescribing Trends.* Drug Saf, 2023. **46**(4): p. 405-416.

16. ANNEXES

16.1 Appendix I: List of preliminary concept definitions

Preliminary list of concept definition for drug of interest

Concept id	Concept Code	Concept Name	Descendants
800878	2626	clozapine	Yes

Preliminary list of concept definitions for outcomes of interest

The primary outcomes of interest are agranulocytosis (narrow definition), neutropenia (narrow definition) and agranulocytosis and neutropenia (broad definition).

The narrow definition will be based on a distinct set of concept codes, specified separately for agranulocytosis and neutropenia. A preliminary list of relevant concept ids for each condition included in the narrow definition is listed in the tables below (refer to the “narrow definition” column). These outcomes of interest occur after the initiation of clozapine and will be considered part of the narrow definition.

The broad phenotype will include a combined set of concept codes covering both agranulocytosis and neutropenia (please refer to the “narrow definition” columns in the tables below).

Additionally, individuals with a record from a preliminary list of prevalent concept ids for agranulocytosis and neutropenia, defined as those already present before clozapine treatment initiation, prior to the index date, will be excluded (please refer to the “prevalent” column in the tables below).

The final phenotype definitions will be determined based on input from the EMA.

Agranulocytosis

Concept id	Concept Code	Concept Name	Descendants	Exclude	Narrow definition	Prevalent
440689	17182001	Agranulocytosis	-	-	Yes	Yes
4224553	421312009	Agranulocytosis associated with AIDS	-	Yes	-	Yes
37017282	713530002	Agranulocytosis co-occurrent with human immunodeficiency virus infection	-	Yes	-	Yes
601107	350691000119103	Agranulocytosis due to and following administration of antineoplastic agent	-	Yes	-	Yes
606394	1144929002	Secondary agranulocytosis	-	-	Yes	Yes

Neutropenia

Concept id	Concept Code	Concept Name	Descendants	Exclude	Narrow definition	Prevalent
604243	1156296001	Acquired neutropenia	-	-	Yes	Yes
4101126	191345000	Acquired neutropenia in newborn	-	Yes	-	Yes
37205095	784392009	Adult chronic idiopathic neutropenia	-	Yes	-	Yes
4030703	14333004	Alloimmune neonatal neutropenia	-	Yes	-	Yes
4125635	234425008	Autoimmune neutropenia	-	Yes	-	Yes
36674950	770947009	Autosomal dominant severe congenital neutropenia	-	Yes	-	Yes
37204524	783201001	Autosomal recessive severe congenital neutropenia due to CSF3R deficiency	-	Yes	-	Yes
37204523	783200000	Autosomal recessive severe congenital neutropenia due to CXCR2 deficiency	-	Yes	-	Yes

37204406	783058007	Autosomal recessive severe congenital neutropenia due to G6PC3 deficiency	-	Yes	-	Yes
37203832	783199003	Autosomal recessive severe congenital neutropenia due to JAGN1 deficiency	-	Yes	-	Yes
604245	1156300000	Benign ethnic neutropenia	-	Yes	-	Yes
4174297	276628009	Chloramphenicol-induced neutropenia	-	Yes	-	Yes
4122939	234423001	Chronic benign neutropenia	-	Yes	-	Yes
4030442	129641006	Chronic benign neutropenia of childhood	-	Yes	-	Yes
4125805	234576008	Chronic familial neutropenia	-	Yes	-	Yes
4030443	129643009	Chronic hypoplastic neutropenia	-	Yes	-	Yes
4032353	129642004	Chronic idiopathic immunoneutropenia in adult	-	Yes	-	Yes
4095623	248693006	Chronic idiopathic neutropenia	-	Yes	-	Yes
434895	89655007	Congenital neutropenia	-	Yes	-	Yes
36678497	775909002	Congenital neutropenia, myelofibrosis, nephromegaly syndrome	-	Yes	-	Yes
432297	191347008	Cyclical neutropenia	-	Yes	-	Yes
4232037	350353007	De Vaal's syndrome	-	Yes	-	Yes
4211401	56918001	Dose-related drug-induced neutropenia	-	-	Yes	Yes
432289	47318007	Drug-induced neutropenia	-	-	Yes	Yes
1340411	OMOP5166066	Exacerbation of neutropenia	-	-	Yes	Yes
4085181	247860002	Familial neutropenia	-	Yes	-	Yes
4250734	409089005	Febrile neutropenia	-	-	Yes	Yes
4276649	65623009	Immune neutropenia	-	Yes	-	Yes
4300659	78378009	Isoimmune neutropenia	-	Yes	-	Yes
36674945	770942003	Kostmann syndrome	-	Yes	-	Yes
4197402	80255009	Maternal transfer neutropenia	-	Yes	-	Yes
4121120	234424007	Metabolic neutropenia	-	Yes	-	Yes

4093513	24974008	Myelokathexis	-	Yes	-	Yes
42514077	9991/3-NULL	Neoplasm defined only by histology: Refractory neutropenia	-	Yes	-	Yes
320073	165517008	Neutropenia	-	-	Yes	Yes
4211956	416729007	Neutropenia associated with AIDS	-	Yes	-	Yes
4008700	111585004	Neutropenia associated with autoimmune disease	-	Yes	-	Yes
432589	46359005	Neutropenia associated with infectious disease	-	-	Yes	Yes
35624757	767658000	Neutropenia due to and following chemotherapy	-	Yes	-	Yes
4143354	267540007	Neutropenia due to irradiation	-	Yes	-	Yes
4171104	276576000	Neutropenia of the small for gestational age baby	-	Yes	-	Yes
42596532	3,12901E+14	Neutropenia with degenerative left shift	-	Yes	-	Yes
4219985	41814009	Neutropenia with dysgranulopoiesis	-	Yes	-	Yes
37117238	725137007	Neutropenia, monocytopenia, deafness syndrome	-	Yes	-	Yes
4119158	303011007	Neutropenic disorder	-	-	Yes	Yes
40484176	443980004	Neutropenic sepsis	-	-	Yes	Yes
4190716	3902000	Non dose-related drug-induced neutropenia	-	-	Yes	Yes
604592	1003381002	Onycho-tricho-dysplasia neutropenia syndrome	-	Yes	-	Yes
45766061	703148008	Periodontitis associated with chronic familial neutropenia	-	Yes	-	Yes
46273778	709608008	Periodontitis co-occurrent with acquired neutropenia	-	Yes	-	Yes
46272545	710927004	Periodontitis co-occurrent with cyclical neutropenia	-	Yes	-	Yes
46272544	710926008	Periodontitis co-occurrent with familial neutropenia	-	Yes	-	Yes
46273559	709535007	Periodontitis co-occurrent with infantile genetic agranulocytosis	-	Yes	-	Yes
36675666	772126000	Poikiloderma with neutropenia	-	Yes	-	Yes
4101125	191338000	Primary splenic neutropenia	-	Yes	-	Yes
36715585	721303001	Refractory neutropenia	-	-	Yes	Yes
4006469	111584000	Reticular dysgenesis	-	Yes	-	Yes

4230679	351287008	Reticular dysgenesis with congenital aleukocytosis	-	Yes	-	Yes
4097998	190996002	Severe combined immunodeficiency with reticular dysgenesis	-	Yes	-	Yes
4232178	89454001	Shwachman syndrome	-	Yes	-	Yes
4135712	32092008	Toxic neutropenia	-	-	Yes	Yes
439149	55444004	Transient neonatal neutropenia	-	Yes	-	Yes
36716753	722925004	Transient neonatal neutropenia due to congenital viral infection	-	Yes	-	Yes
36716754	722926003	Transient neonatal neutropenia due to neonatal bacterial sepsis	-	Yes	-	Yes
37204236	782759001	X-linked dyserythropoietic anemia with abnormal platelets and neutropenia	-	Yes	-	Yes
4120603	234416002	X-linked hypogammaglobulinemia	-	Yes	-	Yes
36714068	719156006	X-linked intellectual disability with hypogammaglobulinemia and progressive neurological deterioration syndrome	-	Yes	-	Yes
36713881	718882006	X-linked severe congenital neutropenia	-	Yes	-	Yes

Preliminary list of concept definitions for pre-specified conditions related to clozapine use

Treatment-resistant schizophrenia

It will be defined as patients with schizophrenia with prior use of ≥2 non-clozapine antipsychotics before starting clozapine. In terms of concepts:

- concept id of schizophrenia: 435783
- prior use of ≥2 different antipsychotics (e.g., Olanzapine 785788, Risperidone 735979)

The treatment-resistant schizophrenia population will include patients as described above, encompassing a wide range of associated psychological and behavioural symptoms commonly observed in this population.

Psychotic disorders in Parkinson's disease (PD)

Individuals with PD diagnosis will be identified and will be checked for a co-occurring diagnosis of psychotic symptoms or schizophrenia within the observation period. In terms of concepts:

- Parkinson's disease: 381270
- Parkinson's disease with psychotic symptoms: 36714473
- Hallucinations in PD: 433031
- Delusions in PD: 444401
- Schizophrenia:435783

Delirium in Parkinson's disease (PD)

Individuals with PD diagnosis will be identified and will be checked for a co-occurring diagnosis of delirium within the observation period. In terms of concepts:

- Parkinson's disease: 381270
- Delirium: 373995

Suicidal or aggressive behaviour in patients with schizophrenia

Individuals with schizophrenia diagnosis will be identified and will be checked for a suicidal or aggressive behaviour code within the observation period. In terms of concepts:

- Schizophrenia: 435783
- Suicide: 440925
- Suicide attempt: 4219484
- Suicide ideation: 4216115, 37399733, 600767, 4037303, 4021339, 4021336
- Aggressive behaviour: 4266361, 36714069

16.2 Appendix II: ENCePP checklist for study protocols

Study title:
DARWIN EU® - Clozapine and the incidence of agranulocytosis over time

EU PAS Register® number: EUPAS1000000549
Study reference number (if applicable): P4-C1-001

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

Comments:

Section 2: Research question	**Yes**	**No**	**N/A**	**Section Number**
2.1 Does the formulation of the research question and objectives clearly explain:	☒	☐	☐	
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)	☒	☐	☐	8
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?	☒	☐	☐	

Comments:

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

Section 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:				
	4.2.1 Study time period	☒	☐	☐	8.3
	4.2.2 Age and sex	☒	☐	☐	8.6
	4.2.3 Country of origin	☒	☐	☐	8.2
	4.2.4 Disease/indication	☒	☐	☐	8.6
	4.2.5 Duration of follow-up	☒	☐	☐	8.4
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)	☒	☐	☐	

Section 5: Exposure definition and measurement		**Yes**	**No**	**N/A**	**Section Number**
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)	☐	☐	☒	
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)	☒	☐	☐	8.2, 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)	☒	☐	☐	8.2, 8.6
	9.1.3 Covariates and other characteristics?	☒	☐	☐	8.2, 8.6
9.2	Does the protocol describe the information available from the data source(s) on:				8.2, 8.6
	9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	8.2, 8.6
	9.2.2 Outcomes? (e.g. date of occurrence, multiple events, severity measures related to event)	☒	☐	☐	8.2, 8.6
	9.2.3 Covariates and other characteristics? (e.g., age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	☒	☐	☐	8.2, 8.6
9.3	Is a coding system described for:				
	9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☒	☐	☐	8.6
	9.3.2 Outcomes? (e.g., International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	☒	☐	☐	8.6
	9.3.3 Covariates and other characteristics?	☒	☐	☐	8.6
9.4	Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☐	☐	☒	

Comments:

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

Section 10: Analysis plan		**Yes**	**No**	**N/A**	**Section Number**
10.7	Does the plan describe methods for handling missing data?	☐	☐	☒	
10.8	Are relevant sensitivity analyses described?	☐	☐	☒	

Comments:

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

Comments:

Section 12: Limitations		**Yes**	**No**	**N/A**	**Section Number**
12.1	Does the protocol discuss the impact on the study results of:				
	12.1.1 Selection bias?	☒	☐	☐	
	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: Dina Vojinovic

Date: 18th March 2025

Signature: