Study Protocol P2 C1-006

27/10/2023

Version 2.1

TABLE OF CONTENTS

1 Contents

DOCUMENT HISTORY

Version	Date	Description
V1.0	20/07/2023	**First submission to EMA**
V2.0	18/08/2023	**Update incorporating EMA feedback after approval**
V2.1	27-10-2023	**EUPAS registration number added**

Study Title	DARWIN EU® - Treatment patterns of drugs used in adult and paediatric population with systemic lupus erythematosus
Protocol version identifier	V2.0
Date of last version of protocol	20/07/2023
EU PAS register number	EUPAS106436
Active substance	N/A
Medicinal product	N/A
Research question and objectives	The overall objective of this study is to characterise paediatric and adult patients with systemic lupus erythematosus (SLE) diagnosed in the period 2013-2022. The specific objectives of this study are: 1. To describe demographic and clinical characteristics of paediatric patients with SLE at the time of diagnosis. 2. To describe demographic and clinical characteristics of adult patients with SLE at the time of diagnosis. 3. To describe the treatment patterns from diagnosis until end of follow up for paediatric patients newly diagnosed with SLE. 4. To describe the treatment patterns from diagnosis until end of follow up for adult patients newly diagnosed with SLE. 5. To describe the use of treatment (including treatment duration, cumulative dose, number of repeated prescriptions for each medication) initiated after a diagnosis of SLE in paediatric patients. 6. To describe the use of treatment (including treatment duration, cumulative dose, number of repeated prescriptions for each medication) initiated after a diagnosis of SLE in adult patients. All results will be reported by country/database, overall and stratified by age and sex when possible.
Country(-ies) of study	France, Germany, Spain, United Kingdom
Author	Eng Hooi (Cheryl) Tan (cheryl.tan@ndorms.ox.ac.uk); Daniel Prieto-Alhambra (d.prietoalhambra@darwin-eu.org)

LIST OF ABBREVIATIONS

Acronyms/terms	Description
ADHD	Attention Deficit Hyperactivity Disorder
CDM	Common Data Model
CDWBordeaux	Clinical Data Warehouse of Bordeaux University Hospital
COPD	Chronic obstructive pulmonary disease
CPRD	Clinical Practice Research Datalink
DA	Disease Analyzer
DARWIN EU®	Data Analysis and Real World Interrogation Network
DMARD	Disease-modifying antirheumatic drug
DOI	Declaration Of Interests
DRE	Digital Research Environment
DUS	Drug utilisation study
EHR	Electronic Health Records
EMA	European Medicines Agency
EULAR	European Alliance of Associations for Rheumatology
GERD	Gastro-esophageal reflux disease
GP	General Practitioner
IMASIS	Institut Municipal Assistència Sanitària Information System
OMOP	Observational Medical Outcomes Partnership
PCT	Primary Care Teams
PSMar	Parc Salut Mar
SIDIAP	Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària
SLE	Systemic lupus erythematosus
SNOMED	Systematized Nomenclature of Medicine

1. TITLE

DARWIN EU® - Treatment patterns of drugs used in adult and paediatric population with systemic lupus erythematosus

2. RESPONSIBLE PARTIES – STUDY TEAM

Study team Role	Names	Organisation
Study Project Manager/Principal Investigator	Eng Hooi (Cheryl) Tan Daniel Prieto-Alhambra	University of Oxford University of Oxford/Erasmus MC
Epidemiologist	Eng Hooi (Cheryl) Tan Daniel Prieto-Alhambra	University of Oxford University of Oxford/Erasmus MC
Clinical Domain Expert	Daniel Prieto-Alhambra	University of Oxford/Erasmus MC
Data Analysts/statisticians	Martí Català Sabaté Mike Du	University of Oxford University of Oxford
Data Partner*	**Names**	**Organisation – Database**
Local Study Coordinator/Data Analyst	James Brash Hanne van Ballegooijen Núria Mercadé Talita Duarte-Salles Miguel-Angel Mayer Angela Leis Juan Manuel Ramirez Romain Griffier Antonella Delmestri Hezekiah Omulo Wai Yi (Teen) Man	IQVIA - DA Germany IQVIA - DA Germany IDIAPJGol - SIDIAP IDIAPJGol - SIDIAP PSMAR - IMASIS PSMAR - IMASIS PSMAR - IMASIS University of Bordeaux - CDWBordeaux University of Oxford – CPRD GOLD University of Oxford – CPRD GOLD University of Oxford – CPRD GOLD

*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 (STAND ALONE SUMMARY OF THE STUDY PROTOCOL)

Title

DARWIN EU® - Treatment patterns of drugs used in adult and paediatric population with systemic lupus erythematosus

Rationale and Background

Systemic lupus erythematosus (SLE) is a multisystem autoimmune disorder of connective tissue characterized by autoantibodies that target nuclear antigens, remissions and flares, and a highly variable clinical presentation, disease course, and prognosis. The disease course is more severe in childhood-onset compared to adult-onset SLE, with higher prevalence of morbidities and lower survival rates.

Therefore, to review new drug applications in this disease area, it would be important for the European Medicines Agency (EMA) to understand the current clinical practice of treating SLE in paediatric population and differences with the treatment in adult population.

Research question and Objectives

The overall objective of this study is to characterise paediatric and adult patients with SLE diagnosed in the period 2013-2022, and to study the treatments they received in this same period.

The specific objectives of this study are:

1. To describe demographic and clinical characteristics of paediatric patients with SLE at the time of diagnosis.
2. To describe demographic and clinical characteristics of adult patients with SLE at the time of diagnosis.
3. To describe the treatment patterns from diagnosis until end of follow up for paediatric patients newly diagnosed with SLE.
4. To describe the treatment patterns from diagnosis until end of follow up for adult patients newly diagnosed with SLE.
5. To describe the use of treatment (including treatment duration, cumulative dose, number of repeated prescriptions for each medication) initiated after a diagnosis of SLE in paediatric patients.
6. To describe the use of treatment (including treatment duration, cumulative dose, number of repeated prescriptions for each medication) initiated after a diagnosis of SLE in adult patients.

All results will be reported by country/database, overall and stratified by age and sex when possible.

Research Methods

Study design

A retrospective cohort study of all patients newly diagnosed with SLE will be conducted. For the description of each treatment objective, a new drug user cohort will be used to characterise patient-level SLE drug utilisation.

Population

The source population will include all individuals eligible in the database between the patient selection period, which is 01/01/2013 and 180 days prior to the end of available data in each database. Eligibility criteria will be applied for each study objective:

New diagnosis cohort

- First diagnosis of SLE in database during patient selection period

- At least 365 days of prior history available before date of first SLE diagnosis

In addition to the criteria above, the paediatric new diagnosis cohort (Cohort 1, Objectives 1 and 3) is aged < 18 years at date of first SLE diagnosis; the adult new diagnosis cohort (Cohort 2, Objectives 2 and 4) is aged ≥ 18 years at date of first SLE diagnosis.

New user cohort

- First diagnosis of SLE in database during patient selection period

- At least 365 days of prior history available before date of first SLE diagnosis

- Initiation of SLE treatment of interest after first diagnosis of SLE

- At least 365 days of washout period at treatment ingredient level prior to date of initiation of SLE treatment of interest

In addition to the criteria above, the paediatric new user cohort (Cohort 3, Objective 5) is aged < 18 years at date of first SLE diagnosis; the adult new user cohort (Cohort 4, Objective 6) is aged ≥ 18 years at date of first SLE diagnosis.

Variables

The main exposure of interest is the treatment of SLE: treatment/s initiated after new diagnosis of SLE. A pre-specified list of SLE treatments will be generated (objectives 3, 4, 5, and 6).

All co-morbidities and co-medications will be used for large-scale patient characterisation, identified as concept/code and descendants. A separate list of pre-specified co-morbidities and co-medications of interest for patients with SLE will also be described.

Data sources

1. IQVIA Disease Analyzer Germany (IQVIA DA Germany), Germany – primary care and specialist data
2. Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària (SIDIAP), Spain – primary care data linked with hospital discharge.
3. Institut Municipal Assistencia Sanitaria Information System (IMASIS), Spain – hospital data
4. Clinical Data Warehouse of Bordeaux University Hospital (CDW Bordeaux), France – hospital data
5. Clinical Practice Research Datalink (CPRD) GOLD, United Kingdom (UK) – primary care data

Sample size

No sample size has been calculated as this is a descriptive Disease Epidemiology Study where we are interested in the characteristics of all incident SLE patients. Based on a preliminary feasibility assessment the expected number of SLE patients in the included databases for this study will be approximately 20,790. (SIDIAP 8,040; CPRD GOLD 5,668; IQVIA DA Germany 5,179; CDW Bordeaux 1,137; IMASIS 766).

Data analyses

Large-scale patient-level characterisation will be conducted (objectives 1 and 2). Medical condition and medication use history will be reported at any time and 365 days prior to index date, respectively.

The number and percentage of patients receiving each of a pre-specified list of SLE treatments and treatment combinations (objectives 3 and 4) will be described per calendar year. Additionally, sunburst plots and Sankey diagrams will be used to describe treatment patterns and sequences over time (objectives 3 and 4).

For the new user cohort (objectives 5 and 6), the index date is the initiation of SLE treatment after SLE diagnosis. Treatment duration, initial dose/strength, cumulative dose, number of prescriptions will be estimated for new users of each SLE treatments at the ingredient level.

For all continuous variables, mean with standard deviation and median with interquartile range will be reported. For all categorical analyses, number and percentages will be reported. A minimum cell count of 5

will be used when reporting results, with any smaller counts reported as "<5". All analyses will be reported by country/database, overall and stratified by age and sex when possible (minimum cell count reached). Additionally, to capture treatments availability and changes over time, sunburst plots, Sankey diagrams will be further stratified by 5-year periods (2013-2017 and 2018-2022).

4. AMENDMENTS AND UPDATES

Number	Date	Section of study protocol	Amendment or update	Reason
Version 2.0	18/08/2023	All	Update	Update following EMA's assessment
Version 2.1	27/10/2023	Document history	Update	EUPAS registration number added

5. MILESTONES

STUDY SPECIFIC DELIVERABLE	TIMELINE
Draft Study Protocol	20 July 2023
Final Study Protocol	18 August 2023
Registration in EUPAS register	
Creation of Analytical code	August 2023
Execution of Analytical Code on the data	September/October 2023
Interim Study Report (if applicable)	Not applicable
Draft Study Report	October 2023
Final Study Report	October/November 2023

6. RATIONALE AND BACKGROUND

Systemic SLE erythematosus (SLE) is a multisystem autoimmune disorder of connective tissue characterized by autoantibodies that target nuclear antigens, remissions and flares, and a highly variable clinical presentation, disease course, and prognosis. The disease course is more severe in childhood-onset compared to adult-onset SLE, with higher prevalence of morbidity and lower survival rates (1).

The European Alliance of Associations for Rheumatology (EULAR) guidelines recommend hydroxychloroquine as first line treatment of adult SLE (2). Glucocorticoids provide rapid symptomatic relief, but long-term safety concerns limit their use. The guidelines also recommend the addition of a disease-modifying antirheumatic drug (DMARD) or immunosuppressant to control disease flares and facilitate glucocorticoid tapering (2). Examples of DMARDs often used are methotrexate, azathioprine, mycophenolate mofetil, or cyclophosphamide. Biological agents such as belimumab should be considered in extrarenal disease, while rituximab might be used off-label in patients with refractory or severe disease, as a result of negative clinical trial outcomes in terms of efficacy (2, 3). Calcineurin inhibitors are recommended as monotherapy or in combination with mycophenolate mofetil in patients at high risk of renal involvement (2). In contrast to adult SLE, there is limited good quality evidence on the treatment of childhood SLE. A European-wide panel of 16 paediatric rheumatologists recommended routine treatment using hydroxychloroquine (4) with the addition of DMARDs if disease cannot be adequately controlled with hydroxychloroquine and corticosteroid tapering. Rituximab was used in a limited number of cases (4).

Therefore, to review new drug applications, it would be important for the European Medicines Agency (EMA) to understand the current clinical practice of treating SLE in paediatric population and differences with the treatment in adult population.

7. RESEARCH QUESTION AND OBJECTIVES

The overall objective of this study is to characterise paediatric and adult patients with SLE diagnosed in the period 2013-2022.

The specific objectives of this study are:

1. To describe demographic and clinical characteristics of paediatric patients with SLE at the time of diagnosis.
2. To describe demographic and clinical characteristics of adult patients with SLE at the time of diagnosis.
3. To describe the treatment patterns from diagnosis until end of follow up for paediatric patients newly diagnosed with SLE.
4. To describe the treatment patterns from diagnosis until end of follow up for adult patients newly diagnosed with SLE.
5. To describe the use of treatment (including treatment duration, cumulative dose, number of repeated prescriptions for each medication) initiated after a diagnosis of SLE in paediatric patients.
6. To describe the use of treatment (including treatment duration, cumulative dose, number of repeated prescriptions for each medication) initiated after a diagnosis of SLE in adult patients.

All results will be reported by country/database, overall and stratified by age and sex when possible.

Table 1: Primary and secondary research questions and objective

Objective:	1. To describe demographic and clinical characteristics of paediatric patients with SLE at the time of diagnosis. 2. To describe demographic and clinical characteristics of adult patients with SLE at the time of diagnosis. 3. To describe the treatment patterns from diagnosis until end of follow up for paediatric patients newly diagnosed with SLE. 4. To describe the treatment patterns from diagnosis until end of follow up for adult patients newly diagnosed with SLE. 5. To describe the use of treatment (including treatment duration, cumulative dose, number of repeated prescriptions for each medication) initiated after a diagnosis of SLE in paediatric patients. 6. To describe the use of treatment (including treatment duration, cumulative dose, number of repeated prescriptions for each medication) initiated after a diagnosis of SLE in adult patients.
Hypothesis:	N/A
Population *(mention key inclusion-exclusion criteria):*	All individuals with a first diagnosis of SLE identified in the database between the patient selection period, which is 01/01/2013 and 180 days prior to the end of available data in each database. Additional eligibility criteria will be applied for each study objective: **New diagnosis cohort** - First diagnosis of SLE in database during patient selection period

	- At least 365 days of prior history available before date of first SLE diagnosis In addition to the criteria above, the paediatric new diagnosis cohort (Cohort 1, Objectives 1 and 3) is aged < 18 years at date of first SLE diagnosis; the adult new diagnosis cohort (Cohort 2, Objectives 2 and 4) is aged ≥ 18 years at date of first SLE diagnosis. **New user cohort** - First diagnosis of SLE in database during patient selection period - At least 365 days of prior history available before date of first SLE diagnosis - Initiation of SLE treatment of interest after first diagnosis of SLE - At least 365 days of washout period at treatment ingredient level prior to date of initiation of SLE treatment of interest In addition to the criteria above, the paediatric new user cohort (Cohort 3, Objective 5) is aged < 18 years at date of first SLE diagnosis; the adult new user cohort (Cohort 4, Objective 6) is aged ≥ 18 years at date of first SLE diagnosis.
Exposure:	SLE treatments [hydroxychloroquine, systemic glucocorticoids, methotrexate, azathioprine, calcineurin inhibitors (tacrolimus, cyclosporine, voclosporin), mycophenolate, cyclophosphamide, rituximab, belimumab]
Comparator:	N/A
Outcome:	N/A
Time *(when follow up begins and ends):*	For objectives 1 to 4, follow-up will start from date of first SLE diagnosis until the earliest of the following: 1) loss to follow-up, 2) end of data availability, or 3) date of death. For objectives 5 and 6, follow-up will start from date of first SLE treatment after SLE diagnosis until the earliest of the following: 1) loss to follow-up, 2) end of data availability, or 3) date of death.
Setting:	Inpatient and outpatient setting from 5 databases in 4 European countries.
Main measure of effect:	Proportions of patients on treatment types and sequences, patient-level drug utilisation.

8. RESEARCH METHODS

8.1 Study Type and Study Design

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

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

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

8.2 Study Setting and Data Sources

This study will be conducted using routinely collected health data from 5 databases in 4 European countries. All databases were previously mapped to the OMOP CDM.

Data sources:

1. IQVIA Disease Analyzer Germany (IQVIA DA Germany), Germany
2. Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària (SIDIAP), Spain
3. Institut Municipal Assistencia Sanitaria Information System (IMASIS), Spain
4. Clinical Data Warehouse of Bordeaux University Hospital (CDWBordeaux), France
5. Clinical Practice Research Datalink (CPRD) GOLD, United Kingdom (UK)

We selected 5 out of the 10 databases onboarded in DARWIN EU® in 2022. The selection of databases for this study was performed based on data reliability and relevance for the proposed research question, as well as sufficient coverage of the paediatric population. The selected databases fulfil the criteria required for a patient-level characterisation study allowing for large-scale characterisation, while covering different settings and regions of Europe.

Complete hospital-based SLE treatment data (needed for objectives 3, 4, 5, and 6) will be available in all databases except CPRD (UK) and SIDIAP (Spain). A proportion of SIDIAP database will have linkage to hospital data to allow for more accurate characterisation, but data on inpatient treatments is not available. In turn, any potential outpatient therapies will be captured in these primary care datasets. In IMASIS, there were small numbers of paediatric patients with SLE, therefore the objectives associated with this population cannot be answered by this database.

Detailed information on the selected data sources and their ability to answer the study research questions are described in **Table 3**.

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	Data lock for the last update	Ability to answer study objectives
DE	IQVIA DA Germany	Covers primary care and outpatient specialist setting with information on SLE diagnoses and treatment.	Primary care and outpatient specialist care	EHR	8.5 million	31/03/2023	1 to 6
ES	SIDIAP	Covers primary care setting with a proportion with hospital linkage, data on SLE diagnoses.	Primary care with hospital linkage	EHR	5.8 million	31/12/2022	1 to 6
ES	IMASIS	Covers secondary care setting, database has information on SLE diagnosis and treatments in the in- and outpatient settings	Secondary care (in and outpatients)	EHR	0.6 million	13/05/2023	2, 4, 6
FR	CDWBordeaux	Covers secondary care setting, database has information on SLE diagnosis and in-hospital treatments	Secondary care (in and outpatients)	EHR	1.9 million	10/04/2023	1 to 6
UK	CPRD GOLD	Covers primary care setting, database has information on SLE diagnosis and treatments	Primary care	EHR	3.1 million	04/07/2022	1 to 6

DE = Germany, ES = Spain, FR = France, NL = The Netherlands, UK = United Kingdom, SIDIAP = Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària, IMASIS = Institut Municipal Assistencia Sanitaria Information System, DA = Disease Analyzer, CDWBordeaux = Clinical Data Warehouse of Bordeaux University Hospital, CPRD = Clinical Practice Research Datalink.

IQVIA Disease Analyser (DA) Germany, Germany

DA Germany is collected from extracts of patient management software used by GPs and specialists practicing in ambulatory care settings (5). Data coverage includes more than 34M distinct person records out of at total population of 80M (42.5%) in the country and collected from 2,734 providers. Patient visiting more than one provider are not cross identified for data protection reasons and therefore recorded as separate in the system. Dates of service include from 1992 through present. Observation time is defined by the first and last consultation dates. Germany has no mandatory GP system and patient have free choice of specialist. As a result, data are collected from visits to 28.8% General, 13.4% Orthopaedic Surgery, 11.8% Otolaryngology, 11.2% Dermatology, 7.7% Obstetrics/Gynaecology, 6.2% various Neurology and Psychiatry 7.0% Paediatric, 4.6% Urology, 3.7% Cardiology, 3.5% Gastroenterology, 1.5% Pulmonary and 0.7% Rheumatology practices. Drugs are recorded as prescriptions of marketed products. No registration or approval is required for drug utilisation studies.

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

SIDIAP is collected from EHR records of patients receiving primary care delivered through Primary Care Teams, consisting of GPs, nurses and non-clinical staff (6). The Catalan Health Institute manages 328 out of

370 such Primary Care Teams with a coverage of 5.8M patients, out of 7.8M people in the Catalan population (74%). The database started to collect data in 2006. The mean follow-up is 15 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.

Institut Municipal Assistència Sanitària Information System (IMASIS), Spain

The Institut Municipal Assistència Sanitària Information System (IMASIS) is the Electronic Health Record (EHR) system of Parc de Salut Mar Barcelona (PSMar) which is a complete healthcare services organisation (7). Currently, this information system includes and shares the clinical information of two general hospitals (Hospital del Mar and Hospital de l'Esperança), one mental health care centre (Centre Dr. Emili Mira) and one social-healthcare centre (Centre Fòrum) including emergency room settings, which are offering specific and different services in the Barcelona city area (Spain). At present, IMASIS includes clinical information more than 1 million patients with at least one diagnosis and who have used the services of this healthcare system since 1990 and from different settings such as admissions, outpatients, emergency room and major ambulatory surgery. The diagnoses are coded using The International Classification of Diseases ICD-9-CM and ICD-10-CM. The average follow-up period per patient in years is 6.37 (SD±6.82). IMASIS-2 is the anonymized relational database of IMASIS which is used for mapping to OMOP including additional sources of information such as the Tumours Registry.

Clinical Data Warehouse of Bordeaux University Hospital (CDWBordeaux), France

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

Clinical Practice Research Datalink GOLD, United Kingdom

The Clinical Practice Research Datalink (CPRD) is a governmental, not-for-profit research service, jointly funded by the National Institute for Health and Care Research and the Medicines and Healthcare products Regulatory Agency, a part of the Department of Health, United Kingdom (UK) (https://cprd.com). CPRD GOLD (8) 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 21 million patients, including 3.1 million currently registered patients (9). Access to CPRD GOLD data requires approval via the Research Data Governance Process.

8.3 Study Period

The study period will be from 01/01/2013 to end of available data in each of the data sources (see **Table 3** for more details).

8.4 Follow-up

For objectives 1 to 4, follow-up will start from date of first SLE diagnosis until the earliest of the following: 1) loss to follow-up, 2) end of data availability, or 3) date of death.

For objectives 5 and 6, follow-up will start from date of first SLE treatment after SLE diagnosis until the earliest of the following: 1) loss to follow-up, 2) end of data availability, or 3) date of death.

Table 4: Operational Definition of Time 0 (index date) and other primary time anchors

Study population name(s)	Time Anchor Description	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 diagnosis cohort (objectives 1 to 4)	Date of first SLE diagnosis	Single entry	Incident	Any time prior to SLE diagnosis	IP, OP, OT	SNOMED	Any	SLE diagnosis	N/A	N/A
New user cohort (objectives 5 and 6)	Date of initiation of SLE treatment after first SLE diagnosis	Single entry	Incident	365 days prior to SLE treatment	IP, OP, OT	RxNorm	N/A	SLE treatment after first SLE diagnosis	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 individuals with a first diagnosis of SLE identified in the database during the patient selection period, which is between 01/01/2013 and 180 days prior to the end of available data in each database. The index dates are defined in **Table 4**.

For this study, patients will be identified based on a record indicating a diagnosis of SLE. Conditions in the OMOP CDM use the Systematized Nomenclature of Medicine (SNOMED) as the standard vocabulary for diagnosis codes. A preliminary code list is provided in **Appendix 1**. We will use a narrow definition of SLE diagnosis codes for the main analysis.

The following eligibility criteria will be applied for each study objective (see Inclusion criteria in **Table 5**):

Objectives 1 and 3

Cohort 1 – New diagnosis cohort (paediatric)

- Aged < 18 years

- First diagnosis of SLE in database during patient selection period

- At least 365 days of prior history available before date of first SLE diagnosis

Objectives 2 and 4

Cohort 2 – New diagnosis cohort (adult)

- Aged ≥ 18 years

- First diagnosis of SLE in database during patient selection period

- At least 365 days of prior history available before date of first SLE diagnosis

Objective 5

Cohort 3 – New user cohort (paediatric)

- Aged < 18 years

- First diagnosis of SLE in database during patient selection period

- At least 365 days of prior history available before date of first SLE diagnosis

- Initiation of SLE treatment of interest after first diagnosis of SLE

- At least 365 days of washout period at treatment ingredient level prior to date of initiation of SLE treatment of interest

Objective 6

Cohort 4 – New user cohort (adult)

- Aged ≥ 18 years

- First diagnosis of SLE in database during patient selection period

- At least 365 days of prior history available before date of first SLE diagnosis

- Initiation of SLE treatment of interest after first diagnosis of SLE

- At least 365 days of washout period at treatment ingredient level prior to date of initiation of SLE treatment of interest

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
Prior database history of 365 days (objectives 1 to 4)	Study participants will be required to have 365 days of prior history observed before contributing observation time	After index date is determined	365 days	IP, OP, OT	N/A	N/A	All study participants with first SLE diagnosis	N/A	N/A
New user of SLE treatment (objectives 5 and 6)	Only participants with no use of SLE treatment at the ingredient level in the 365 days prior to initiation of SLE treatment (index date) will be included	After index date is determined	365 days	IP, OP, OT	RxNorm	N/A	All study participants with first SLE diagnosis	N/A	N/A

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

8.6 Variables

8.6.1. Exposure/s

SLE treatments will include hydroxychloroquine, systemic glucocorticoids, methotrexate, azathioprine, calcineurin inhibitors (tacrolimus, cyclosporine, voclosporin), mycophenolate, cyclophosphamide, rituximab, belimumab. For the new diagnosis cohort, no washout period will be applied. Treatment patterns of SLE drugs of interest will be described after first diagnosis of SLE. For the new user cohort, washout period of 365 days at the ingredient level will be applied after first diagnosis of SLE, therefore it will not include patients who are prevalent users of treatment, if there are treatments initiated before diagnosis of SLE is recorded. Please see **Table 6** for definitions of exposure and **Appendix 1 Table 2** for a preliminary list of codes to identify these treatments.

8.6.2. Outcome/s

N/A

Table 6. Operational Definitions of Exposure

Exposure group name(s)	Details	Washout window	Assessment Window	Care Setting[1]	Code Type[2]	Diagnosis position[3]	Applied to study populations:	Incident with respect to…	Measurement characteristics/ validation	Source of algorithm
SLE treatment	Preliminary code lists in Appendix 1 Table 2	n/a	[0, censor]	IP and OP	RxNorm	n/a	New diagnosis cohort	n/a	n/a	n/a
SLE treatment	Preliminary code lists in Appendix 1 Table 2	365 days	[0, censor]	IP and OP	RxNorm	n/a	New user cohort	SLE diagnosis	n/a	n/a

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

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

Other covariates

Age at SLE diagnosis will be calculated. The following age grouping will be used: 0-4; 5-12; 13-17; 18-39; 40-49; 50-59; 60-69; 70 and over. The sex (male/ female) of study participants will also be reported.

All co-morbidities and co-medications recorded prior and at index date will be used for large-scale patient characterisation, identified as concept/code and descendants (**Table 7**). Additionally, a list of pre-specified co-morbidities and co-medications relevant for patients with SLE will be described. These will include:

- Medical History: Asthma, Cardiovascular disease, Chronic obstructive pulmonary disease (COPD), Chronic Liver disease, Crohn's Disease, Diabetes mellitus, Gastro-esophageal reflux disease (GERD), GI-Bleeding, Human Immunodeficiency Virus (HIV), Hyperlipidemia, Hypertension, Obesity, Osteoarthritis, Pneumonia, Psoriasis, Renal impairment, Ulcerative Colitis, Urinary Tract infection, Viral Hepatitis, Visual system disorder, Schizophrenia, Dementia, Parkinson, Depressive disorder, Anxiety, Attention Deficit Hyperactivity Disorder (ADHD), Any cancer except non-melanoma skin cancer.
- Medication use: Agents acting on the renin-angiotensin system, Antibacterials for systemic use, Antidepressants, Antiepileptics, Antiinflammatory and antirheumatic products, Antineoplastic agents, Antipsoriatics, Antithrombotic agents, Beta blocking agents, Calcium channel blockers, Diuretics, Drugs for acid related disorders, Drugs for obstructive airway diseases, Drugs used in diabetes, Immunosuppressants, Lipid modifying agents, Opioids, Psycholeptics, Psychostimulants, agents used for ADHD and nootropics .

Confounders

N/A

Effect modifiers

N/A

Table 7. Operational Definitions of Covariates

Characteristic	Details	Type of variable	Assessment window	Care Settings[1]	Code Type	Diagnosis Position	Applied to study populations:	Measurement characteristics /validation	Source for algorithm
Co-morbidities	Large-scale patient-level characterisation with regard to underlying comorbidities	Counts	At index date (ID); Before ID: 30 to 1 day, 365 to 31 days at any time and up to 366 days; After ID: 1 to 30 days, 31 to 90 days, 91 to 180 days, 181 to 365 days, 366+ days	OP, IP, OT	SNOMED	N/A	N/A	N/A	N/A
Concomitant medication	Large-scale patient-level characterisation with regard to use of concomitant drugs	Counts	At index date (ID); Before ID: 30 to 1 day, 365 to 31 days at any time and up to 366 days; After ID: 1 to 30 days, 31 to 90 days, 91 to 180 days, 181 to 365 days, 366+ days	OP, IP, OT	RxNorm	N/A	N/A	N/A	N/A

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

8.7 Study size

No sample size has been calculated as this is a descriptive Disease Epidemiology Study with the objective of characterising all available incident SLE patients. Based on a preliminary feasibility assessment the expected number of SLE records in the included databases for this study will be approximately 20,790 (SIDIAP 8,040; CPRD GOLD 5,668; IQVIA DA Germany 5,179; CDWBordeaux 1,137; IMASIS 766). Please note that this number is based on the overall number of SLE patients in each database with at least 365 days of prior observation at time of diagnosis and no filter by study period.

8.8 Analysis

Table 8. Description of Study Types and Type of analysis

STUDY TYPE	STUDY CLASSIFICATION	TYPE OF ANALYSIS
Patient-level characterisation and DUS	Off-the-shelf (C1)	- Large-scale characterisation - Patient-level treatment patterns - Patient-level drug utilisation

8.8.1 Federated Network Analyses

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

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

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

8.8.2 Patient privacy protection

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

8.8.3 Statistical model specification and assumptions of the analytical approach considered

R-packages

We will use R packages for the patient-level characterization of demographics and clinical characteristics; "DrugUtilisation" (https://github.com/darwin-eu/DrugUtilisation) for the patient-level drug utilisation analyses including patient-level characterisation and treatment duration, cumulative dose, number of repeated prescriptions for each medication; "TreatmentPatterns" (https://github.com/darwin-eu-

dev/TreatmentPatterns) for the patient-level characterisation of treatments including combination and sequence of therapy.

Patient-level characterisation

Large-scale patient-level characterisation will be conducted (objectives 1 and 2). Age and sex at time of SLE diagnosis will be described for each of the generated study cohorts. The index date will be the date of the first SLE diagnosis for each patient. Medical condition and medication use history will be assessed for anytime –and up to 365 days before index date, for 365 to 31 days before index date, for 30 to 1 day before index date, and at index date. We will also report medical condition and medication use for 1 to 30, 31 to 90, 91 to 180, 181 to 365 days, and 366 days to anytime post index date. These time windows were defined based on the options currently available in the standard analytical tools that will be used in this project. For the main study report, medical conditions any time prior to index date and medication use 365 days prior to index date will be presented. The other time windows will be available in an interactive dashboard. Co-variates to be presented in a summary baseline characteristics table will be pre-defined as described in section 8.6.3.

Patient-level drug utilisation

The number and percentage of patients receiving each of a pre-specified list of SLE treatments (see Appendix 1) and treatment combinations (objectives 3 and 4) will be described per calendar year. Additionally, sunburst plots and Sankey diagrams will be used to describe treatment patterns and sequences over time (objectives 3 and 4). Sankey diagrams will be censored at end of treatment or end of follow-up as described in section 8.

For the new user cohort (objectives 5 and 6), the index date is the initiation of each SLE treatment after SLE diagnosis. Treatment duration, initial dose/strength, cumulative dose, number of prescriptions will be estimated for new users of SLE treatments at the ingredient level.

Drug exposure calculations

Drug eras will be defined as follows: Exposure starts at date of the first prescription after the first SLE diagnosis. For each prescription, the estimated duration of use is retrieved from the drug exposure table in the CDM. Drug exposure diagnostics will be reviewed for parenteral medications to consider a default duration of the dosing schedule. Subsequent prescriptions for the same drug will be combined into continuous exposed episodes (drug eras) using the following specifications:

Two drug exposures will be merged into one continuous drug era if the distance in days between end of the first exposure and start of the second exposure is ≤ 30 days. The time between the two joined exposures will be considered as exposed to the first era as shown in Figure 1.

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

If two exposures overlap, the overlap time will be considered exposed to the first exposure (Figure 2). No time will be added at the end of the combined drug era to account for the overlap.

If two exposures start at the same date, the overlapping period will be considered exposed to both. We will not consider repetitive exposure. Complex dosing schedule for rituximab will not be considered in constructing drug eras as this medication is off-label for SLE and rarely prescribed. Thus, only the first drug era will be considered for rituximab.

Overlap mode	Schematics	Dose overlap
"first"		d_1
"second"		d_2
"both"		$d_1 + d_2$
"maximum"		$\max(d_1, d_2)$

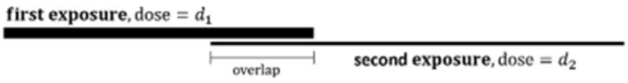

Figure 2. Gap era overlap mode

To construct treatment pathways, various parameters can be defined in the TreatmentPatterns package (Figure 3).

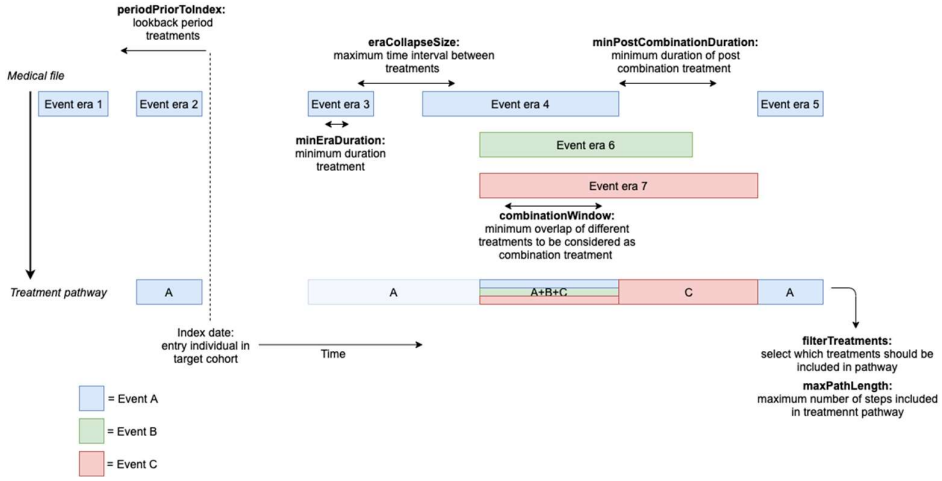

Figure 3. Parameters in TreatmentPatterns package

The following parameters will be defined in this study. The target cohort refers to the specified study population, i.e. patients with first diagnosis of SLE whereas the event(s) refer to treatment(s) of interest. (10)

Individual pathway settings

periodPriorToIndex	The period (number of days) prior to the index date of the target cohort from which treatments should be included	0
minEraDuration	Minimum time (days) an event era should last to be included in the analysis	0
eraCollapseSize	Maximum gap (days) within two eras of the same event cohort which would still allow the eras to be collapsed into one era	30
combinationWindow	Minimum time (days) that two event eras need to overlap to be considered a combination treatment	30
minPostCombinationDuration	Minimum time (days) that an event era before or after a generated combination treatment should last to be included in the pathway as a separate treatment	30
filterTreatments	Select which treatments should be included in pathway first time occurrences of treatments ('First'), remove sequential repeated treatments ('Changes'), all treatments ('All')	First
maxPathLength	Maximum number of treatments included in pathway	5
Aggregate pathway settings		
minCellCount	Minimum number of persons with a specific treatment pathway for the pathway to be included in analysis	5
minCellMethod	Select to completely remove / sequentially adjust (by removing last step as often as necessary) treatment pathways below minCellCount	Adjust
groupCombinations	Select to group all non-fixed combinations in one category 'other' in the sunburst plot	TRUE/10
addNoPaths	Select to include untreated persons without treatment pathway in the sunburst plot	TRUE

For all continuous variables, mean with standard deviation and median with interquartile range will be reported. For all categorical analyses, number and percentages will be reported. A minimum cell count of 5 will be used when reporting results, with any smaller counts reported as "<5". All analyses will be reported by country/database, overall and stratified by age and sex when possible (minimum cell count reached). Additionally, to capture treatments availability and changes over time, sunburst plots, Sankey diagrams will be further stratified by study periods (2013-2017 and 2018-2022).

8.8.4 Sensitivity analysis

We will repeat all statistical analyses for patients with a broad definition of SLE diagnosis codes (Appendix 1 Table 1).

8.9 Evidence synthesis

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

9. DATA MANAGEMENT

9.1 Data management

All databases are mapped to the OMOP common data model. This enables the use of standardised analytics and tools across the network since the structure of the data and the terminology system is harmonised. The OMOP CDM is developed and maintained by the Observational Health Data Sciences and Informatics (OHDSI) initiative and is described in detail on the wiki page of the CDM: https://ohdsi.github.io/CommonDataModel and in The Book of OHDSI: http://book.ohdsi.org

The analytic code for this study will be written in R. Each data partner will execute the study code against their database containing patient-level data and will then return the results set which will only contain aggregated data. The results from each of the contributing data sites will then be combined in tables and figures for the study report.

9.2 Data storage and protection

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

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

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

10. QUALITY CONTROL

General database quality control

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

Study specific quality control

When defining SLE, 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 CDM so as to find potentially relevant codes. The codes returned will be reviewed by two clinical epidemiologists to consider their relevance. In addition, we will run cohort diagnostics to assess the use of different codes across the databases contributing to the study and identify any codes potentially omitted in error. This will allow for a consideration of the validity of the study cohort of patients with SLE in each of the databases, and inform decisions around whether multiple definitions are required.

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 for the SLE treatments.

The study code will be based on two R packages currently being developed to (1) characterise demographic and clinical characteristics, (2) characterise treatment patterns. 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, such as reliability and relevance must be considered. In particular, the identification of SLE patients and the recording of the co-morbidities may vary across databases and while relatively few false positives would be expected (i.e. those recorded with a condition who do not truly have the condition), false negatives (i.e. those with a condition that is not recorded) may be more likely especially for databases without patient-level linkage from primary care to secondary care data. There is scarce data on the validation of the SLE phenotype in

administrative databases in Europe (11, 12). The SLE phenotype used in this study is defined using coding only and does not based upon other clinical data such as symptoms and autoantibody laboratory tests.

In addition, the recording of comorbid conditions and medication use defined for patient characterisation may vary across databases; and in databases with information on SLE treatment, the recording of treatment use may be incomplete. This may occur particularly for primary care databases such as CPRD GOLD without linkage to hospital data. Characterisation of baseline co-morbidities in pre-specified time periods before the index date represent partial prevalence of conditions and and not complete prevalence. To mitigate selection bias, we have also specified a new user design for drug utilisation to exclude prevalent users. However, the definition of incident users is subject to the initial treatment being recorded in the data sources.

12. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

Adverse events/adverse reactions will not be collected or analyzed as part of this evaluation. The nature of this non-interventional evaluation, through the use of secondary data, does not fulfill 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).

13. GOVERNANCE BOARD ASPECTS

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

14. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

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

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

15. OTHER ASPECTS

N/A

16. REFERENCES

1. Smith EMD, Lythgoe H, Midgley A, Beresford MW, Hedrich CM. Juvenile-onset systemic lupus erythematosus: Update on clinical presentation, pathophysiology and treatment options. Clin Immunol. 2019;209:108274.

2. Fanouriakis A, Kostopoulou M, Alunno A, Aringer M, Bajema I, Boletis JN, et al. 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus. Ann Rheum Dis. 2019;78(6):736-45.
3. Gordon C, Amissah-Arthur M-B, Gayed M, Brown S, Bruce IN, D'Cruz D, et al. The British Society for Rheumatology guideline for the management of systemic lupus erythematosus in adults. Rheumatology. 2017;57(1):e1-e45.
4. Groot N, de Graeff N, Avcin T, Bader-Meunier B, Brogan P, Dolezalova P, et al. European evidence-based recommendations for diagnosis and treatment of childhood-onset systemic lupus erythematosus: the SHARE initiative. Ann Rheum Dis. 2017;76(11):1788-96.
5. Rathmann W, Bongaerts B, Carius HJ, Kruppert S, Kostev K. Basic characteristics and representativeness of the German Disease Analyzer database Int J Clin Pharmacol Ther. 2018;56(10):459-66.
6. Recalde M, Rodríguez C, Burn E, Far M, García D, Carrere-Molina J, et al. Data Resource Profile: The Information System for Research in Primary Care (SIDIAP). Int J Epidemiol. 2022;51(6):e324-e36.
7. Martín-Baranera M, Planas I, Palau J, Sanz F. IMASIS computer-based medical record project: dealing with the human factor. Medinfo. 1995;8 Pt 1:333.
8. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). Int J Epidemiol. 2015;44(3):827-36.
9. Clinical Practice Research Datalink. CPRD GOLD July 2022 (Version 2022.07.001) [Data set] Clinical Practice Research Datalink. https://doi.org/10.48329/ty3h-h728. 2022.
10. Markus AF, Verhamme KMC, Kors JA, Rijnbeek PR. TreatmentPatterns: An R package to facilitate the standardized development and analysis of treatment patterns across disease domains. Computer Methods and Programs in Biomedicine. 2022;225:107081.
11. Rees F, Doherty M, Grainge M, Davenport G, Lanyon P, Zhang W. The incidence and prevalence of systemic lupus erythematosus in the UK, 1999-2012. Ann Rheum Dis. 2016;75(1):136-41.
12. Nightingale AL, Davidson JE, Molta CT, Kan HJ, McHugh NJ. Presentation of SLE in UK primary care using the Clinical Practice Research Datalink. Lupus Sci Med. 2017;4(1):e000172.
13. European Medicines Agency and Heads of Medicines Agencies. Guideline on good pharmacovigilance practices (GVP) – Module VI (Rev 2). EMA/873138/2011 Rev 2 2017. Available from: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guideline-good-pharmacovigilance-practices-gvp-module-vi-collection-management-submission-reports_en.pdf.

17. ANNEXES

Appendix I: Definition of SLE Diagnosis and Treatments

Table 1: Preliminary code list for SLE.

Note: We will consider two concept sets – narrow definition for main analysis and broad definition for sensitivity analysis. The concept sets are subject to change after review of cohort diagnostics.

CONCEPT_ID	CONCEPT_NAME	narrow	broad
4295179	Acute systemic lupus erythematosus	TRUE	TRUE
36676444	Autosomal systemic lupus erythematosus	TRUE	TRUE
4346976	Bullous systemic lupus erythematosus	TRUE	TRUE
46270532	Cheilitis due to lupus erythematosus	FALSE	TRUE
37110504	Chorea co-occurrent and due to systemic lupus erythematosus	TRUE	TRUE
4044056	Chorea in systemic lupus erythematosus	TRUE	TRUE
37110517	Demyelination of central nervous system co-occurrent and due to systemic lupus erythematosus	TRUE	TRUE
4269448	Dilated cardiomyopathy due to systemic lupus erythematosus	TRUE	TRUE
46273369	Endocarditis due to systemic lupus erythematosus	TRUE	TRUE
4296502	Fulminating systemic lupus erythematosus	TRUE	TRUE
37019030	Gingival disease co-occurrent and due to lupus erythematosus	FALSE	TRUE
37016279	Glomerular disease due to systemic lupus erythematosus	TRUE	TRUE
4055640	Lung disease with systemic lupus erythematosus	TRUE	TRUE
4299106	Lupus disease of the lung	TRUE	TRUE
255891	Lupus erythematosus	FALSE	TRUE
45768793	Lupus erythematosus of oral mucous membrane	FALSE	TRUE
4291306	Lupus erythematosus overlap syndrome	FALSE	TRUE
4301142	Lupus erythematosus-associated necrotizing vasculitis	FALSE	TRUE
4295305	Lupus erythematosus-associated urticarial vasculitis	FALSE	TRUE
4057084	Lupus hepatitis	TRUE	TRUE
4344399	Lupus panniculitis	TRUE	TRUE
4344495	Lupus vasculitis	TRUE	TRUE
4105023	Myopathy due to disseminated lupus erythematosus	TRUE	TRUE
4316373	Neonatal lupus erythematosus	FALSE	TRUE
46270384	Nephropathy co-occurrent and due to systemic lupus erythematosus	TRUE	TRUE
37399735	Nephrosis co-occurrent and due to systemic lupus erythematosus	TRUE	TRUE
37395585	Nephrotic syndrome co-occurrent and due to systemic lupus erythematosus	TRUE	TRUE
4101469	Pericarditis secondary to systemic lupus erythematosus	TRUE	TRUE

4105637	Polyneuropathy in disseminated lupus erythematosus	TRUE	TRUE
4319305	Rash of systemic lupus erythematosus	TRUE	TRUE
4145240	Renal tubulo-interstitial disorder in systemic lupus erythematosus	TRUE	TRUE
4217054	Retinal vasculitis due to systemic lupus erythematosus	TRUE	TRUE
37117740	Secondary autoimmune hemolytic anemia co-occurrent and due to systemic lupus erythematosus	TRUE	TRUE
4285717	SLE glomerulonephritis syndrome	TRUE	TRUE
4250483	SLE glomerulonephritis syndrome, WHO class I	TRUE	TRUE
4186940	SLE glomerulonephritis syndrome, WHO class II	TRUE	TRUE
4297164	SLE glomerulonephritis syndrome, WHO class III	TRUE	TRUE
4267801	SLE glomerulonephritis syndrome, WHO class IV	TRUE	TRUE
4178133	SLE glomerulonephritis syndrome, WHO class V	TRUE	TRUE
4002526	SLE glomerulonephritis syndrome, WHO class VI	TRUE	TRUE
257628	Systemic lupus erythematosus	TRUE	TRUE
4318863	Systemic lupus erythematosus encephalitis	TRUE	TRUE
44784527	Systemic lupus erythematosus in remission	FALSE	FALSE
4301051	Systemic lupus erythematosus of childhood	TRUE	TRUE
4344400	Systemic lupus erythematosus with multisystem involvement	TRUE	TRUE
4344158	Systemic lupus erythematosus with organ/system involvement	TRUE	TRUE
4149913	Systemic lupus erythematosus with pericarditis	TRUE	TRUE
44814064	Systemic lupus erythematosus/Sjogren's overlap syndrome	TRUE	TRUE
4300204	Systemic lupus erythematosus-associated antiphospholipid syndrome	TRUE	TRUE
4219859	Systemic lupus erythematosus-related syndrome	TRUE	TRUE

SLE Treatments

Table 2: Preliminary code list for SLE treatments.

Class	Treatment	WHO ATC code	Ingredient ConceptID
Antimalarial	Hydroxychloroquine	P01BA02	1777087
DMARD	Methotrexate	L01BA01 L04AX03	1305058
	Azathioprine	L04AX01	19014878
	Mycophenolate Mycophenolate mofetil	L04AA06	19068900 19003999
	Cyclophosphamide	L01AA01	1310317
Calcineurin inhibitors	Tacrolimus	L04AD02	950637
	Cyclosporine	L04AD01	19010482
	Voclosporin	L04AD03	739590
Biologic agents	Rituximab	L01FA01	1314273

	Belimumab	L04AA26	40236987
Glucocorticoids	Betamethasone	H02AB01	920458
	Dexamethasone	H02AB02	1518254
	Fluocortolone	H02AB03	19055344
	Methylprednisolone	H02AB04	1506270
	Paramethasone	H02AB05	19027186
	Prednisolone	H02AB06	1550557
	Prednisone	H02AB07	1551099
	Triamcinolone	H02AB08	903963
	Hydrocortisone	H02AB09	975125
	Cortisone	H02AB10	1507705
	Prednylidene	H02AB11	19011127
	Rimexolone	H02AB12	977421
	Deflazacort	H02AB13	19086888
	Cloprednol	H02AB14	19050907
	Meprednisone	H02AB15	19009116
	Cortivazol	H02AB17	19061907

Appendix II: ENCePP checklist for study protocols

Study title:
DARWIN EU® - Treatment patterns of drugs used in adult and paediatric population with systemic lupus erythematosus

EU PAS Register® number: N/A
Study reference number (if applicable): N/A

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

Comments:

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

Comments:

[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:

This is a descriptive study and no measure of association or collection or reporting of adverse events/reactions will be reported.

Section 4: Source and study populations		Yes	No	N/A	Section Number
4.1	Is the source population described?	☒	☐	☐	8.2/8.5
4.2	Is the planned study population defined in terms of:				8.5
	4.2.1 Study time period	☒	☐	☐	
	4.2.2 Age and sex	☒	☐	☐	
	4.2.3 Country of origin	☒	☐	☐	
	4.2.4 Disease/indication	☒	☐	☐	
	4.2.5 Duration of follow-up	☒	☐	☐	
4.3	Does the protocol define how the study population will be sampled from the source population? (e.g., event or inclusion/exclusion criteria)	☒	☐	☐	8.5

Comments:

Section 5: Exposure definition and measurement		Yes	No	N/A	Section Number
5.1	Does the protocol describe how the study exposure is defined and measured? (e.g. operational details for defining and categorising exposure, measurement of dose and duration of drug exposure)	☒	☐	☐	8.8

Section 5: Exposure definition and measurement	Yes	No	N/A	Section Number
5.2 Does the protocol address the validity of the exposure measurement? (e.g., precision, accuracy, use of validation sub-study)	☐	☒	☐	
5.3 Is exposure categorised according to time windows?	☐	☒	☐	
5.4 Is intensity of exposure addressed? (e.g., dose, duration)	☒	☐	☐	8.8
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:

This is a descriptive study with no comparison of exposures.

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?	☐	☐	☒	
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:

This is a descriptive study detailing patient characterisation and drug utilisation, with no outcomes.

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)	☒	☐	☐	11

Section 7: Bias	Yes	No	N/A	Section Number
7.3 Does the protocol address information bias? (e.g. misclassification of exposure and outcomes, time-related bias)	☒	☐	☐	11

Comments:

This is a descriptive study and no associations will be assessed.

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:

This is a descriptive study and no associations will be assessed.

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
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?	☒	☐	☐	8.2
9.2 Does the protocol describe the information available from the data source(s) on:				
9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	8.2
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)	☒	☐	☐	8.2
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))	☐	☐	☒	
9.3.3 Covariates and other characteristics?	☒	☐	☐	8.6

Section 9: Data sources	Yes	No	N/A	Section Number
9.4 Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☐	☐	☒	

Comments:

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

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
11.3 Is there a system in place for independent review of study results?	☐	☒	☐	

Comments:

Section 12: Limitations	Yes	No	N/A	Section Number
12.1 Does the protocol discuss the impact on the study results of:				
12.1.1 Selection bias?	☒	☐	☐	11

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

Comments:

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

Comments:

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

Comments:

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

Comments:

Name of the main author of the protocol: Eng Hooi Tan

Date: 17/08/2023

Signature: E.H. Tan