Study Protocol P2 C1-007

25/03/2024

Version 1.2

TABLE OF CONTENTS

DOCUMENT HISTORY

Version	Date	Description
V1.0	31/08/2023	**First submission to EMA**
V1.1	22/09/2023	**Version with minor amendments for archiving**

Study Title	DARWIN EU® - Natural history of dermatomyositis (DM) and polymyositis (PM) in adults and paediatric populations
Protocol version identifier	V1.1
Date of last version of protocol	22/09/2023
EU PAS register number	EUPAS107454
Active substance	N/A
Medicinal product	N/A
Research question and objectives	The <u>overall objective</u> of this study is to describe and characterise dermatomyositis (DM), polymyositis (PM) and their juvenile forms (JDM and JPM), in terms of prevalence, natural history of the disease, disease severity, and treatment. The <u>specific objectives</u> of this study are: 1. To estimate the yearly prevalence of DM and PM in adult (18+ years) and paediatric populations (0 to less than 2 years, 2 to less than 6 years, 6 to less than 12 years, 12 to less than 18 years), overall, and by sex. 2. To characterise patients and describe age at disease onset, for DM, PM, JDM and JPM. 3. To describe the occurrence in adults and children of biomarker measurements (e.g. creatinine kinase, tests for myositis auto-antibodies - INF-b levels, INF type I gene signature) before, at the time of, and after a diagnosis of DM, PM, JDM and JPM. 4. To describe the occurrence of clinical manifestations (muscle inflammation, muscle weakness, connective tissue disease overlap, presence of calcinosis in children) before, at the time, and after a diagnosis of DM, PM, JDM and JPM. 5. To describe disease severity including organ involvement (skin, joints, lung, heart, GI tract) before, at the time, and after a diagnosis of DM, PM, JDM and JPM. 6. To describe treatment administered (including combinations and sequences) after a diagnosis of DM, PM, JDM and JPM All results will be reported by database, overall, and by study periods (2006-2013, 2013-2020, and 2020-2022), and stratified by age and sex when possible.

Country(-ies) of study	Estonia, France, Germany, Spain, United Kingdom
Authors	Albert Prats-Uribe (albert.prats-uribe@ndorms.ox.ac.uk) Lucia Bellas (lucia.bellas@ndorms.ox.ac.uk) Daniel Prieto-Alhambra (d.prietoalhambra@darwin-eu.org)

LIST OF ABBREVIATIONS

Acronyms/terms	Description
ACR	American College of Rheumatology
ADM	Amyopathic dermatomyositis
ATC	Anatomical Therapeutic Chemical Classification System
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
DM	Dermatomyositis
DMARD	disease-modifying antirheumatic drug
csDMARD	conventional synthetic disease-modifying antirheumatic drug
DOI	Declaration Of Interests
EBB	Estonian BioBank
EHR	Electronic Health Records
EMA	European Medicines Agency
EULAR	European League Against Rheumatism
GERD	Gastro-esophageal reflux disease
GP	General Practitioner
IBM	Inclusion body myositis
IIM	Idiopathic inflammatory myopathies
IMASIS	Institut Municipal Assistència Sanitària Information System
JDM	Juvenile Dermatomyositis
JPM	Juvenile Polymyositis
LOINC	Logical Observation Identifiers Names and Codes
OMOP	Observational Medical Outcomes Partnership
PM	Polymyositis
PCT	Primary Care Teams
RxNorm	Medical prescription normalized
SIDIAP	Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària
SNOMED	Systematized Nomenclature of Medicine

1. TITLE

DARWIN EU® - Natural history of dermatomyositis (DM) and polymyositis (PM) in adults and paediatric populations

2. RESPONSIBLE PARTIES – STUDY TEAM

Study team Role	Names	Organisation
Study Project Manager/ Principal Investigator	Albert Prats-Uribe	University of Oxford
Epidemiologist	Albert Prats-Uribe Ed Burn	University of Oxford University of Oxford
Clinical Domain Expert	Daniel Prieto-Alhambra Lucía Bellas	University of Oxford University of Oxford
Data Analysts/statisticians	Ed Burn Mike Du	University of Oxford University of Oxford

Data Partner*	Names	Organisation – Database
Local Study Coordinator/Data Analyst	James Brash	IQVIA - DA Germany
	Jasmine Gratton	IQVIA - DA Germany
	Talita Duarte Salles	IDIAPJGol - SIDIAP
	Laura Pérez Crespo	IDIAPJGol - SIDIAP
	Romain Griffier	University of Bordeaux - CDWBordeaux
	Antonella Delmestri	University of Oxford – CPRD GOLD
	Hezekiah Omulo	University of Oxford – CPRD GOLD
	Wai Yi (Teen) Man	University of Oxford – CPRD GOLD
	Raivo Kolde	Estonian Biobank

*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® - Natural history of dermatomyositis (DM) and polymyositis (PM) in adults and paediatric populations

Rationale and Background

Idiopathic inflammatory myopathies (IIM) are rare and diverse autoimmune disorders characterized by muscle inflammation, weakness, and extramuscular manifestations affecting organs like skin, lungs, heart, and joints (Lundberg 2021, Sasaki 2018). The subgroups include dermatomyositis, antisynthetase syndrome, immune-mediated necrotizing myopathy, inclusion body myositis, polymyositis, and overlap

myositis (Lundberg 2021). Despite their rarity, understanding the epidemiology of these disorders is essential to identify patterns and determinants.

These diseases are challenging because of their associated morbidity and mortality (Aggarwal et al., 2017). Classification criteria developed by the European League Against Rheumatism (EULAR) and the American College of Rheumatology (ACR) help identify major IIM subgroups. Diagnostic tools involve elevated muscle-derived enzymes in serum, antinuclear antibodies, muscle biopsy, electromyography, and MRI (Lundberg 2021, Papadopoulo 2023).

The pathogenesis, treatment responses, and organ involvement vary among IIM subtypes, necessitating a deeper understanding of molecular pathways and auto-antigens (Lundberg 2021, Findlay 2015). Glucocorticoids are commonly used as first-line treatment, often combined with immunosuppressive agents like methotrexate, azathioprine, and others (Oldroyd 2022, Sasaki 2018). Rituximab shows promise in refractory cases (Valiyil R 2010-, Mok 2007). Although TNF's role is implicated, anti-TNF treatments' efficacy is limited (Lundberg 2021).

Pediatric cases require special consideration. Juvenile idiopathic inflammatory myopathies affect children and young individuals, involving muscles, skin, and other organs. Differences exist between juvenile and adult forms in terms of pathogenesis, autoantibody profiles, and treatment responses. Consensus guidelines help guide diagnosis and management (Belluti et al.).

In conclusion, idiopathic inflammatory myopathies encompass a spectrum of rare autoimmune disorders affecting muscles and various organs. Understanding their epidemiology, classification, diagnostic criteria, and treatment approaches is essential for improving patient outcomes and tailoring treatments, especially in paediatric cases.

Research question and Objectives

The overall objective of this study is to describe and characterise dermatomyositis (DM), polymyositis (PM) and their juvenile forms (JDM and JPM), in terms of prevalence, natural history of the disease, disease severity, and treatment patterns.

The specific objectives of this study are:

1. To estimate the yearly prevalence of DM and PM in adult (18+ years) and paediatric populations (less than 2 years, 2 to less than 6 years, 6 to less than 12 years, 12 to less than 18 years) overall and by sex.
2. To characterise patients and describe age at disease onset, for DM, PM, JDM and JPM.
3. To describe the occurrence in adults and children of biomarker measurements (e.g. creatinine kinase, tests for myositis auto-antibodies - INF-b levels, INF type I gene signature) before, at the time of, and after a diagnosis of DM, PM, JDM and JPM.
4. To describe the occurrence of clinical manifestations (muscle inflammation, muscle weakness, connective tissue disease overlap, presence of calcinosis) before, at the time of, and after a diagnosis of DM, PM, JDM and JPM.

5. To describe disease severity including organ involvement (skin, joints, lung, heart, GI tract) before, at the time of, and after a diagnosis of DM, PM, JDM and JPM.
6. To describe treatment administered (including combinations and sequences) after a diagnosis of DM, PM, JDM and JPM

All results will be reported by database, overall, and by study periods (2006-2012, 2013-2019, and 2020-2022), and stratified by age and sex when possible.

Research Methods

Study design

Cohort study. We will include cohorts of first diagnosed DM, PM, JDM, JPM and new user cohorts of their treatments (for objective 6).

Population

The source population will include all individuals eligible in the database between 01/01/2006 and end of the available date in each database. For objective 1, all patients active in the database at the start of all calendar year will be included. For objectives 2-5, two cohorts with be characterised, one with a 90-day prior history requirement from diagnosis date, and one without this requirement. For objective 6, a washout period of 365 days at the treatment ingredient level will be applied to capture new users of DM, PM, JDM and JPM treatment.

Variables

DM, PM, JDM and JPM will be assessed as first occurrence of the codes specified in Annex 1. Additional age criteria, <18 year old at time of first diagnosis will be applied for JDM and JPM and >18 at time of first diagnosis for DM and PM. Co-morbidities and co-medications will be used for large-scale patient characterisation, identified as concept/code and descendants. A list of pre-specified co-morbidities, measurements, clinical manifestations, and severity markers will also be characterised and is included in Annex 1. Treatments of DM, PM, JDM, JPM will be identified using the codes included in Annex 1.

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. Clinical Data Warehouse of Bordeaux University Hospital (CDW Bordeaux), France
4. Clinical Practice Research Datalink (CPRD) GOLD, United Kingdom (UK)
5. Estonian Biobank (EBB), Estonia

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 DM, PM, JDM and JPM patients. Based on a preliminary feasibility assessment the expected number of patients in the included databases for this study will be approximately 6,000 for DM, and 5,000 for PM. We will define JDM and JPM based on disease diagnostic codes and age at diagnosis as they are likely to not be coded specifically as juvenile, so we are not able to determine a concrete sample size for the juvenile forms (we expect around 3-10% of them to be juvenile).

Data analyses

Point prevalence of each outcome of interest (DM, PM, JDM, JPM), with every individual deemed to have the diagnosis from first occurrence until end of follow-up calculated on an annual basis as of the 1st January for each year, estimated overall and stratified by age and sex.

Age and sex at time of DM, PM, JDM, JPM diagnosis (index date) will be described for each of the generated study cohorts (Objective 2). Large-scale patient-level characterisation will be conducted for objectives 3 to 5. Occurrence of co-morbidities, measurements, clinical manifestations, and severity markers will be assessed for anytime –and up to 365 days before index date, for 364 to 91, for 90 to 31, and for 30 to 1 day before index date, and at index date. We will also report them for 1 to 90, 91 to 180, 181 to 365 days, 366 to 1095, 1096 to 1825 days, and 1826 days to any time post index date.

The number and percentage of patients receiving each of a pre-specified list of DM, PM, JDM and JPM treatments (see Appendix 1) and treatment combinations will be described at index date, 1 to 90, 91 to 180, 181 to 365 days, 366 to 1095, 1096 to 1825 days, and 1826 days to any time post index date. Additionally, sunburst plots and Sankey diagrams will be used to describe treatment patterns and sequences over time (objective 6).

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 groups and sex when possible (minimum cell count reached). Additionally, to capture treatments availability and changes over time, analyses will be further stratified by study periods (2006-2013, 2013-2020, and 2020-2022).

4.   AMENDMENTS AND UPDATES

Number	Date	Section of study protocol	Amendment or update	Reason
V1.2	25-03-2024	Document History	Update	Added the EUPAS registration number

5. MILESTONES

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

6. RATIONALE AND BACKGROUND

Idiopathic inflammatory myopathies (IIM) are a heterogeneous group of rare and chronic autoimmune disorders with clinical manifestations such as muscle inflammation, muscle weakness and extra-muscular manifestations including involvement of organs such as the skin, lung, heart, gastrointestinal tract and joints (Lundberg 2021, Sasaki 2018).

On the basis of clinical, histopathological and serological features, IIM can be classified into several subgroups - dermatomyositis (including amyopathic dermatomyositis), antisynthetase syndrome, immune-mediated necrotizing myopathy, inclusion body myositis, polymyositis and overlap myositis (Lundberg 2021).

Although these conditions are believed to be very rare, the epidemiological features of IMs have been poorly studied and synthesis of the existing data regarding incidence and prevalence are lacking. Epidemiological studies of rare diseases are essential to identify their geographical and population disparities as well as clusters and time trends and hence their possible key determinants. Such epidemiological patterns may provide useful clues towards improving our understanding of IMs. (Meyer, 2015) . A wide range of estimates of incidence and prevalence as well as of risk factors for disease have been published. Globally, albeit with a majority of studies from Asia, Europe and North America, the incidence estimates range from 11 to 660 patients with newly diagnosed inflammatory miosisitis per 1,000,000 person-years and between 2.9 and 34 individuals per 100,000 population are suggested to have the disease. All available data suggest that PM, DM and IMNM are more common in women than in men. The incidence increases with age and the peak age of incidence is ~50 years of age in both Europe and North America. (Lundberg 2021)

Despite significant morbidity and mortality associated with DM/PM, there are currently no specific therapies approved for these syndromes by the European Medicines Agency or the Food and Drug Administration based on randomised controlled trials. (Aggarwal et al, 2017)

Polymyositis (PM), dermatomyositis (DM), and inclusion body myositis (IBM) are major IIM subsets (Sasaki 2018). Juvenile onset DM (JDM) and PM (JPM) are sub-categories of DM and PM whereas IBM only occurs in adults. JDM and JPM can also encompass juvenile patients with overlap myositis (ie, DM or PM signs and symptoms but co-existing with other autoimmune diseases such as lupus or scleroderma)

DM is defined by the presence of characteristic cutaneous manifestations (Gottron papules, "V sign", heliotrope rash among others) and myositis. While muscle and skin involvement coexist in the prototype of DM (classic DM), DM can exist without muscle disease (amyopathic DM) or overt muscle symptoms despite evidence of myositis on laboratory testing (hypomyopathic DM). Amyopathic DM and hypomyopathic DM are defined when the conditions last for ≥6 months and are collectively termed as clinically amyopathic DM (Gerami et al)

PM could be defined as a myositis phenotype with chronic muscle weakness without skin involvement and involving predominant cytotoxic T cell mechanisms (Lundberg 2021). However, most historical studies of PM included samples from patients now classified as Immune-Mediated Necrotizing Myopathy and antisynthetase syndrome without a rash, diseases now recognized to be pathologically distinct from each other. Thus, future studies will be required to define the risk factors and mechanisms underlying muscle inflammation and damage more completely in PM. (Aggarwal et al, 2017)

Diagnosis of myositis is made when typical clinical and laboratory parameters are present and other possible causes are excluded. However, formal diagnostic criteria do not exist, and classification criteria are used for guidance instead. The most recent classification criteria developed by the European League Against Rheumatism (EULAR) and the American College of Rheumatology (ACR) published in 2017 (Lundberg 2017) can aid in identifying the major myositis subgroups: DM, PM, ADM, juvenile myositis and IBM. (Lundberg 2021)

Muscle-derived enzymes in serum are elevated in most patients with active muscle disease. Creatine kinase is the most sensitive marker and has diagnostic and monitoring utility. Antinuclear antibodies (identified by indirect inmunofluorescence) have also diagnostic utility. Muscle biopsy is an important tool to diagnose IIM, confirm signs of inflammation, identify signs of the different subtypes of IIM and, importantly, exclude other myopathies. Abnormal electrical activities of muscle fibres, signs of muscle oedema or immune-mediated changes in histopathological specimens can be detected by electromyography, imaging and muscle biopsy. (Lunberg, 2021). MRI is now favoured as a diagnostic tool. (Papadopoulo 2023). In a systematic revision, Meyer et al found that the median time to diagnosis of overall IM varied between 3 and 6 months. A considerably delayed diagnosis was a constant feature of IBM. The mean duration of symptoms before diagnosis varied between 4.1 and 8 years.

The prognoses, treatment responses and organ manifestations vary among IIM subtypes, implicating different pathophysiological mechanisms in each subtype. A deeper understanding of the molecular pathways underlying the pathogenesis and identifying the auto-antigens of the immune reactions in these subgroups is crucial to improving outcomes (Lundberg 2021, Findlay 2015).

In terms of treatments and care management of patients with IIM, glucocorticoids are used empirically as the first-line treatment despite their various adverse effects. Glucocorticoid dose should be weaned when disease activity, considered across all domains, substantially improves, usually after around 6 weeks of treatment initiation. (Alexander GS). Concomitant treatment with steroid-sparing immunosuppressive agents, including methotrexate, azathioprine, calcineurin inhibitors, mycophenolate mofetil, and cyclophosphamide, reduces successfully initial glucocorticoid doses for the remission induction, the relapse risk during glucocorticoid tapering, and adverse effects of glucocorticoids. (Oldroyd 2022, Sasaki 2018).

Evidence does not exist to allow recommendation of specific csDMARDs as first-/ second-/third-line for adults. DMARDs should be prescribed and monitored according to existing age appropriate BSR guidelines. (Alexander GS)

Rituximab depletes CD20 B cells that are likely to be involved in the pathogenesis of some myositis subgroups. Several open-label studies have reported safety and efficacy in patients with severe and refractory myositis (Valiyil R 2010-, Mok 2007). Although TNF has been implicated in the pathogenesis of myositis, the efficacy of anti-TNF agents (such as etanercept and infliximab) is somewhat disappointing. Currently, anti-TNF treatment is not typically recommended or considered in patients with adult myositis, although it may have a role in the treatment of calcinosis in juvenile DM (Lundberg 2021).

In the context of pediatric extrapolation in clinical drug development, it is important to understand the potential for disease similarity (or differences) between juvenile and adult forms of DM and PM, as well as between JPM and JDM: The childhood-onset or juvenile idiopathic inflammatory myopathies are a heterogenous group of rare and serious autoimmune diseases of children and young people that predominantly affect the muscles and skin but can also involve other organs, including the lungs, gut, joints, heart and central nervous system. Juvenile idiopathic inflammatory myopathies can differ from adult-onset myopathies in terms of the pathogenesis, autoantibody profile, disease phenotype and treatment response, but these differences need to be further defined. A Single Hub and access point for Paediatric Rheumatology in Europe (SHARE) initiative-based consensus guideline has set out recommendations for diagnosis (Belluti et al). Ultimately, a combination of a better understanding of disease mechanisms, biomarkers that accurately track disease activity, including subclinical disease, and definitions of outcomes that include the patient perspective will be needed to deliver a personalized approach to managing myositis in children, and in the young people and adults they become.

7. RESEARCH QUESTION AND OBJECTIVES

The overall objective of this study is to describe and characterise dermatomyositis (DM), polymyositis (PM) and their juvenile forms (JDM and JPM), in terms of prevalence, natural history of the disease, disease severity, and treatment.

The specific objectives of this study are:

1. To estimate the yearly prevalence of DM and PM in adult (18+ years) and paediatric populations (0 to less than 2 years, 2 to less than 6 years, 6 to less than 12 years, 12 to less than 18 years) overall and by sex.
2. To characterise patients and describe age at disease onset, for DM, PM, JDM and JPM.

3. To describe the occurrence in adults and children of biomarker measurements (e.g. creatinine kinase, tests for myositis auto-antibodies - INF-b levels, INF type I gene signature) before, at the time, and after a diagnosis of DM, PM, JDM and JPM.
4. To describe the occurrence of clinical manifestations (muscle inflammation, muscle weakness, connective tissue disease overlap, presence of calcinosis in children) before, at the time, and after a diagnosis of DM, PM, JDM and JPM.
5. To describe disease severity including organ involvement (skin, joints, lung, heart, GI tract) before, at the time, and after a diagnosis of DM, PM, JDM and JPM.
6. To describe treatment administered (including combinations and sequences) after a diagnosis of DM, PM, JDM and JPM

All results will be reported by database, overall, and by study periods (2006-2013, 2013-2020, and 2020-2022), and stratified by age and sex when possible.

Table 1: Primary and secondary research questions and objective

Objective:	1. To estimate the yearly prevalence of DM and PM in adult (18+ years) and paediatric populations (0 to less than 2 years, 2 to less than 6 years, 6 to less than 12 years, 12 to less than 18 years) overall and by sex. 2. To characterise patients and describe age at disease onset, for DM, PM, JDM and JPM. 3. To describe the occurrence in adults and children of biomarker measurements (e.g. creatinine kinase, tests for myositis auto-antibodies - INF-b levels, INF type I gene signature) before, at the time, and after a diagnosis of DM, PM, JDM and JPM. 4. To describe the occurrence of clinical manifestations (muscle inflammation, muscle weakness, connective tissue disease overlap, presence of calcinosis in children) before, at the time, and after a diagnosis of DM, PM, JDM and JPM. 5. To describe disease severity including organ involvement (skin, joints, lung, heart, GI tract) before, at the time, and after a diagnosis of DM, PM, JDM and JPM. 6. To describe treatment administered (including combinations and sequences) after a diagnosis of DM, PM, JDM and JPM
Hypothesis:	N/A
Population *(mention key inclusion-exclusion criteria):*	All individuals with a diagnosis of DM, PM, JDM and JPM identified in the database between 01/01/2006 and 31/12/2022 or end of the available date in each database. For objective 1, all patients active in the database at start of each available year will be used to form the denominator. For objectives 2-5, two cohorts for each of the DM, PM, JDM and JPM groups will be characterised, one with a 90-day prior history requirement from diagnosis date, and one without this requirement. For the treatment cohorts, a washout period of 365 days at the

	treatment ingredient level will be applied to capture new users of DM, PM, JDM and JPM treatments (objective 6).
Exposure:	DM, PM, JDM and JPM assessed as first occurrence of the codes specified in Appendix 1 Table 1. DM, PM, JDM and JPM treatments (Appendix 1 Table 4): • Hydroxychloroquine, chloroquine • Systemic glucocorticoids (prednisone, methylprednisolone) • Methotrexate • Azathioprine • Calcineurin inhibitors (tacrolimus, cyclosporine) • Mycophenolate • Cyclophosphamide • Immunoglobulins • DMARDs: Rituximab (anti-CD20), etanercept (anti-TNF), infliximab (anti-TNF), abatacept (anti-CTLA-4)
Comparator:	N/A
Outcome:	N/A
Time *(when follow up begins and ends):*	For objectives 1 to 5, follow-up will start from date of first DM, PM, JDM or JPM diagnosis until the earliest of the following: 1) loss to follow-up, 2) end of data availability, or 3) date of death. For objective 6, follow-up will start from date of first DM, PM, JDM and JPM treatment after DM, PM, JDM and JPM diagnosis until the earliest of the following: 1) loss to follow-up, 2) end of data availability, or 3) date of death.
Setting:	Inpatient, outpatient, and biobank setting from 5 databases in 5 European countries.
Main measure of effect:	Prevalence of DM, PM, JDM and JPM. Biomarker, clinical severity, and clinical manifestation occurrence after and before DM, PM, JDM and JPM diagnosis. Proportions of patients on treatment types and sequences, patient-level drug utilisation.

8. RESEARCH METHODS

8.1 Study Type and Study Design

The study will consist of a retrospective cohort designed including patients with a diagnosis of DM, PM, JDM, and JPM. We will perform **a population level descriptive epidemiology** and a **patient-level characterisation** study classified as "off-the-shelf" (C1) and as described in the DARWIN EU® Complete Catalogue of Standard Data Analyses. A retrospective cohort study of all DM, PM, JDM and JPM cases will be conducted.

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

STUDY TYPE	STUDY DESIGN	STUDY CLASSIFICATION
Population level descriptive epidemiology	Cohort analysis.	Off-the-shelf (C1)
Patient-level characterisation	Cohort analysis.	Off-the-shelf (C1)
Patient-level treatment patterns	Cohort analysis.	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 5 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. Clinical Data Warehouse of Bordeaux University Hospital (CDWBordeaux), France
4. Clinical Practice Research Datalink (CPRD) GOLD, United Kingdom (UK)
5. Estonian Biobank (EBB), Estonia

We selected 5 out of the 10 databases currently onboarded in DARWIN EU® in 2022, as of 27/07/2023. 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 DM, PM, JDM and JPM treatment data 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.

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 DM, PM diagnoses and treatment.	Primary care and outpatient specialist care	EHR	8.5 million	31/12/2022	2 to 6
ES	SIDIAP	Covers primary care setting with a proportion with hospital linkage, data on DM, PM diagnoses and some treatments.	Primary care with hospital linkage	EHR	5.8 million	31/12/2022	1 to 6
FR	CDWBordeaux	Covers secondary care setting, database has information on DM,PM diagnosis and in-hospital treatments	Secondary care (in and outpatients)	EHR	1.9 million	10/04/2023	2 to 6
UK	CPRD GOLD	Covers primary care setting, database has information on DM,PM diagnosis and treatments	Primary care	EHR	3.1 million	04/07/2022	1 to 6
EE	EBB	Covers primary care and hospital data for a biobank cohort, containing data on DM, and PM diagnoses and treatment.	Primary care and secondary care.	Biobank cohort.	200,000	31/12/2021	1to 6

DE = Germany, ES = Spain, FR = France, NL = The Netherlands, UK = United Kingdom, EE= Estonia, SIDIAP = Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària, DA = Disease Analyzer, CDWBordeaux = Clinical Data Warehouse of Bordeaux University Hospital, CPRD = Clinical Practice Research Datalink, EBB = Estonian Biobank.

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

. 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.

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.

Estonian Biobank (EBB), Estonia

The Estonian Biobank (EBB) is a population-based biobank of the Estonian Genome Center at the University of Tartu (EGCUT. Its cohort size is currently close to 200,000 participants ("gene donors" ≥ 18 years of age), which closely reflects the age, sex and geographical distribution of the Estonian population. Estonians represent 83%, Russians 14%, and other nationalities 3% of all participants. Genomic GWAS analyses have been performed on all gene donors. The database also covers health insurance claims, digital prescriptions, discharge reports, information about incident cancer cases, and causes of death from national sources for each donor.

8.3 Study Period

The study period will be from 01/01/2006 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 5, follow-up will start from date of first DM, PM, JDM and JPM diagnosis until the earliest of the following: 1) loss to follow-up, 2) end of data availability, or 3) date of death.

For objective 6, follow-up will start from date of first treatment after a first DM, PM, JDM and JPM diagnosis until the earliest of the following: 1) loss to follow-up, 2) end of data availability, or 3) date of death.

8.5 Study Population with inclusion and exclusion criteria

For objective 1, the study population will include all individuals identified in the database with at least 90 days of previous data visibility between 01/01/2006 and end of available data in each database. For a sensitivity analysis, the 90 days restriction will be removed.

For objective 2-6, The study population will include all individuals with a first diagnosis of DM, PM, JDM and JPM identified in the database between 01/01/2006 and end of available data in each database.

For objective 6, patients will be identified based on a record indicating a diagnosis of DM, PM, JDM and JPM. Diagnoses will be identified using condition and observation codes in the OMOP CDM that use the Systematized Nomenclature of Medicine (SNOMED) as the standard vocabulary for these. A preliminary code list is provided in appendix 1 table 1.

Additional eligibility criteria will be applied for each study objective:

For objectives 2-4, two cohorts per diagnosis, one with at least 90 days of prior history available before date of new DM, PM, JDM and JPM diagnosis will be applied for large-scale characterisation and another without this requirement; For objective 6, a washout period of 365 days at the treatment ingredient level will be applied to capture new treatment users of DM, PM, JDM and JPM treatments (specified in appendix 1. table 4).

8.6 Variables

8.6.1. Exposure/s

For Objective 6, DM, PM, JDM and JPM treatments will be identified using RxNorm codes. These include hydroxychloroquine, chloroquine, systemic glucocorticoids (prednisone, methylprednisolone), methotrexate, azathioprine, calcineurin inhibitors (tacrolimus, cyclosporine, voclosporin), mycophenolate, cyclophosphamide, inmunoglobulins , and DMARDs: rituximab (anti-CD20), etanercept (anti-TNF), infliximab (anti-TNF), abatacept (anti-CTLA-4). A preliminary list of codes to identify these treatments can be found in Appendix 1 Table 4.

8.6.2. Outcome/s

For Objective 1, the outcome will be a diagnosis of indicating a diagnosis of DM, PM, JDM and JPM as defined in Appendix 1 Table 1.

For Objective 2, Age at DM, PM, JDM, JPM diagnosis will be described.

For Objective 3, the outcome will be the presence of a biomarker measurement code as defined by SNOMED and LOINC codes, including autoantibodies, inflammation markers, and muscular disease markers. A preliminary list of codes can be found in Appendix 1 Table 3.

For Objective 4 and 5, clinical manifestations and disease severity including organ involvement will be identified as defined by SNOMED codes. A preliminary list of codes can be found in Appendix 1 Table 2.

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

The following age grouping will be used: 0-1; 2-5; 6-11; 12-17; ; 18-29 30-39; 40-49; 50-59; 60-69; 70-79; 80 and over. The sex (male/ female) of study participants will also be identified. 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.

8.7 Study 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 DM, PM, JDM, and JPM patients. Based on a preliminary feasibility assessment the expected number of DM records in the included databases for this study will be approximately 6,900 (CDWBordeaux 400; EBB 200; CPRD GOLD 1,000; IQVIA DA Germany 3,500; SIDIAP 1,800). For PM, the number of records in the included databases for this study will be approximately 5,400 (CDWBordeaux 300; EBB <100; CPRD GOLD 1000; IQVIA DA Germany 3,500; SIDIAP 600). Juvenile forms potential records are difficult to approximate as they will be defined based on age. Please note that this number is based on the rounded number of PM, DM, JDM and JPM patients in each database with no filter by study period or prior observation restriction.

8.8 Analysis

Table 9. Description of Study Types and Type of analysis

STUDY TYPE	STUDY CLASSIFICATION	TYPE OF ANALYSIS
Disease epidemiology - Population-level characterization	Off-the-shelf (C1)	- Prevalence of disease
Disease epidemiology - Patient-level characterization	Off-the-shelf (C1)	- Large-scale characterisation
Drug epidemiology - Patient-level characterization	Off-the-shelf (C1)	- Patient-level treatment patterns

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

<u>Population-level epidemiology</u>

For objective 1, point prevalence (number of people diagnosed with the disease divided by total population active) of each outcome of interest (DM, PM, JDM, JPM) with complete persistence (an individual is deemed to have the diagnosis from first occurrence until end of follow-up) calculated on an annual basis as of the 1st of January for each year, will be estimated overall and stratified by age and sex alongside with its 95% confidence intervals using the Wilson method. We will also estimate period prevalence as a sensitivity analysis, estimated on an annual basis (between the 1st January and 31st December for each year). Period prevalence will first be estimated with participants required to contribute a minimum of only one day of the period to be included. All analyses will be further stratified by sex and by paediatric and adult populations (0 to less than 2 years, 2 to less than 6 years, 6 to less than 12 years, 12 to less than 18 years, 18 or more years).

<u>Patient-level characterisation</u>

Large-scale patient-level characterisation will be conducted for objectives 2 to 5. Age and sex at time of PM, DM, JDM and JPM diagnosis will be described for each of the generated study cohorts. The index date will be the date of the PM, DM, JDM and JPM diagnosis for each patient. Medical condition, biomarkers and medication use history will be assessed for anytime –and up to 365 days before index date, for 364 to 91, for 90 to 31, and for 30 to 1 day before index date, and at index date. We will also report them for 1 to 90, 91 to 180, 181 to 365 days, 366 to 1095, 1096 to 1825 days, and 1826 days to any time post index date. Occurrence of biomarker measurements, clinical manifestations codes, and disease severity codes specified as described in the outcomes section will be also measured in these timeframes.

<u>Patient-level drug utilisation</u>

For Objective 6, The number and percentage of patients receiving each of a pre-specified list of PM, DM, JDM and JPM treatments as defined in the outcomes section will be described at index date (date of diagnosis), 1 to 90, 91 to 180, 181 to 365 days, 366 to 1095, 1096 to 1825 days, and 1826 days to any time post index date with the denominator being the patients still observed at each time point. Additionally, sunburst plots and Sankey diagrams will be used to describe treatment patterns over time. Sankey diagrams will be censored at end of treatment or end of follow-up.

<u>Drug exposure calculations</u>

Drug eras will be defined as follows: Exposure starts at date of the first prescription after the first PM, DM, JDM and JPM diagnosis. For each prescription, the estimated duration of use is retrieved from the drug exposure table in the CDM. 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 show in 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	$d_1 \cdot x_1 + d_2 \cdot x_2$	$x_1 + x_2$

first exposure, time $= x_1$, dose $= d_1$ | gap, time $= x_{12}$ | second exposure, time $= x_2$, dose $= d_2$
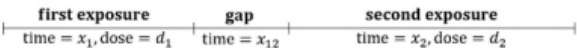
Figure 1. Gap era joint mode

If two eras overlap, the overlap time will be considered exposed by the first era (Figure 2). No time will be added at the end of the combined drug era to account for the overlap. If two eras start at the same date, the overlapping period will be considered exposed by both. We will not consider repetitive exposure.

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

first exposure, dose $= d_1$
overlap
second exposure, dose $= d_2$
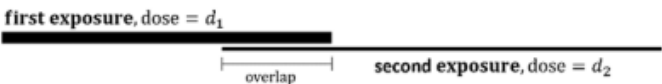
Figure 2. Gap era overlap mode

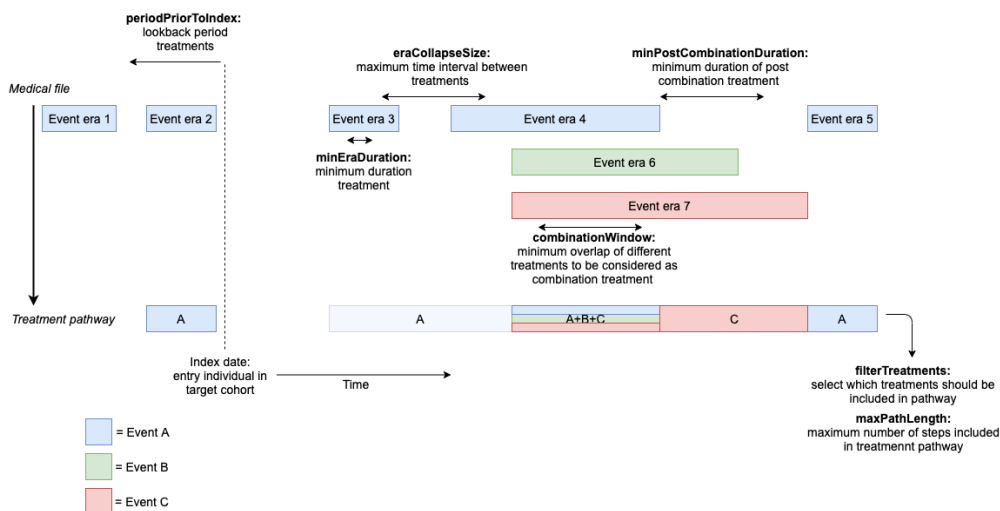To construct treatment pathways, various parameters can be defined in the TreatmentPatterns package (Figure 3).

Figure 3. Parameters in TreatmentPatterns package

The following parameters will be defined in this study.

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	Time (days) that two event eras need to overlap to be considered a combination treatment	30
minPostCombination Duration	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 (2006-2013, 2013-2020, and 2020-2022).

Software

All analyses will be performed in R. "IncidencePrevalence" (https://github.com/darwin-eu/IncidencePrevalence) will be used for the computation of prevalence; "TreatmentPatterns" (https://github.com/darwin-eu-dev/TreatmentPatterns) will be used for the patient-level characterisation of treatments including combination and sequence of therapy.

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 DM, PM, JDM, and JPM, 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 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 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 must be considered. In particular, the identification of PM, DM, JDM and JPM patients and the recording of the co-morbidities may vary across databases and while relatively few false positives would be expected, false negatives may be more likely especially for databases without patient-level linkage to secondary care data. There is scarce data on the validation of rheumatological phenotypes in administrative databases in Europe (10, 11). The possibility of some of the hospitals covered in the datasets being referents for the diagnosis and

treatment of these disease may artificially increase the prevalence, and underreporting in primary care databases if the diagnosis is done in secondary care and not fed back to primary care could artificially underestimate the prevalence.

In addition, the recording of events defined for patient characterisation may vary across databases; and in databases with information on PM, DM, JDM and JPM 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.

12. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

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

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

13. GOVERNANCE BOARD ASPECTS

All data sources require approval from their respective IRB boards, 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

14.0 Study Report

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

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

15. OTHER ASPECTS

N/A

16. REFERENCES

Aggarwal R, Charles-Schoeman C, Schessl J et al. Prospective, double-blind, randomized, placebocontrolled phase III study evaluating efficacy and safety of octagam 10% in patients with dermatomyositis ("ProDERM Study"). Medicine (Baltimore) 2021;100: e23677.

Aggarwal R, Rider LG, Ruperto N, Bayat N, Erman B, Feldman BM, Oddis CV, Amato AA, Chinoy H, Cooper RG, Dastmalchi M, Fiorentino D, Isenberg D, Katz JD, Mammen A, de Visser M, Ytterberg SR, Lundberg IE, Chung L, Danko K, García-De la Torre I, Song YW, Villa L, Rinaldi M, Rockette H, Lachenbruch PA, Miller FW, Vencovsky J; International Myositis Assessment and Clinical Studies Group and the Paediatric Rheumatology International Trials Organisation. 2016 American College of Rheumatology/European League Against Rheumatism criteria for minimal, moderate, and major clinical response in adult dermatomyositis and polymyositis: An International Myositis Assessment and Clinical Studies Group/Paediatric Rheumatology International Trials Organisation Collaborative Initiative. Ann Rheum Dis. 2017 May;76(5):792-801. doi: 10.1136/annrheumdis-2017-211400.

Alexander G S Oldroyd and others, British Society for Rheumatology guideline on management of paediatric, adolescent and adult patients with idiopathic inflammatory myopathy, Rheumatology, Volume 61, Issue 5, May 2022, Pages 1760–1768, https://doi.org/10.1093/rheumatology/keac115.

Ashton C. et al. Idiopathic inflammatory myopathies: a review, Intern Med J. 2021 Jun;51(6):845-852. doi: 10.1111/imj.15358.

Bellutti Enders F, Bader-Meunier B, Baildam E, Constantin T, Dolezalova P, Feldman BM, Lahdenne P, Magnusson B, Nistala K, Ozen S, Pilkington C, Ravelli A, Russo R, Uziel Y, van Brussel M, van der Net J, Vastert S, Wedderburn LR, Wulffraat N, McCann LJ, van Royen-Kerkhof A. Consensus-based recommendations for the management of juvenile dermatomyositis. Ann Rheum Dis. 2017 Feb;76(2):329-340. doi: 10.1136/annrheumdis-2016-209247.

Findlay A.R. et al. An overview of polymyositis and dermatomyositis, Muscle Nerve. 2015 May;51(5):638-56. doi: 10.1002.

Gerami, P., Schope, J. M., McDonald, L., Walling, H. W. & Sontheimer, R. D. A systematic review of adult-onset clinically amyopathic dermatomyositis (dermatomyositis sine myositis): a missing link within the spectrum of the idiopathic inflammatory myopathies. J. Am. Acad. Dermatol. 54, 597–613 (2006).

Lundberg I. E. et al. 2017 European League Against Rheumatism/American College of Rheumatology classification criteria for adult and juvenile idiopathic inflammatory myopathies and their major subgroups. Ann. Rheum. Dis. 76, 1955–1964 (2017). This paper describes the international classification criteria recently revised for IIM.

Lundberg I.E. et al. Idiopathic inflammatory myopathies, Nat Rev Dis Primers. 2021 Dec 2;7(1):86. doi: 10.1038.

Meyer A. and others, Incidence and prevalence of inflammatory myopathies: a systematic review, Rheumatology, Volume 54, Issue 1, January 2015, Pages 50–63, https://doi.org/10.1093/rheumatology/keu289.

Mok, C. C., Ho, L. Y. & To, C. H. Rituximab for refractory polymyositis: an open-label prospective study. J. Rheumatol. 34, 1864–1868 (2007).

Papadopoulou, C., Chew, C., Wilkinson, M.G.L. et al. Juvenile idiopathic inflammatory myositis: an update on pathophysiology and clinical care. Nat Rev Rheumatol 19, 343–362 (2023). https://doi.org/10.1038/s41584-023-00967-9.

Sasaki H. et al. Current diagnosis and treatment of polymyositis and dermatomyositis, Mod Rheumatol. 2018 Nov;28(6):913-921. doi: 10.1080.

Valiyil, R., Casciola-Rosen, L., Hong, G., Mammen, A. & Christopher-Stine, L. Rituximab therapy for myopathy associated with anti-signal recognition particle antibodies: a case series. Arthritis Care Res. 62, 1328–1334 (2010).

17.   ANNEXES

Appendix I: Definition of DM, PM, JDM and JPM Diagnosis, Clinical Manifestations & Complications, Biomarkers and Treatments

Table 1: Preliminary code list for DM.
Note: We will consider three concept sets – narrow definition, broad definition, and prevalent condition. The concept sets are subject to change after review of cohort diagnostics

CONCEPT_ID	CONCEPT_NAME
Dermatomyositis	
80182	Dermatomyositis
4081250	Dermatomyositis sine myositis
4344161	Dermatomyositis with malignant disease
46270398	Disorder of respiratory system due to dermatomyositis
4084268	Idiopathic dermatomyositis
4270868	Adult onset dermatomyositis
Polymyositis	
80800	Polymyositis
4055369	Lung disease with polymyositis
4084780	Idiopathic polymyositis
4346977	Polymyositis associated with autoimmune disease
Juvenile DM: **- Dermatomyositis + age at diagnosis < 18 OR**	
4005037	Childhood type dermatomyositis
40405496	Juvenile dermatomyositis
36674477	Neonatal dermatomyositis
606434	Hypomyopathic juvenile dm
Juvenile PM **- Polymyositis + age at diagnosis < 18 OR**	
42538014	Juvenile polymyositis

Table 2: DM, PM, JDM, JPM Clinical manifestations & Complications

CONCEPT_ID	CONCEPT_NAME
Dermatomyositis, Polymyositis, JDM, JPM	
40642539	Dilated cardiomyopathy due to dermatomyositis (disorder)
4344161	Dermatomyositis with malignant disease (disorder)
	Calcification of muscle due to adult dermatomyositis (disorder)
606376	Calcinosis due to adult type dermatomyositis (disorder)

46270398	Disorder of respiratory system due to dermatomyositis (disorder)
4140598	Sinus tachycardia
4145510	Supraventricular tachycardia
141038	Diastolic dysfunction
40482477	Antisynthetase syndrome
4143972	Raynaud's phenomenon
3531038	Interstitial lung disease
42537657	Interstitial lung disease due to systemic disease (disorder)
4159647	Thromboembolic disorder
4231363	Thrombosis
4121618	Pulmonary embolism
79908	Muscle weakness
442752	Muscle pain
4280820	Calcinosis
31317	Dysphagia (disorder)
4083305	Gottron's papules (finding)
4135412	Periorbital erythema
37311082	Erythematous rash
4314120	Poikilodermatomyositis
4178680	Degenerative disorder of musculoskeletal system
40483721	Autoimmune inflammation of skeletal muscle
4208786	Musculoskeletal and connective tissue disorder
4125793	Acute sarcoid polymyositis
606431	Antisynthetase syndrome due to polymyositis
4055369	Lung disease with polymyositis
44783568	Secondary nonischemic congestive cardiomyopathy
320136	Disorder of respiratory system
73854	Progressive myositis ossificans

Table 3: DM, PM, JDM Biomarkers

CONCEPT_ID	CONCEPT_NAME
Dermatomyositis, Polymyositis, DM y PM	
35624342	AntiMi2
37392868	Extractable nuclear antigen antibody level
37394412	Antibody to extractable nuclear antigen typing
4023909	Extractable nuclear antibody
4163958	Antinuclear antibody
37394192	Creatine – Kinase levels
37393853	ESR = erythrocyte sedimentation rate
4208414	C-reactive protein

4197913	Ca125
40315390	LDH blood level
40462064	AST-GOT blood level
445187004	Ac. Anti-histidyl-tRNA synthetases (Jo1)
4221398	Myoglobine
3946209/42529588/3048888/42870556	Ac-Anti-isoleucil tRNA synthetases (OJ)
42529529	Ac. Anti-glycyl-tRNA synthetases (EJ)
3030939	Ac. Anti-alanyl tRNA synthetase (PL12)
1988758/42529058	Ac. Anti-threonyl RNA synthetase (PL-7)
4012406	Ac. Anti-SAE1
4123191	Polymyositis-scleroderma antibody
37393417	Pm-1 - Polymyositis-scleroderma antibody level
4024409	PM-SCL extractable nuclear antibody
3033229	Ac. Anti-SRP
445187004	Ac. Anti-histidyl-tRNA synthetases (Jo1)
45932003	Histidyl t-RNA synthetase antibody Jo-1 antibody
4124105	Ac. Anti-Ku

Table 4: Preliminary code list for PM, DM, JDM, JPM treatments.

Class	Treatment	ATC code	ConceptID
Antimalarial	Hydroxychloroquine	P01BA02	21604889
	Chloroquine	P01BA02	
Glucocorticoids	Prednisone	A07EA03	1551099
	Methylprednisolone	H02AB04	21602732
Calcineurin inhibitors	Tacrolimus	L04AD02	950637
	Cyclosporine	L04AD01	19010482
DMARDs	Cyclophosphamide	L01AA01	1310317
	Mycophenolate mofetil	L04AA06	4304148
	Azathioprine	L04AX01	45999143
	Methotrexate	L04AX03	1305058
Biologic agents	Rituximab	L01FA01	1314273
	Beta interferone -1a	L03AB07	34297544
	Beta interferone-1b	L03AB08	4305098
	Etanercept	L04AB01	37017299
	Infliximab	L04AB02	4297538
	Adalimumab	L04AB04	4251303
	Basiliximab	L04AC02	4305237
	Eculizumab	L04AA25	4141829
	Abatacept	L04AA24	1186087

| Inmunoglobulines | Inmunoglobulines | J06BA | 40529754 |
| Topical treatment | Brimionidine | S01EA05 | 950637 |

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				5
1.1.1 Start of data collection[1]	☐	☒	☐	
1.1.2 End of data collection[2]	☒	☐	☐	
1.1.3 Progress report(s)	☐	☐	☒	
1.1.4 Interim report(s)	☐	☐	☒	
1.1.5 Registration in the EU PAS Register®	☐	☒	☐	
1.1.6 Final report of study results.	☒	☐	☐	

Comments:

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

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)	☐	☐	☒	
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:

No exposures are described. The use of medicines is described as outcomes in this protocol.

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?	☒	☐	☐	8.6
6.3 Does the protocol address the validity of outcome measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, use of validation sub-study)	☐	☐	☒	
6.4 Does the protocol describe specific outcomes relevant for Health Technology Assessment? (e.g. HRQoL, QALYs, DALYS, health care services utilisation, burden of disease or treatment, compliance, disease management)	☐	☐	☒	

Comments:

Section 7: Bias	**Yes**	**No**	**N/A**	**Section Number**
7.1 Does the protocol address ways to measure confounding? (e.g., confounding by indication)	☐	☐	☒	
7.2 Does the protocol address selection bias? (e.g. healthy user/adherer bias)	☐	☐	☒	
7.3 Does the protocol address information bias? (e.g. misclassification of exposure and outcomes, time-related bias)	☒	☐	☐	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)	☐	☐	☒	8.6
9.1.2 Outcomes? (e.g., clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	☒	☐	☐	8.6
9.1.3 Covariates and other characteristics?	☒	☐	☐	8.6
9.2 Does the protocol describe the information available from the data source(s) on:				
9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☐	☐	☒	8.6
9.2.2 Outcomes? (e.g. date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	8.6
9.2.3 Covariates and other characteristics? (e.g., age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	☒	☐	☐	8.6
9.3 Is a coding system described for:				
9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☒	☐	☐	8.6
9.3.2 Outcomes? (e.g., International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	☒	☐	☐	8.6
9.3.3 Covariates and other characteristics?	☒	☐	☐	8.6
9.4 Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☐	☐	☒	

Comments:

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Section Number**
10.1 Are the statistical methods and the reason for their choice described?	☒	☐	☐	8.8
10.2 Is study size and/or statistical precision estimated?	☐	☐	☒	8.7
10.3 Are descriptive analyses included?	☒	☐	☐	8.8
10.4 Are stratified analyses included?	☒	☐	☐	8.8
10.5 Does the plan describe methods for analytic control of confounding?	☐	☐	☒	
10.6 Does the plan describe methods for analytic control of outcome misclassification?	☐	☐	☒	
10.7 Does the plan describe methods for handling missing data?	☐	☐	☒	
10.8 Are relevant sensitivity analyses described?	☒	☐	☐	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
12.1.2 Information bias?	☒	☐	☐	
12.1.3 Residual/unmeasured confounding? (e.g., anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods).	☐	☐	☒	
12.2 Does the protocol discuss study feasibility? (e.g. study size, anticipated exposure uptake, duration of follow-up in a cohort study, patient recruitment, precision of the estimates)	☒	☐	☐	8.2

Comments:

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

Comments:

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

Comments:

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

Comments:

Name of the main author of the
protocol: Albert Prats Uribe

Date: 23/08/2023

Signature: ALBERT PRATS URIBE