VALSE (VALNAC09344)

Non-Interventional retrospective longitudinal study in the UK and France to investigate the therapeutic strategies after discontinuation of valproate and related substances in clinical practice: VALSE study

Protocol

Version 8.0, 8 November 2022

Sponsor: Consortium of MAHs

Bordeaux PharmacoEpi
Plateforme de recherche en Pharmaco-épidémiologie
CIC Bordeaux CIC1401
INSERM – Université de BORDEAUX – CHU de Bordeaux – Adera
Bâtiment Le Tondu case 41 – 146 rue Léo Saignat – 33076 Bordeaux cedex
BPE platform is certified ISO 9001:2015
for its research activities in pharmacoepidemiology

PASS / General information

TITLE	Non-Interventional retrospective longitudinal study in the UK and France to investigate the therapeutic strategies after discontinuation of valproate and related substances in clinical practice: VALSE study (VALNAC09344)
PROTOCOL VERSION IDENTIFIER	8.0
DATE OF LAST VERSION OF PROTOCOL	15 Oct 2021
EU PAS REGISTER NUMBER	EUPAS 37438
ACTIVE SUBSTANCE	Valproate and related substances: ATC code: N03AG01 and N03AG02
MEDICINAL PRODUCT	Valproate and related substances*: - magnesium valproate - sodium valproate - valproic acid - sodium valproate/ valproic acid - valproate semisodium - valpromide *All substances will be summarized under the term “valproate”
PRODUCT REFERENCE	Information is detailed in the cover letter’s Annex
PROCEDURE NUMBER	EMA/H/A-31/1454
MARKETING AUTHORISATION HOLDER(S) / SPONSOR	APOTEX EUROPE B.V.; ARISTO PHARMA GMBH; ARROW GENERIQUES; BETAPHARM ARZNEIMITTEL GMBH/DR.REDDY'S; BIOGARAN; BIOMO PHARMA GMBH; CONSILIENT HEALTH LIMITED, CRESCENT PHARMA, DESITIN ARZNEIMITTEL GMBH; GENERIS FARMACEUTICA S.A.; G.L. PHARMA GMBH; LUPIN HEALTHCARE, MYLAN SAS; NEURAXPHARM ARZNEIMITTEL GMBH; ORION CORPORATION; PHARMASWISS Ceska republika s.r.o.; SANDOZ/HEXAL AG; SANOFI AVENTIS GROUPE; STADA ARZNEIMITTEL AG; TECNIFAR S.A.; TEVA PHARMACEUTICALS EUROPE; WOCKHARDT UK LIMITED. Of note, Alfasigma and Glenmark left the Consortium on 31 December 2019.
JOINT PASS	Yes
RESEARCH QUESTION AND OBJECTIVES	The research question is to investigate the therapeutic strategies implemented when valproate is discontinued in clinical practice. The objectives and study population will be split for each indication of valproate (epilepsy or bipolar disorder) in the

	overall population of valproate WCBP chronic users and in a subpopulation of pregnant women. The **primary study objective** is to determine the clusters of patients that are the most likely to reflect a success in epilepsy/ bipolar disorder management after valproate discontinuation based on: (i) the description of the overall treatment patterns in the year following valproate discontinuation, (ii) the categorization of patients according to their treatment patterns (clusters), and (iii) the description of patients' and treatment characteristics at baseline, and clinical relapse occurrence, pregnancy occurrence, and other healthcare resources in the follow-up period in each of these clusters. For each cluster, Success/Failure in epilepsy/BD management after valproate discontinuation, will be defined based on the absence of valproate reintroduction in the follow-up period. This will be contextualized according to several clinical and pharmaceutical parameters such as: clinical relapse, number of hospitalizations, polypharmacy. Results will be then discussed with the Scientific Committee to determine which cluster(s) is (are) the most likely to reflect a success in epilepsy/ bipolar disorder management after valproate discontinuation. The **secondary study objectives** is to identify the baseline factors (e.g., patients', Epilepsy/BD treatments, disease characteristics) associated with the potential successful / unsuccessful clusters.
COUNTRY OF STUDY	UK and France
AUTHOR	**Sandrine Colas, on behalf of the Consortium of MAHs** Epidemiology & Benefit- Risk Evaluation, Sanofi R&D sandrine.colas@sanofi.com **Dr Patrick Blin** Chief Scientific Officer, BPE platform patrick.blin@u-bordeaux.fr **Dr Pauline Bosco-Lévy** Scientific Officer, BPE platform pauline.bosco-levy@u-bordeaux.fr

Marketing authorisation holder(s)/ Sponsor

MARKETING AUTHORISATION HOLDER(S)	refer to Annex 3
MAH CONTACT PERSON	Amélie Prieur, Deputy QPPV . SANOFI R&D, 1, avenue P. Brossolette, F - 91385 Chilly Mazarin

1. TABLE OF CONTENTS

List of tables

List of figures

2. LIST OF ABBREVIATIONS

ABE	Augmented Backward Elimination
AEDs	Antiepileptic drugs
AHC	Agglomerative Hierarchical Clustering
ANSM	*Agence Nationale de Sécurité du Médicament et des produits de santé*
APC	Admission Patient Care
ATC	Anatomical Therapeutic Chemical
BD	Bipolar Disorder
BNF	British National Formulary
CI	Confidence Interval
CMDh	Coordination Group for Mutual Recognition and Decentralized Products
CMU-c	*Couverture Maladie Universelle complémentaire*
CNAM	*Caisse Nationale d'Assurance Maladie*
CNIL	*Commission Nationale de l'Informatique et des Libertés*
CPRD	Clinical Practice Research Datalink
CRO	Contract Research Organisation
DDD	Defined Daily Dose
DEP	Data Extraction Plan
EC	European Commission
EEG	Electroencephalogram
EGB	*Echantillon Généraliste de Bénéficiaires*
EMR	Electronic Medical Records
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
ER	Emergency Room
EU PAS	European Union Post Authorisation Studies
GP	General Practitioner
GVP	Guideline on good pharmacovigilance practices
HES	Hospital Episode Statistics
HR	Hazard Ratio

ICD	International Classification of Diseases
ICMJE	International Committee of Medical Journal Editors
ICH	International Conference on Harmonisation
INDS	*Institut National des Données de Santé*
ISAC	Independent Scientific Advisory Committee
ISPE	International Society for Pharmacoepidemiology
IUD	IntraUterine Device
LTD	Long-term disease or ALD "*affection longue durée*" in French, registration for major chronic diseases with full insurance coverage of all claims related to disease
MAH	Marketing Authorisation Holder
MHRA	Medicines and Healthcare products Regulatory Agency
MPR	Medication Possession Ratio
MRI	Magnetic Resonance Imaging
MSA	*Mutualité Sociale Agricole*
OM	Optimal Matching
OP	Outpatient Care
OR	Odd Ratio
PAM	Partitioning Around Medoids
PASS	Post Authorisation Safety Study
PMSI	French national hospital-discharge summaries database system (*Programme de Médicalisation des Systèmes d'Information*)
PRAC	Pharmacovigilance Risk Assessment Committee
QPPV	Qualified Person for PharmacoVigilance
RSI	*Régime Social des Indépendants*
SAP	Statistical Analysis Plan
SAR	Statistical Analysis Report
SAS	Statistical Analysis System software
SC	Scientific Committee
SLM	*Sections Locales Mutualistes*
SNDS	*Système National des Données de Santé*
SSA	State Sequence Analysis

STROBE	STrengthening the Reporting of OBservational studies in Epidemiology
UK	United Kingdom
VPA	Valproate
WCBP	Women of ChildBearing Potential

3. RESPONSIBLE PARTIES

SCIENTIFIC COMMITTEE	
Prof. Dr. Bettina Schmitz	Chefärztin der Klinik für Neurologie, Stroke Unit und Zentrum für Epilepsie Vivantes Humboldt-Klinikum Am Nordgraben 2 13509 Berlin - Germany Bettina.Schmitz@vivantes.de
Prof. Bruno Etain, MD, PhD	Université Paris Diderot Centres Experts Troubles Bipolaires & Dépressions Résistantes, Département de Psychiatrie et de Médecine Addictologique Hôpital Fernand WIDAL 200 rue du Faubourg Saint Denis 75010 Paris - France Bruno.Etain@inserm.fr

COORDINATING CENTRE	
Bordeaux PharmacoEpi (BPE) ***Plateforme de recherche en pharmaco-épidémiologie*** **INSERM CIC1401**	Université de Bordeaux Bât Le Tondu - Case 41 146, rue Léo Saignat 33076 Bordeaux cedex - France ☎ +33 (0)557 574 675 / Fax: +33 (0)557 574 740 http://www.pharmacoepi.eu
Director Founder of BPE Pr. Nicholas Moore	☎ +33 (0)5 57 57 15 60 nicholas.moore@u-bordeaux.fr
Director Cécile Droz-Perroteau	☎ +33 (0)5 57 57 47 37 cecile.droz@u-bordeaux.fr
Chief Scientific Officer Dr. Patrick Blin	☎ +33 (0)5 57 57 95 63 patrick.blin@u-bordeaux.fr
Scientific Officer Dr. Pauline Bosco-Lévy	☎ +33 (0)5 57 57 95 64 pauline.bosco-levy@u-bordeaux.fr
Biostatistics & Data management Chief Régis Lassalle	☎ +33 (0)5 57 57 47 64 regis.lassalle@u-bordeaux.fr
Project Leader Magali Rouyer	☎ +33 (0)5 57 57 47 67 magali.rouyer@u-bordeaux.fr
Assistant Project Leader Adeline Grolleau	☎ +33 (0)5 57 57 92 03 adeline.grolleau@u-bordeaux.fr
Study Manager Emmanuelle Bignon	☎ +33 (0)5 57 57 92 22 emmanuelle.bignon@u-bordeaux.fr
Senior statistician Marie-Agnès Bernard	☎ +33 (0)5 57 57 48 56 marie-agnes.bernard@u-bordeaux.fr
Statistician Clélia Favary	☎ +33 (0)5 47 30 43 15 clelia.favary@u-bordeaux.fr

SPONSOR

Consortium of MAHs (Annex 3)	
Epidemiology Analytist	Cindy Li Epidemiology Analytics Sanofi Aventis 55 Corporate Drive, Bridgewater NJ 08807, USA Xinyu.Li@sanofi.com

4. ABSTRACT

TITLE	VALSE (VALNAC09344): Non-Interventional retrospective longitudinal study in the UK and France to investigate the therapeutic strategies after discontinuation of valproate and related substances in clinical practice
RATIONALE AND BACKGROUND	Valproate and related substances have been licensed since 1967 to treat epilepsy and since 1995 to treat bipolar disorder in Europe. In March 2017, a referral under Article 31 of Directive 2001/83/European Commission was initiated and the Pharmacovigilance Risk Assessment Committee (PRAC) assessed the impact of the risk minimization measures in the current pregnancy exposure to the treatment with medicinal products containing substances related to valproate and their impact on the benefit-risk balance. Several consultations including a Public Hearing and two Scientific Advisory Group meetings with Neurologists and Psychiatrists were held in September and October 2017. During these consultations, considerations were discussed with clinicians about the case when a woman of childbearing potential (WCBP) (aged 13 to 49 years) treated with valproate is unable to comply with an effective contraception method or is willing to become pregnant or finds out she is pregnant. It was highlighted that the currently available recommendations regarding switching or discontinuation of valproate are insufficient. The outcome of the Referral procedure was approved on 31 May 2018, recommended new restrictions on the use of valproate and set-up of a pregnancy prevention program. Also, further studies to characterise the nature and extent of the risks posed by valproate are imposed to all Marketing Authorisation Holders (MAHs). Among those, a retrospective study that aims to evaluate and identify the best practices for therapeutic management after valproate discontinuation in clinical practice was proposed.
RESEARCH QUESTION AND OBJECTIVES	The **research question** is to investigate the therapeutic strategies implemented when valproate is discontinued in clinical practice. The objectives and study population will be split for each indication of valproate (epilepsy or bipolar disorder) in the overall population of valproate WCBP chronic users and in a subpopulation of pregnant women. The **primary study objective** is to determine the clusters of patients that are the most likely to reflect a success in epilepsy/ bipolar disorder management after valproate discontinuation based on: (i) the description of the overall treatment patterns in the year following valproate discontinuation, (ii) the categorization of patients according to their treatment patterns after valproate discontinuation (clusters), and (iii) the description of patients' and treatment characteristics at baseline, and clinical relapse occurrence, pregnancy occurrence, and other healthcare resources in the follow-up period in each of these clusters. For each cluster, Success/Failure in epilepsy/Bipolar Disorder (BD) management after valproate discontinuation, will be defined based on the absence of valproate reintroduction in the follow-up period. This will be contextualized according to several clinical and pharmaceutical parameters such as: clinical relapse, number of hospitalizations, polypharmacy. Results will be then discussed with the Scientific Committee to determine which cluster(s)

is (are) the most likely to reflect a success in epilepsy/ bipolar disorder management after valproate discontinuation.

The **secondary study objectives** is to identify the baseline factors (e.g., patients', Epilepsy/BD treatments, disease characteristics) associated with the potential successful / unsuccessful clusters.

STUDY DESIGN

This is a cohort study of WCBP, chronic users of valproate for either epilepsy or bipolar disorder, who have discontinued valproate during the inclusion period, with a follow-up of one year after inclusion.

This study will be conducted with secondary data from an English electronic medical records database, the CPRD (Clinical Practice Research Datalink) and from the French nationwide claims database, the SNDS (*Système National des Données de Santé*).

The index date will be defined at the last supply day of the last valproate prescription/dispensing before discontinuation during the inclusion period from January 1, 2014, to December 31, 2017, in each database.

Each woman will have a pre-index period of at least 1 year up to 5 years before the index date depending on the availability of patient historical data in each database (i.e., 5 years in SNDS, and from 1 to 5 years in CPRD). Each woman will be followed for 1-year after the index date, or until the date of death or database eligibility lost, whichever came first. Therefore, data will be extracted from January 1, 2009, to December 31, 2018 (full study period) for patients identified in CPRD database and from January 1, 2009, to December 31, 2019, for patients identified in SNDS database to ensure a period of 9 months in addition to the 1-year of follow-up, necessary to accurately define all pregnancies, the delivery date being included in the pregnancy identification algorithm.

POPULATION

The study population will include all WCBP, i.e., aged 13 to 49 years, chronic users of valproate who have discontinued valproate during the inclusion period.

Pregnant subpopulation will be all women of the study population who will be pregnant during the inclusion or follow-up period, i.e., with an estimated date of pregnancy start between 9 months prior to index date and the end of the 1-year follow-up (using data sources-specific information and pregnancy identification algorithm).

Study population and pregnant subpopulation will be categorized in two cohorts using the sources of diagnoses available in each database (hospitalization, outpatient or Long-Term Disease diagnoses):

Epilepsy cohort will include all women with at least 1 diagnosis code recorded for epilepsy during the pre-index period.

Bipolar disorder cohort will include all women with at least 1 diagnosis code recorded for bipolar disorder (either Schizoaffective disorder or Bipolar affective disorder or Manic episode) during the pre-index period.

Will be excluded from the cohort:

- Patients with less than 1 year of historical data (prior to index date) in the database.
- Patients with both epilepsy and bipolar disorder.

VARIABLES

Exposure

- Drugs of interest:
  - Valproate and its related substances,
  - Other antiepileptic drugs,
  - Drugs indicated in bipolar disorder (antidepressants, mood stabilizers and neuroleptic drugs).
- Valproate exposure (before index date and during follow-up):
  - Index date (i.e., date of last valproate prescription/dispensing plus the number of day's supply of the last prescription/dispensing).
  - Chronic use of valproate (i.e., being continuously exposed to valproate during the year before the index date: MPR > 60% and no valproate discontinuation) followed by a valproate discontinuation after a 60-day grace period,
  - Valproate dose-tapering phase before valproate discontinuation.
  - Valproate discontinuation (i.e., absence of new prescription/dispensing of valproate for at least 60 days after the days supplied by the last prescription/dispensing) assessed before index date and after valproate reintroduction.

Main outcomes

Given the complexity of the treatment patterns that can be observed in real-life settings, a description of the overall treatment patterns after valproate discontinuation at a macro-level is necessary to have an overview of the situation. To this end, different treatment sequences are defined below:

- Monotherapy: only 1 drug of interest for epilepsy/bipolar disorder (excluding valproate);
- Double therapy: 2 distinct drugs of interest for epilepsy/bipolar disorder (excluding valproate);
- Combination therapy: ≥3 distinct drugs of interest for epilepsy/bipolar disorder (excluding valproate);
- Different monotherapy: 1 medication that differs from the previous sequence treatment;
- Different double therapy: ≥ 1 of the 2 medications that differs from the previous sequence treatment;
- Different combination therapy: ≥ 1 of the medications of the combination therapy that differs from the previous sequence treatment;
- Valproate reintroduction;
- Monotherapy + Valproate reintroduction;
- Double therapy + Valproate reintroduction;
- Combination therapy + Valproate reintroduction;
- Different monotherapy + Valproate reintroduction;
- Different double therapy + Valproate reintroduction;

- Different combination therapy + Valproate reintroduction;
- Not exposed to any epilepsy/bipolar disorder medication (including valproate);

These **treatment sequences** will be identified at regular intervals, based on the frequency of prescription/dispensing of the drugs of interest (i.e., 1 month in SNDS), from the day after index date to the end of the follow-up (i.e., end of the 1-year follow-up period, death or database eligibility lost, whichever comes first).

This overview of the overall treatment patterns will be represented graphically, using a sequence index plot, to illustrate the succession of treatment sequences over time for each patient.

This step will be then followed by the determination of clusters defined as groups of patients sharing homogeneous treatment patterns of the different predefined sequences according to time periods, using an unsupervised clustering method.

All the identified clusters and their patient's characteristics will be reviewed by two independent clinical experts (Scientific Committee), each of them being specialist in each studied disease (epilepsy and bipolar disorders) to determine which ones are the most relevant according to their experience in clinical practice. Relevant cluster(s) that is (are) the most likely to reflect a success in epilepsy/bipolar disorder management after valproate discontinuation will be identified based on the non-reintroduction of valproate, contextualized with other factors.

Indeed, the success or failure after valproate discontinuation will be contextualized according to several clinical and pharmaceutical parameters such as the number of hospitalizations, polypharmacy. This list of parameters may be supplemented by SC experts with regard to the results and their interpretation.

Secondary outcomes

The following **secondary outcomes** will be assessed during the year of follow-up for either epilepsy or bipolar disorder in the overall cohort and by clusters, depending on data availability and completeness in each database, to complement the main outcome:

- First occurrence of valproate reintroduction.
- Occurrence of clinical relapse.
- Occurrence of pregnancy.
- Hospitalization and discharge diagnoses.
- Emergency room (ER) visits (CPRD and SNDS) and diagnoses, unscheduled care (CPRD).
- Number of office visits to GP, Neurologist or Psychiatrist or other medical specialty or other relevant health care professionals (CPRD and SNDS).
- Death from any cause (CPRD and SNDS).
- Sick-leave days (SNDS).

Other variables

The following characteristics will be considered according to availability and completeness in each database:

- Patient characteristics: age at index date, deprivation index, region (CPRD and SNDS).
- Disease characteristics: Age at diagnosis date and duration of disease (CPRD); Number and cumulative length of stay of hospitalizations related to each indication; Time elapsed between last clinical event (hospital admission/discharge specific to each indication) and index date (CPRD and SNDS).
- Psychotherapy / Psychosocial support or education (CPRD).
- History of anti-epileptic or antidepressants, mood stabilizers or neuroleptic treatment (CPRD and SNDS).
- Record of contraceptive method; record of folate prescription (CPRD and SNDS).
- Psychiatric diagnosis associated with a consultation or hospital admission or discharge before index date (CPRD and SNDS).
- Epilepsy-related diagnosis associated with hospital admission or discharge before index date (CPRD and SNDS).
- Somatic comorbidities other than epilepsy (CPRD and SNDS).
- History of suicide attempt/ self-injury (CPRD and SNDS).
- History of previous pregnancies; history of fertility treatment/visit to fertility clinic (CPRD, 5-year history for SNDS).

DATA SOURCE

All data will be obtained from two European databases:

- CPRD (The Clinical Practice Research Data Link Gold, plus Hospital Episodes Statistics (HES) linked Data) from the UK.
- SNDS (*Système National des Données de Santé*) from France, the nationwide Claims and hospital database.

STUDY SIZE

In UK, the prevalent use of valproate was approximately estimated at 25 to 30 per 10,000 women aged 18-45 years and 10 to 15 per 10,000 women aged 12-17 years over 2010-2016 (MHRA data). In France, the prevalent use of valproate decreased in WCBP from 122,382 in 2007 to 83,712 by 2015.

The expected number of valproate WCBP chronic users who have discontinued valproate during the inclusion period (from January 1, 2014 to December 31, 2017) will be a result of the study analysis.

The sample size of each cluster will be the output of the clustering analysis, and the precision of each criterion included in definition of success will be also estimated according to the size of clusters.

DATA ANALYSIS

The following analyses will be performed for the epilepsy and bipolar disorder cohorts and the corresponding subpopulation of pregnant women:

- A flowchart depicting patients' identification and inclusion.

- A description of patient's characteristics at index date and during pre-inclusion period (comorbidities, treatments, pregnancy…).
- A description of the overall treatment patterns during the follow-up using a sequence index plot.
- An identification of patients with similar treatment sequences using an unsupervised clustering method.
- For each cluster, a description of patients' characteristics, reviewed by clinical experts of the Scientific Committee who will assess the relevance of selected clusters.
- For each cluster, a description of the treatment sequences by the type of medications using appropriate figures (e.g., Sankey diagram, sunburst, etc.).
- The proportion of patients with the primary or the secondary outcomes during follow-up will be assessed according to treatment patterns (where applicable).
- A description (number of occurrences and proportion of patients with at least one occurrence) of each secondary outcomes (clinical relapse, valproate reintroduction, occurrence of pregnancy in the subgroup of pregnant women) during the follow-up in the overall cohort and by cluster.
- A description of the other outcomes during follow-up in the overall cohort and by cluster: hospitalization, ER visits, medical visits, other relevant health care professionals visits, death, sick leave….
- The identification of covariates associated with the most relevant clusters (multivariable multinomial logistic regression model), reviewed by clinical experts of the Scientific Committee. Results will be expressed as odd ratios (OR) with 95% CI and Wald test p-value. The quality of model will be estimated using the Akaike criterion.

MILESTONES

• Study Protocol submission to PRAC	28 November 2018
• Study Protocol updated submission to PRAC	(V2.0) 09 August 2019 (V3.0) 18 December 2019 (V4.0) 11 May 2020 (V5.0) 21 April 2021 (V6.0) 15 October 2021 (V7.0) 1 July 2022 (V8.0) November 2022
• PRAC initial endorsement of the Study Protocol	09 July 2020
• <Registration in the EU PAS register>	30 September 2020

• Statistical Analysis Plan (SAP)	(V1.0) UK (CPRD data) and France (SNDS data) - 08 December 2020 (V2.0) – UK (CPRD data) 17 September 2021 and France (SNDS data) 24 January 2022 (V3.0) – UK (CPRD data) and France (SNDS data) - 31 January 2023
• Start of data extraction (CPRD)	29 March 2021
• Regulatory aspects and data extraction follow-up with CNAM (SNDS)	August 2021
• Interim report submission to PRAC	
○ CPRD section	09 October 2021
○ SNDS section	31 January 2023
• Final report of study results submission to PRAC	31 July 2023

5. AMENDMENTS AND UPDATES

Version	Date	Summary of changes
Amendment 01 (Protocol Version 5.0)	21 April 2021	Updated due date for the first interim report. The following sections are updated: Section 4, 5, 6 and Annex 1.
Amendment 02 (Protocol version 6.0)	15 October 2021	Updated due date for the second interim report and the final report. The following sections are updated: Section 4, 5, 6 and Annex 1.
Amendment 03 (Protocol version 8.0)	November 2022	Updated due date for the second interim report and the final report. Main changes concern primary outcome definition, new definition of VPA discontinuation with extended grace period, and overall treatment patterns observed in real-life settings throughout the whole follow-up period following valproate discontinuation, without attempting to categorized patients beforehand.

To note: the protocol version 7.0 was not approved by the PRAC, the modifications which had been made to it are therefore not considered.

Table summarizing all substantial modifications of amendment 03

Number	Date	Section of study protocol	Amendment or update	Reason
03	Nov 2022	Title PASS	Title of the Study	To be aligned with the main research question.
03	Nov 2022	5. Amendments and update	Amendments' table updated	Section updated to summarize all substantial amendments.
03	Nov 2022	6. Milestones	Update of milestones	Update of milestones agreed with PRAC.
03	Nov 2022	PASS 7.1. Background 7.2. Rationale, 8. Research question and objectives	Update of research question	To better reflect the new methodological approach proposed.
03	Nov 2022	PASS 8. Research question and objectives 9.5 Study size (1st §)	Update of objectives	To account for real word clinical practice.

		9.7.1 Generalities (of Data analysis)		
03	Nov 2022	9.1. Study design 9.3.1. Exposure	New definition of index date and grace period.	Index date revised to describe overall treatment patterns since the last supply day of the last valproate prescription/dispensing before discontinuation, and grace period (30 days replaced by 60 days) extended to better define the discontinuation of valproate.
03	Nov 2022	9.3 Variables 9.7.3 Treatment patterns	Valproate exposure and outcomes were updated.	To align with the research question and primary objectives.
03	Nov 2022	9.5. Study size	Removal of the section related to the calculation of the sample size based on proportion of successful switch. Adding of a table providing information on the precision of a success criteria according to the size of clusters	The sample size of each cluster will be the output of the clustering analysis. The precision of each criterion included in definition of success is also estimated according to the size of clusters.
03	Nov 2022	Data analysis 9.7.1. Generalities (of Data analysis) 6. Milestones, (last §)	Adding of R software Update of interim and final reports contents'	R Software is needed to perform clustering analysis. Content of interim and final reports updated according to the research question and primary objectives newly defined.
03	Nov 2022	Data analysis 9.7.4. Outcomes (of Data analysis)	Section updated and split into primary and secondary outcomes	This update is justified by the methodological choices regarding primary and secondary outcomes. Precision was added about the involvement of Scientific committee.
03	Nov 2022	Data analysis 9.7.5. Identification of risk factors (Data analysis)	Replacing of risk factors by covariates and update of statistical methods used for analysis. The section related to the multivariable Poisson regression was removed.	Update according to the newly defined outcomes. Adding of methodological details on modelling strategy (exploratory and conditioned by the size of the clusters and the number of candidate covariates), and Scientific Committee role.

03	Nov 2022	9.7.6. Missing data	Adding information about missing data	To clarify how missing data were approached.

6. MILESTONES

The conditions to market authorization stated in the Coordination Group for Mutual Recognition and Decentralized Products (CMDh) approval states the following reporting requirement: the study protocol is to be submitted in accordance with Article 107n (1) of Directive 2001/83/EC within 6 months after CMDh agreement/ Commission decision (received on 31 May 2018). The first interim report shall be submitted to the PRAC within 12 months after endorsement of the study protocol. Further interim reports, if any, should be submitted to the PRAC 6-monthly thereafter for the first 2 years and the final study report shall be submitted to the PRAC within 48 months after endorsement of the study protocol.

MILESTONES	PLANNED DATE
Study Protocol submission to PRAC	28 November 2018
Study Protocol updated submission to PRAC	(V2.0) 09 August 2019 (V3.0) 18 December 2019 (V4.0) 11 May 2020 (V5.0) 21 April 2021 (V6.0) 15 October 2021 (V7.0) 1 Jun 2022 (V8.0) November 2022
PRAC initial endorsement of the Study Protocol	09 July 2020
<Registration in the EU PAS register>	30 September 2020
Statistical Analysis Plan (SAP)	(V1.0) UK (CPRD data) and France (SNDS data) 08 December 2020 (V2.0) UK (CPRD data) 17 September 2021, and France (SNDS data) 24 January 2022 (V3.0) – UK (CPRD data) and France (SNDS data) 31 January 2023
Start of data extraction (CPRD)	29 March 2021
Regulatory aspects and data extraction follow-up with CNAM (SNDS)	August 2021
Interim report (IR) submission to PRAC: • IR1: CPRD section • IR2: SNDS section	09 October 2021

	31 January 2023
Final report of study results submission to PRAC	31 July 2023

The interim reports will present the results related to the description of the treatment patterns, the characterization of the different clusters identified in each database (if available) during follow-up. The final report will present the overall results including those related to the identification of baseline factors (e.g., patients', Epilepsy/Bipolar Disorder (BD) treatments, disease characteristics) associated with the potential successful / unsuccessful clusters.

7. RATIONALE AND BACKGROUND

7.1. Background

Valproate and related substances have been licensed since 1967 to treat epilepsy and since 1995 to treat bipolar disorder in Europe.

In October 2014, the PRAC concluded a review under Article 31 of Directive 2001/83/EC of all available data from published literature, spontaneous reports as well as the views of the relevant experts on the safety and efficacy of valproate and related substances in women of childbearing potential (WCBP) and pregnant women, due to the risk of malformations and developmental disorders in babies exposed to valproate in utero. Restrictions on the use of valproate containing substances were applied to the label and risk minimization measures such as educational materials for physicians and patients were implemented.

In March 2017, a referral under Article 31 of Directive 2001/83/EC was initiated and the Pharmacovigilance Risk Assessment Committee (PRAC) assessed the impact of these risk minimization measures in the current pregnancy exposure of the treatment with medicinal products containing substances related to valproate and their impact on the benefit-risk balance.

Several consultations including a Public Hearing and two Scientific Advice Group meetings with Neurologists and Psychiatrists were held in September-October 2017. As part of the referral procedure, switching or discontinuation of valproate has been discussed with clinical experts in consultation meetings. It has been agreed that the currently available recommendations provide limited information for switching or discontinuation in clinical practice.

In Epilepsy, the Task Force of the European Academy of Neurology provided a recommendation regarding switching of valproate in WCBP, which is based on expert opinion (1). According to this recommendation, the switch of valproate to an alternative treatment will commonly occur over at least 2-3 months. The new medication is usually introduced as add-on to valproate until a potentially effective dose of the second drug has been achieved and after this, an attempt can be made to gradually taper down and discontinue valproate. No specific recommendations are available regarding switching of valproate during pregnancy. It is acknowledged that switching from valproate during pregnancy is a very challenging clinical issue and discontinuation during pregnancy might not be possible. Indeed, a study based on the EURAP registry (observational international registry of AEDs) and pregnancy, reported a signal that withdrawal of or switch from valproate during the first trimester could lead to loss of seizure control (1). Furthermore, it was reported that the maternal death rate from epilepsy among women with epilepsy was estimated at 100 per 100,000 maternities (2). Uncontrolled epilepsy, particularly in young people, carries a risk of Sudden Unexpected Death and so leaving seizures, especially tonic-clonic seizures, uncontrolled is not an acceptable option during pregnancy. For some patients, particularly those with idiopathic generalized epilepsy, valproate may prove to be the only successful treatment that will control their seizures. In some circumstances, tonic-clonic seizures may cause miscarriage, trauma related to falls and blood conditions that can affect the developing baby.

In bipolar disorder, no specific recommendations are available regarding switching valproate in patients with diagnosis of bipolar disorder. Based on the available general recommendation regarding discontinuation of mood stabilisers (3) in women with bipolar disorder, allowing sufficient time for a gradual discontinuation and close monitoring of the patients for possible relapse seem to be also relevant for discontinuation of valproate. In addition, no specific guidelines of switch according to the reason for switch (lack of efficacy, tolerance and type of side effect, pregnancy…) are available and clinical management mainly rely on empirical approaches.

The need to obtain more information and to identify the best practices regarding the switching or discontinuation of valproate was discussed and several approaches were considered. It is not possible to conduct clinical trials in pregnant women with epilepsy or bipolar disorder due to methodological and ethical considerations. Experiences from clinical practice might provide helpful information on switching and discontinuation of valproate in WCBP and pregnant women in real-life practice as well as insight to identify the best practices for discontinuation or switch of valproate

in both indications and support the release of useful guidelines. It was therefore proposed to conduct an observational retrospective study in order to generate real-world evidence related to overall specific treatment patterns observed after valproate discontinuation to determine which treatment pattern is more likely to reflect the success or the failure of a therapeutic management in both epilepsy and bipolar disorder populations.

On 08 February 2018, the PRAC issued a recommendation including revised prescribing conditions in WCBP in the product information, a pregnancy prevention program and revised educational materials. Also, further studies to characterise the nature and extent of the risks posed by valproate are imposed to all MAHs. The CMDh endorsed the PRAC recommendations on 21 March and the EC decision was adopted on 31 May 2018. Among those, the MAHs of medicinal products with substances related to valproate shall conduct an observational study to describe and identify treatment patterns that are more likely to reflect the best practices for therapeutic management after valproate discontinuation in clinical practice.

When the results of the proposed study will become available, the data will be shared with EU experts in epilepsy and in bipolar disorder to potentially complement the expert consensus guidelines and also, to update the recommendations described in the Health Care Professional guide regarding treatment management after valproate discontinuation.

7.2. Rationale

For such a study, the requirements are:

- A sample with a sufficient size to allow the description in a large number of cases of a relatively rare event or characteristic,
- An accurate prescription chronology including start and end dates of various medications in the indications of interest,
- Specific and complete records of diagnoses.

The UK is one of the European countries with the highest prevalent use of valproate (approximately 25 to 30 per 10,000 women aged 18-45 years and 10 to 15 per 10,000 women aged 12-17 years over 2010-2016 (MHRA data)). In terms of disease management, General Practitioners (GPs) are the health care “gatekeeper” and have the major responsibility for medication continuation or change. The CPRD is an electronic medical records database of about 700 primary care physicians.

France is the European country with the highest prevalence of valproate use although the frequency of exposure among WCBP has decreased (from 122,382 in 2007 to 83,712 by 2015) according to SNIIRAM data (*Système national d’information inter-régimes de l’Assurance maladie*) from National Agency for the Safety of Medicines and Health Products (ANSM). The claims for reimbursement of ambulatory care for all the health insurance schemes are centralized in a unique national medico-administrative database, as well as all public and private hospitalizations.

Both data sources are widely used in pharmacoepidemiology and are recognized as valid for potential for use in medicines regulation (4).

Therefore, due to the large use of valproate in UK and France among the European countries and the availability of health database in these both countries, CPRD and SNDS databases were chosen for reach the study objectives.

This is a mandatory non-interventional PASS imposed as an obligation that is to be conducted according to the Guideline on good pharmacovigilance practices (GVP) – Module VIII (Rev 3) dated 9 October 2017 (EMA/813938/2011 Rev 3) and the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology (rev 6, EMA/95098/2010).

The study rationale led to the reason for considering valproate continuous users as the source population (chronic users). The specific treatment patterns should be described in a population of

female patients who were considered as stable in term of valproate treatment and who had to discontinue possibly for trying alternative therapies or planning a pregnancy. It was evaluated that a one-year follow-up would be sufficient to have an adequate overview of the most frequent treatment patterns.

Based on experiences from clinical practice, most of switches, when they occur, are often before valproate is fully discontinued. In both indications, it is critical to have the patient controlled with medication(s) so if decided, valproate dosage would be most often gradually decreased while the new medication(s) is/are increasingly administered. As a result, the initiation of other medication(s) may occur up to some months before valproate discontinuation. In few cases, the switch or discontinuation of valproate would have to be managed over a few weeks or days, for instance in case the patient is found to be pregnant during a consultation. Still in these “urgent” situations, it is unlikely that the clinician would take the risk of a treatment gap and would prescribe a new therapy in parallel to decreasing valproate in a few days.

In both cases, the time-period for the switch depends on various factors such as patient’s history of disease and clinician/patient decision. In addition, history of previous treatment(s) (within indication) and previous clinical events (e.g., time since last seizure, which may reach several years) may be taken into account. As a result, the study period and particularly the pre-index date period must be optimized, meaning go back as far as possible in each database. However, due to different capabilities, they will be specific to each data source.

8. RESEARCH QUESTION AND OBJECTIVES

The study aims to is to investigate the therapeutic strategies implemented when valproate is discontinued in clinical practice for WCBP. The characteristics of patients, the disease management and the outcomes of interest are substantially different between the population with epilepsy or bipolar disorder. Therefore, the objectives and study population will be split for each indication of valproate use, in the overall population of valproate WCBP chronic users and in a subpopulation of pregnant women.

The **primary study objective** is to determine the clusters of patients that are the most likely to reflect a success in epilepsy/ bipolar disorder management after valproate discontinuation based on: (i) the description of the overall treatment patterns in the year following valproate discontinuation, (ii) the categorization of patients according to their treatment patterns after valproate discontinuation (cluster), and (iii) the description of patients’ and treatment characteristics at baseline, and clinical relapse occurrence, pregnancy occurrence, and other healthcare resources in the follow-up period in each of these clusters.

For each cluster, Success/Failure in epilepsy/BD management after valproate discontinuation, will be defined based on the absence of valproate reintroduction in the follow-up period. This will be contextualized according to several clinical and pharmaceutical parameters such as: clinical relapse, number of hospitalizations, polypharmacy. Results will be then discussed with the Scientific Committee to determine which cluster(s) is (are) the most likely to reflect a success in epilepsy/ bipolar disorder management after valproate discontinuation.

The **secondary study objectives** is to identify the baseline factors (e.g., patients’, Epilepsy/BD treatments, disease characteristics) associated with the potential successful / unsuccessful clusters.

9. RESEARCH METHODS

9.1. Study design

This is a cohort study of WCBP chronic users of valproate for either epilepsy or bipolar disorder, who have discontinued valproate during the inclusion period, with a follow-up of up to one year after inclusion. This study will be conducted with secondary data from an English electronic medical records database, the CPRD (Clinical Practice Research Datalink) or the French nationwide claims, the SNDS (*Système National des Données de Santé*).

The index date will be defined at the last supply day of the last valproate prescription/dispensing before discontinuation (defined in §9.3.1) during the inclusion period (a patient must be included only once) from January 1, 2014, to December 31, 2017 (4 years) in each database.

Each woman will have a pre-index period of at least 1 year up to 5 years before the index date depending on the availability of patient historical data in each database (i.e., 5 years for all included patients in SNDS, and from 1 to 5 years for patients included in CPRD). Each woman will be followed for 1 year after the index date, or until the date of death or database eligibility lost, whichever came first. Therefore, data will be extracted from January 1, 2009, to December 31, 2018, for patients identified in CPRD database and from January 1, 2009, to December 31, 2019, for patients identified in SNDS database to ensure a period of 9 months in addition to the 1-year of follow-up, necessary to accurately define all pregnancies, the delivery date being included in the pregnancy identification algorithm.

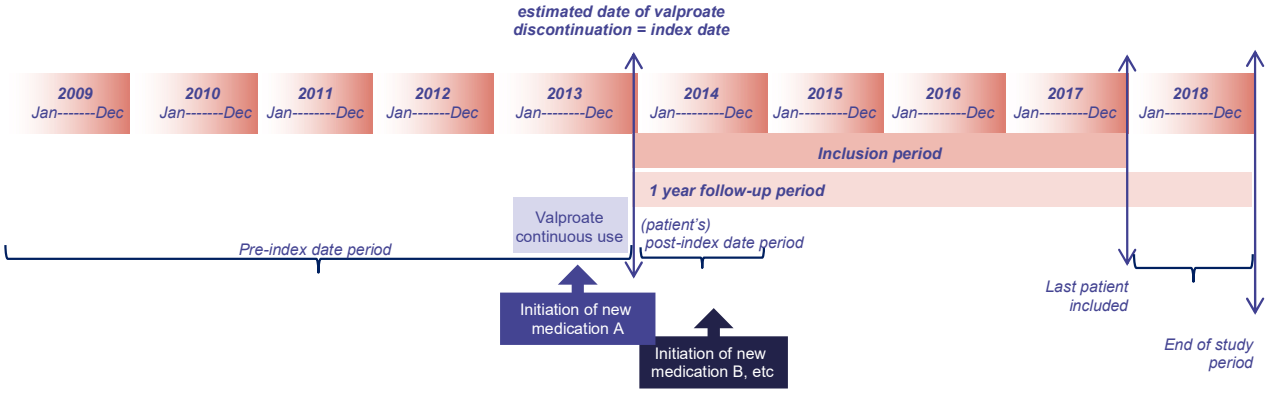

Figure 1. Study design

Note: this scheme illustrates the study periods to apply to CPRD database. For patients identified in SNDS database, data will be extracted in 2019 in addition to the other study periods to ensure an accurate identification of all pregnancies.

9.2. Setting

The study population will be defined using the following inclusion and exclusion criteria:

9.2.1. Inclusion criteria

The **study population** will include all WCBP, i.e., aged 13 to 49 years, chronic users of valproate who have discontinued valproate during the inclusion period.

Pregnant subpopulation will be all women of the study population who will be pregnant during the inclusion or follow-up period, i.e., with an estimated date of pregnancy start between 9 months prior to index date and the end of the 1-year follow-up (using data sources-specific information and pregnancy identification algorithm).

Study population and corresponding pregnant subpopulation will be categorized in two cohorts using the sources of diagnoses available in each database (hospitalization, outpatient or Long-Term Disease diagnoses):

- Epilepsy cohort will include all women with at least 1 diagnosis code recorded for epilepsy (refer to code list in appendix 3) during the pre-index period.
- Bipolar disorder cohort will include all women with at least 1 diagnosis code recorded for bipolar disorder (either Schizoaffective disorder, Bipolar affective disorder, Manic episode; refer to code list in appendix 3) during the pre-index period.

Additional items specific to each database could be used to improve identification of these two groups. For instance, in France, the brand name of valproate is different according to its indication (Depakine® for epilepsy and Depakote® or Depamide® for bipolar disorder). In the UK, the brand names are Epilim® for epilepsy and Depakote® for bipolar disorder.

9.2.2. Exclusion criteria

For both cohorts, exclusion criteria will be:

- Patients with less than 1 year of historical data (prior to index date) in the database.
- Patients with both epilepsy and bipolar disorder.

9.3. Variables

9.3.1. Exposure

9.3.1.1. Drugs of interest

Exposure will be examined as the use of the following drugs of interest, which will be identified according to the prescription in the CPRD (British National Formulary [BNF]) and pharmacy dispensing in SNDS (Anatomical Therapeutic Chemical [ATC] Classification) during the study period:

- Valproate and its related substances will include magnesium valproate, sodium valproate, valproic acid, sodium valproate/ valproic acid, valproate semisodium, valpromide (refer to code list in appendix 3). All substances will be summarized under the term “valproate”.
- Other antiepileptic drugs (listed in appendix 3).
- Antidepressants, mood stabilizers and neuroleptic drugs (listed in appendix 3).

9.3.1.2. Valproate exposure (before index date and during follow-up)

Valproate exposure will be examined before index date and during follow up as follows:

- **Index date** will be defined as the date of last valproate prescription/dispensing plus the number of day’s supply of the last prescription/dispensing of valproate (Figure 2).
- **Valproate chronic users** will be defined as patients being continuously exposed to valproate during the year before the index date, i.e., with a 1-year Medication Possession Ratio (MPR) > 60% and no valproate discontinuation (as defined below) within the 1-year period prior to index date. In the calculation of the 1-year MPR, the numerator will be the total number of days of valproate supply during the 1-year period prior to index date and the denominator the total number of calendar days (365 days).
- **Valproate dose-tapering phase** will be identified before index date depending on data availability in each database. The dose tapering phase will be defined as the period between the start of the last VPA dose reduction and the index date. For CPRD, all prescriptions

within the dose-tapering phase must be actual daily doses (can't be imputed). For SNDS, daily doses will be estimated using the dispensing information, see detail calculation in the SAP.

- **Valproate discontinuation** will be defined as the absence of new prescription/dispensing of valproate for at least 60 days after the days supplied by the last prescription/dispensing, and will be assessed before index date and after valproate reintroduction.

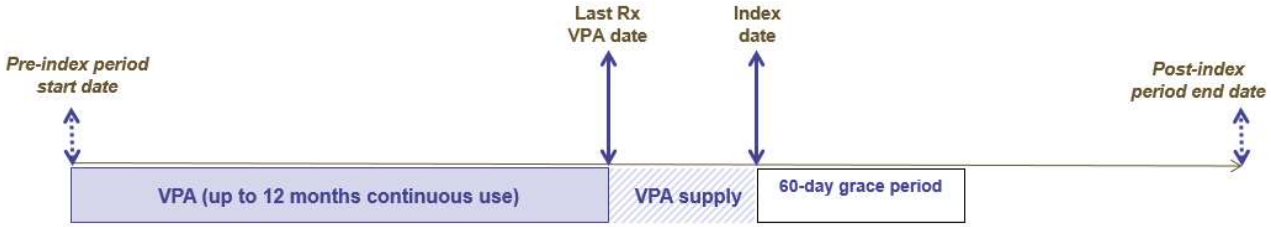

Figure 2. Description of valproate exposure

Abbreviation: VPA = valproate

9.3.2. Outcomes

9.3.2.1. Main outcome

Given the complexity of the treatment patterns that can be observed in real-life settings, a description of the overall treatment patterns after valproate discontinuation at a macro-level is necessary to have an overview of the situation. To this end, different treatment sequences will be defined in comparison with the latest "on treatment" sequence (when available) (see Figure 3). These sequences are the following:

- Monotherapy: only 1 drug of interest for epilepsy/bipolar disorder (excluding valproate);
- Double therapy: 2 distinct drugs of interest for epilepsy/bipolar disorder (excluding valproate);
- Combination therapy: ≥ 3 distinct drugs of interest for epilepsy/bipolar disorder (excluding valproate);
- Different monotherapy: 1 medication that differs from the previous sequence treatment;
- Different double therapy: ≥ 1 of the 2 medications that differs from the previous sequence treatment;
- Different combination therapy: ≥ 1 of the medications of the combination therapy that differs from the previous sequence treatment;
- Valproate reintroduction;
- Monotherapy + Valproate reintroduction;
- Double therapy + Valproate reintroduction;
- Combination therapy + Valproate reintroduction;
- Different monotherapy + Valproate reintroduction;
- Different double therapy + Valproate reintroduction;
- Different combination therapy + Valproate reintroduction;
- Not exposed to any epilepsy/bipolar disorder medication (including valproate).

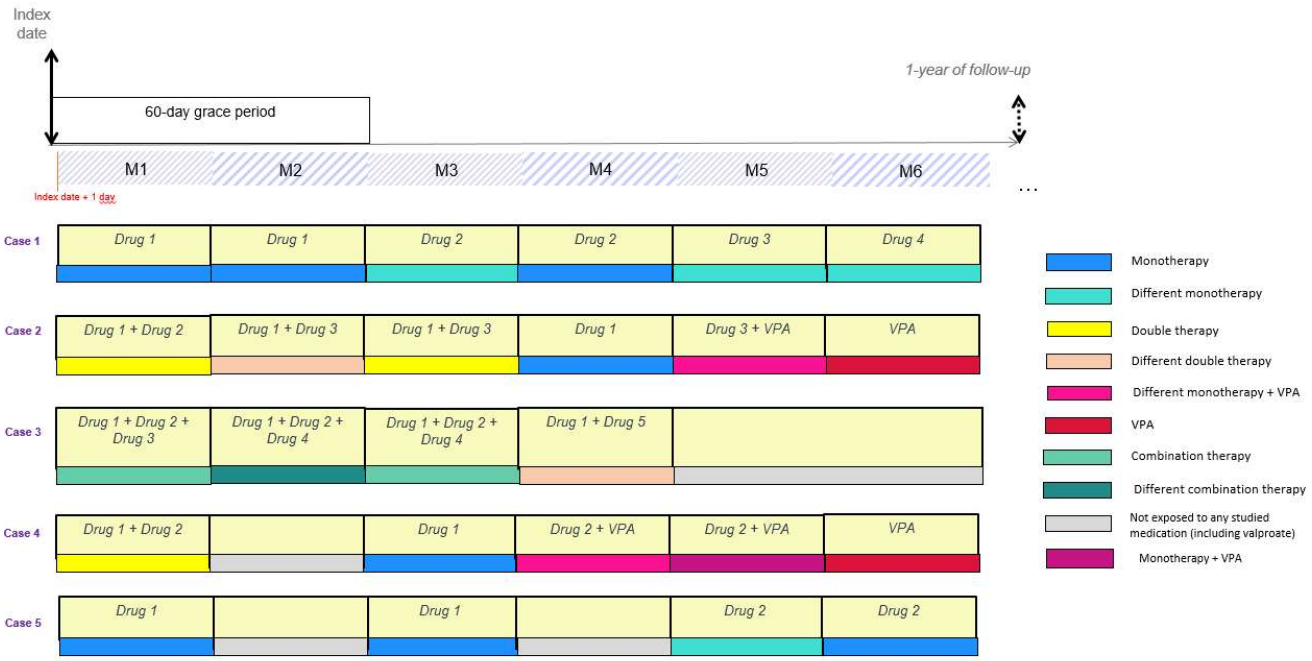

Figure 3. Example of treatment sequences definition in the 1-year follow-up

By defining these treatment sequences, we ensured that all situations of therapeutical management that can be observed in clinical practice will be represented, while having a reasonable number of treatment sequences (n=14) to perform the unsupervised clustering method. These **treatment sequences** will be identified at regular intervals, based on the frequency of prescription/dispensing of the drugs of interest (i.e., 1 month in SNDS), from the day after index date to the end of the follow-up (i.e., end of the 1-year follow-up period, death or database eligibility lost, whichever comes first).

This overview of the overall treatment patterns will be represented graphically, using a sequence index plot, to illustrate the succession of treatment sequences over time for each patient.

This step will be then followed by the determination of clusters defined as groups of patients sharing homogeneous treatment patterns of the different predefined sequences according to time periods. A state sequence analysis (SSA), based on unsupervised machine learning methods (clustering analysis)(5) will be used as the type and number of homogenous treatment sequence clusters are not known *a priori*. Indeed, in this case, unsupervised learning methods are more suitable than supervised learning techniques that fit in the fields of predictive methods.

The unsupervised clustering methods that will be used are detailed in §9.7.

All the identified clusters and their patient's characteristics will be reviewed by the Scientific Committee comprising two independent clinical experts, each of them being specialist in each studied disease (epilepsy and bipolar disorders), to determine which ones are the most relevant according to their experience in clinical practice. Therefore, they could identify the cluster(s) that is (are) the most likely to reflect a success in epilepsy/bipolar disorder management after valproate discontinuation based on the combination of a set of indicators including valproate non reintroduction, contextualized with others factors. Indeed, the success or failure after valproate discontinuation will be contextualized according to several and pharmaceutical parameters such as the number of hospitalizations, polypharmacy. This list of parameters may be supplemented by the Scientific Committee experts with regard to the results and their interpretation.

Medications used at each treatment sequence will be further detailed in each cluster and represented using appropriate graphics (Sankey diagram, sunburst, etc.).

9.3.2.2. Secondary outcomes

Secondary outcomes will be assessed during the 1-year follow-up period for either epilepsy or bipolar disorder, in the overall cohort and by clusters, depending on data availability and completeness in each database, to complement the main outcome:

- First occurrence of **valproate reintroduction.**
- **Occurrence of clinical relapse**. Relapse will be identified by at least one hospitalization or ER visit for epilepsy in the epilepsy cohort, and at least one hospitalization or ER visit for either Schizoaffective disorders, Bipolar affective disorder, Manic episode, or suicide attempt in the bipolar disorder cohort.
    - (CPRD only) the following variables related to specialist outpatient care delivered at hospitals will be considered to build an algorithm for proxy of clinical relapse in patients without linkage to inpatient hospital data: type of outpatient consultation appointment dates, main specialty and treatment specialty under which the patient was treated, referral source, waiting times, clinical diagnosis and procedures performed)
- Occurrence of pregnancy.
- Hospitalization and discharge diagnoses (including epilepsy-related diagnosis, Psychiatric-related diagnosis, fall, fracture, trauma).
- ER visits and diagnoses, unscheduled care.
- Number of office visits to GP, Neurologist or Psychiatrist or other medical specialty (gynecologist) or other relevant health care professionals (epilepsy nurse in the UK).
- Death.
- Sick-leave days.

The pregnancy outcomes (maternal and new-born) are of high importance for the clinician but will not be considered for the objectives of this study.

9.3.3. Other variables

The following characteristics will be considered according to availability and completeness in each database:

- Patient characteristics:
    - Age at index date.
    - Deprivation index.
    - Region.
    - Psychotherapy / psychosocial support or education.
- Disease characteristics:
    - Age at diagnosis.
    - Duration of disease (date since diagnosis).
    - Number and cumulative length of stay of hospitalizations related to each indication.
    - Time elapsed between last clinical event (hospital admission/discharge specific to each indication) and index date.
    - Psychiatric diagnosis associated with hospital admission or discharge before index date: depressive disorder, Schizophrenia, Phobic anxiety disorders, Other anxiety

disorders, Obsessive-compulsive disorder, Reaction to severe stress, and adjustment disorders, Eating disorders, Nonorganic sleep disorders, Mental and behavioral disorders due to use of alcohol or Care involving use of rehabilitation procedures or Alcohol rehabilitation, Mental and behavioral disorders due to use of opioids, or other psychoactive substance; personality disorder.

  - Epilepsy-related diagnosis associated with hospital admission or discharge before index date: Localization-related (focal)(partial) idiopathic epilepsy and epileptic syndromes with seizures of localized onset, Localization-related (focal)(partial) symptomatic epilepsy and epileptic syndromes with simple partial seizures, Localization-related (focal)(partial) symptomatic epilepsy and epileptic syndromes with complex partial seizures, Generalized idiopathic epilepsy and epileptic syndromes, Other generalized epilepsy and epileptic syndromes, Special epileptic syndromes, Grand mal seizures, unspecified (with or without petit mal), Petit mal, unspecified, without grand mal seizures, other epilepsy; EEG, MRI procedure.
  - History of anti-epileptic or antidepressants, mood stabilizers or neuroleptic treatment.
  - History of suicide attempt/ self-injury.
- Obstetric and gynecological characteristics:
  - Record of contraceptive method.
  - Record of folate prescription.
  - History of previous pregnancies (leading to live births or discontinued (miscarriages or abortions)).
  - History of fertility treatment/visit to fertility clinic.
  - Access to perinatal mental health unit (for subgroup of pregnant women only).
- Other comorbidities: diabetes, metabolic disorders, overweight/obesity.

It should be noted that the aetiology, sub-types of epilepsy or clinical events (such as seizures, characteristics or phases of bipolar disorder) might not be captured with sufficient completeness in healthcare databases.

9.4. Data sources

The study will be based on secondary data collection from a United Kingdom electronic medical records database, the CPRD (Clinical Practice Research Datalink) or the French nationwide claims, the SNDS (*Système National des Données de Santé*) (table 1).

9.4.1. CPRD

The CPRD is a primary care database containing anonymized patient records for about 6% of the UK population. Its strengths as a research tool include its size, representativeness of patient and practice characteristics, and a virtually complete medical history of patients due to the recording of referral to secondary care. More than 700 participating general practices are required to record (i) each episode of illness, or new occurrence of a symptom, and (ii) all significant morbidity events, e.g., all significant clinical contacts, all significant diagnoses and abnormal test results, all referrals to outpatient clinics and hospital admissions practices (6). The medical conditions of interest as well as the drug classes have already been described in previous publications (7–9). Data are retrieved by means of the READ classification system.

The CPRD has developed the Pregnancy Register that includes identification and details of pregnancy episodes. Because the precise time period during which a woman is pregnant can be difficult to ascertain in the data, a pregnancy identification algorithm has been developed to identify and maximise the use of records relating to the timing and duration of pregnancy, the type of pregnancy outcome (live birth, stillbirth or pregnancy loss), and additional features pertaining to the pregnancy.

Hospitalization events may be better recorded in secondary versus primary care, which requires record linkage with other health-related patients' data sets, namely Hospital Episodes Statistics (HES) Admission Patient Care (APC). Also, the linkage to HES- Outpatient Care (OP) may inform about specialists' outpatient care delivered by some hospitals in England. In 2018, about 8.0 out of 10.6 million patients (76%) were eligible for CPRD standard linked datasets (10).

9.4.2. SNDS

The SNDS database is the nationwide healthcare insurance system database with individual anonymous information on all reimbursed outpatient claims linked to the national hospital-discharge summaries database system (PMSI) and the national death registry, using a unique national pseudonymised identifier. It currently includes 98.8% of the French population, more than 66 million persons from birth (or immigration) to death (or emigration), even if a subject changes occupation or retires.

Of the 66 million inhabitants in France at the end of 2015, the general scheme covers salaried employees of the private sector and their dependents (i.e., about 76% of the population living in France), as well as people covered by "*Sections Locales Mutualistes*" (SLM), essentially civil servants, employees of territorial collectivities and public hospitals and students, i.e., about 11% of the population. The two other main healthcare schemes ("*Régime Social des Indépendants"* [RSI] - for craftsmen, shopkeepers, liberal professions and their dependents - and "*Mutualité Sociale Agricole*" [MSA] - for farmers and agricultural workers) cover 11% of the population.

The SNDS contains individual pseudonymised information on (11,12):

- General characteristics: gender, year of birth, affiliation scheme, area of residence; deprivation status (*Couverture Maladie Universelle complémentaire*, CMU-c).
- Date of death for those concerned and very soon cause of death.
- Long-term disease (LTD, or ALD in French, and associated ICD-10 codes) with starting and ending date. LTD mainly concerned costly chronic diseases. LTD registration is obtained at the request of a patient's practitioner and validated by the health insurance system physician. Once registered, patients receive full (i.e., 100%) reimbursement for expenditure related to the LTD. The LTD information is specific for the diagnosis (very low risk of false positives), but not sensitive because not all patients with the disease ask to benefit from a LTD.
- Outpatient reimbursed healthcare expenditures: visits, medical procedures, nursing acts, physiotherapy, medical imageries, lab tests, drugs, medical devices, transports, sick leaves… with prescriber and professional caregiver information (medical or paramedical specialty, private/public practice), dates (prescription and dispensing), and codes (but not the medical indication nor result).
- Hospital-discharge summaries from the PMSI: ICD-10 diagnosis codes (primary, linked and associated diagnosis) for all private and public medical, obstetric and surgery hospitalizations, psychiatric hospitalizations, at-home hospitalizations and rehabilitation centre hospitalizations, with the date and duration of hospitalization, medical procedures, and cost coding system, as well as most of very costly drugs. The hospital discharge summary includes the medical unit summaries when the patient is hospitalized successively in several medical units. Primary diagnosis is the health problem that motivated the admission in the hospital. It is determined at hospital discharge. For patients hospitalized successively in several medical units, the primary diagnosis of the hospitalization, as well as all medical unit primary diagnoses, are generally

taken into account to define the occurrence of an outcome in a pharmacoepidemiology study. A linked diagnosis can exist only if the primary diagnosis is a care procedure with a code Z of the ICD-10 classification (e.g., chemotherapy session) for a chronic or LTD disease. It indicates the pathology at the origin of the care procedure. As primary diagnosis is taken into account to define the occurrence of an outcome in a pharmacoepidemiology study. Associated diagnoses are specified if they represent specific healthcare resources. They are mainly underlying chronic diseases. Associated diagnoses can be used to define chronic diseases but are generally not taken into account to define the occurrence of an outcome in a pharmacoepidemiology study (many being false positives for the studied outcome).

The SNDS contains data on pregnancy outcomes using discharge diagnoses and medical procedures, but the pregnancy start dates are lacking. To address this issue, a pregnancy identification algorithm has been developed in the SNDS to capture all types of pregnancy outcome (live birth, stillbirth, elective abortion, therapeutic abortion, spontaneous abortion, and ectopic pregnancy) and estimate the date of pregnancy start (13).

Non-hospital data are updated every month with a lag of at least 6 months to have 98% of information uploaded and hospital-discharge summaries yearly at end of Q3 for the previous year. Access to SNDS is regulated and needs approval from Institute of Health Data (*Institut National des Données de Santé* - INDS) and French data protection commission (*Commission Nationale de l'Informatique et des Libertés* - CNIL).

Table 1: Overview of databases to be used for the study

	SNDS	CPRD
Country	France	UK
Type of database	Claims	EMR
Number of inhabitants	66.6 million	66.8 million
Number of patients in databases	66 million	5,7 million
HES (APC or OP)	100% (PMSI)	≈ 55%
Database updates	Yearly (Q3)	Yearly (Q2)
Patients' characteristics		
Date of birth	Yes	Yes
Region	Department	County
Pregnancy	Yes (validated algorithm)	Yes (pregnancy identification algorithm within pregnancy register)
Exposure		
Prescription/dispensing	Dispensing	Prescription
Coding of drugs	ATC	BNF
Dosing regimen	Yes	Mostly available
Outcomes		
Hospitalizations	Yes (ICD-10 codes)	Yes (HES-APC) (ICD-10 codes)
Out-patient diagnoses (primary care)	No	Yes (READ codes)
Long-Term Disease diagnoses	Yes (ICD-10 codes)	No
Outpatient diagnoses (specialist care)		YES (HES-OP) (ICD-10 codes)

EMR = Electronic medical records; PMSI = *Programme de Médicalisation des Systèmes d'Information*; HES = Hospital Episodes Statistics; ATC = Anatomical therapeutic chemical; BNF = British National Formulary; ICD = International Classification of Diseases; APC = Admitted Patient Care; OP = Outpatient care

9.5. Study size

The primary objective of the study is to describe treatment patterns, to identify those which are similar to define clusters of patients, after valproate discontinuation in valproate WCBP chronic

users and in a subpopulation of pregnant women, and then to determine which ones are more likely to reflect a success of the therapeutic management.

Analyses will be performed separately in each database and the different datasets will not be pooled. All patients who fulfil the inclusion and exclusion criteria will be included in the analysis.

The calculation of the sample size is thus made for each database and for each indication cohort, and relies on the precision of the 95% confidence interval of the proportion of interest.

In UK, the prevalent use of valproate was approximately estimated at 25 to 30 per 10,000 women aged 18-45 years and 10 to 15 per 10,000 women aged 12-17 years over 2010-2016 (MHRA data, (14)). In France, the prevalent use of valproate decreased in WCBP from 122,382 in 2007 to 83,712 by 2015 (15). The expected number of valproate WCBP chronic users who have discontinued valproate during the inclusion period (from January 1, 2014 to December 31, 2017) will be a result of the study analysis.

The sample size of each cluster will be the output of the clustering analysis, and the precision of each criterion included in definition of success will be also estimated according to the size of clusters using the following formula, in which n is the sample size of a cluster, t is the t-test value for a given confidence interval (t=1.96 for a confidence interval of 95%), p is the proportion of patients with the considered criterion, and e is the error margin: $e = t.\sqrt{\frac{p.(1-p)}{n}}$. The Table 2 illustrates the precision obtained for a given cluster size and an observed percentage.

Table 2: Precision obtained for a given cluster size and an observed percentage of considered criterion

Cluster size	Observed percentage										
	1%	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%
50	2.8%	6.0%	8.3%	9.9%	11.1%	12.0%	12.7%	13.2%	13.6%	13.8%	13.9%
100	2.0%	4.3%	5.9%	7.0%	7.8%	8.5%	9.0%	9.3%	9.6%	9.8%	9.8%
200	1.4%	3.0%	4.2%	4.9%	5.5%	6.0%	6.4%	6.6%	6.8%	6.9%	6.9%
300	1.1%	2.5%	3.4%	4.0%	4.5%	4.9%	5.2%	5.4%	5.5%	5.6%	5.7%
400	1.0%	2.1%	2.9%	3.5%	3.9%	4.2%	4.5%	4.7%	4.8%	4.9%	4.9%
500	0.9%	1.9%	2.6%	3.1%	3.5%	3.8%	4.0%	4.2%	4.3%	4.4%	4.4%
600	0.8%	1.7%	2.4%	2.9%	3.2%	3.5%	3.7%	3.8%	3.9%	4.0%	4.0%
700	0.7%	1.6%	2.2%	2.6%	3.0%	3.2%	3.4%	3.5%	3.6%	3.7%	3.7%
800	0.7%	1.5%	2.1%	2.5%	2.8%	3.0%	3.2%	3.3%	3.4%	3.4%	3.5%
900	0.7%	1.4%	2.0%	2.3%	2.6%	2.8%	3.0%	3.1%	3.2%	3.3%	3.3%
1000	0.6%	1.4%	1.9%	2.2%	2.5%	2.7%	2.8%	3.0%	3.0%	3.1%	3.1%
1500	0.5%	1.1%	1.5%	1.8%	2.0%	2.2%	2.3%	2.4%	2.5%	2.5%	2.5%
2000	0.4%	1.0%	1.3%	1.6%	1.8%	1.9%	2.0%	2.1%	2.1%	2.2%	2.2%
2500	0.4%	0.9%	1.2%	1.4%	1.6%	1.7%	1.8%	1.9%	1.9%	2.0%	2.0%
3000	0.4%	0.8%	1.1%	1.3%	1.4%	1.5%	1.6%	1.7%	1.8%	1.8%	1.8%
3500	0.3%	0.7%	1.0%	1.2%	1.3%	1.4%	1.5%	1.6%	1.6%	1.6%	1.7%
4000	0.3%	0.7%	0.9%	1.1%	1.2%	1.3%	1.4%	1.5%	1.5%	1.5%	1.5%
4500	0.3%	0.6%	0.9%	1.0%	1.2%	1.3%	1.3%	1.4%	1.4%	1.5%	1.5%
5000	0.3%	0.6%	0.8%	1.0%	1.1%	1.2%	1.3%	1.3%	1.4%	1.4%	1.4%
5500	0.3%	0.6%	0.8%	0.9%	1.1%	1.1%	1.2%	1.3%	1.3%	1.3%	1.3%
6000	0.3%	0.6%	0.8%	0.9%	1.0%	1.1%	1.2%	1.2%	1.2%	1.3%	1.3%
6500	0.2%	0.5%	0.7%	0.9%	1.0%	1.1%	1.1%	1.2%	1.2%	1.2%	1.2%
7000	0.2%	0.5%	0.7%	0.8%	0.9%	1.0%	1.1%	1.1%	1.1%	1.2%	1.2%
7500	0.2%	0.5%	0.7%	0.8%	0.9%	1.0%	1.0%	1.1%	1.1%	1.1%	1.1%
8000	0.2%	0.5%	0.7%	0.8%	0.9%	0.9%	1.0%	1.0%	1.1%	1.1%	1.1%
8500	0.2%	0.5%	0.6%	0.8%	0.9%	0.9%	1.0%	1.0%	1.0%	1.1%	1.1%
9000	0.2%	0.5%	0.6%	0.7%	0.8%	0.9%	0.9%	1.0%	1.0%	1.0%	1.0%
9500	0.2%	0.4%	0.6%	0.7%	0.8%	0.9%	0.9%	1.0%	1.0%	1.0%	1.0%
10000	0.2%	0.4%	0.6%	0.7%	0.8%	0.8%	0.9%	0.9%	1.0%	1.0%	1.0%

9.6. Data management

This study will follow the relevant chapters of the ENCePP and the International Conference on Harmonisation (ICH) guidelines for data management.

9.6.1. Database-specific data management

The Clinical Practice Research Data Link Gold plus Hospital Episodes Statistics (HES) Data (CPRD) – UK

The processes for database management will be detailed in the specific protocol or the SAP. Generally, the data are stored at the database level and analysed locally. Statistical Analysis System (SAS®) Software will be utilized for access to the raw data, to manage the analytic datasets and to conduct data analysis.

The *Système National des Données de Santé* (SNDS) – France

Database extraction criteria will be described in a Data Extraction Plan (DEP) approved prior to initiating extraction. Data extraction will be done by the CNAM (*Caisse Nationale d'Assurance Maladie*).

Data transformation, including decision rules, diseases definitions, exposure definitions, outcomes, risk factors, healthcare resources and calculated variables will be detailed in a statistical analysis plan (SAP).

Raw data and transformed data will be stored on a network meeting security standard as required by the French law.

9.7. Data analysis

9.7.1. Generalities

Statistical analysis will be performed using SAS® (SAS Institute, latest current version, North Carolina, USA) and R software (R Foundation, latest current version). A SAP will be developed and will be validated before the interim and final analyses.

Datasets will not be pooled across countries and analyses will be performed separately for each database. The following analyses will be performed for the epilepsy and bipolar disorder cohorts and the corresponding subpopulation of pregnant women.

An interim report will present the results related to the description (and graphic representations) of the treatment patterns as well as the identification and characterization of the different clusters (similar treatment patterns) identified in each database (if available) during follow-up. The final report will present the overall results including those related to the identification of covariates associated with the clusters, and those associated with the studied outcomes.

Qualitative and ordinal variables will be summarized by frequencies and proportions of each modality, taking into account missingness as a modality (concerns only CPRD database). Continuous variables will be summarized by size, number of patients with missing data, arithmetic mean, standard deviation, median, interquartile ranges and extreme values. 95% Confidence intervals (CI) will be estimated using Normal approximation for quantitative and qualitative relevant parameters.

9.7.2. Population description

The selection of study populations for data analysis will be presented in a flowchart. The descriptive analyses of the study population will include:

- A description of patients at index date: demographic characteristics.
- A description of patients during the pre-inclusion period: medical history, medication, and healthcare consumption (refer to variables listed in section 9.3.3).

- (CPRD only) A stratification of descriptive analysis by mutually exclusive sub-populations with the following availability of linkage:
    - with HES-APC linkage
    - with HES-OP linkage and without HES-APC linkage
    - without HES-APC or HES-OP linkage.

9.7.3. Main outcome

The overall succession of the treatment sequences observed for each patient during the follow-up will be broadly described using a sequence index plot, where a simple ordering strategy will be applied to prioritize patients with similar treatment patterns according to time period. State sequence analysis (SSA), based on unsupervised machine learning methods (clustering analysis)(5) will be used to determine clusters of patients with similar treatment sequences, as the type and number of homogenous treatment sequence clusters are not known *a priori*.
SSA allows the comparison of sequences among the different subjects and the identification of common patterns.(16,17) In this case, once individual treatment sequences are created, i) the distance and dissimilarity between each pair of sequences is estimated, and ii) a clustering method is applied on the dissimilarity matrix containing all the pairwise distances among the sequences.

(i) The Optimal Matching (OM) method (18,19) based on Levenshtein distance (20) will be used, this is one of the most used methods to estimate a distance between 2 sequences.

(ii) There are a multitude of unsupervised learning algorithms for clustering (partitioning-based, hierarchical, density-based, grid-based, model-based,…) (21–23) such as Partitioning Around Medoids (PAM), Agglomerative Hierarchical Clustering (AHC) or K-means for the best known. The PAM algorithm will be preferentially used in first instance as it is easy to use with dedicated R package and works better on small-medium sized datasets (7,000-10,000 subjects per cohort in SNDS data for this study, 200-500 subjects for CPRD). In order to determine the quality of the partitioning and the most efficient number of clusters, internal validity indices (*ie.* silhouette metric, sum of square, R-squared, Hubert-Somers D, Hubert's C…) will be estimated. These indices are measures based on the goodness-of-fit between each clustering and the data, they evaluate clustering results by using only features and information inherent in a dataset.
The AHC method will be performed as a sensitivity analysis as it is also adapted for small-medium sized databases and is frequently used in SSA analyses. The optimal clusters obtained with AHC will be compared with those obtained from the PAM algorithm. In case of major discrepancies between both methods, the opinion of clinical experts of Scientific Committee will be asked.

All these methods will be detailed in the SAP.

Patients' characteristics will be described according to the cluster previously identified. This description will be reviewed by clinical experts of the Scientific Committee who will assess the relevance of selected clusters according to what they know from real-life healthcare in the considered domains. Small-sized and non-relevant clusters may be removed from the secondary analyses (associated factors).

A specific description of treatment patterns in each cluster will be performed including the details of the medication(s) dispensed/prescribed and according to the type of associated treatments categorized in mono-, double- and combination therapy). Two levels of treatment pattern will be used based on the number of subjects observed in each pre-defined treatment pattern during the follow-up period:

- Low-level patterns will concern all treatment patterns with low frequency (<5% or <30 patients),

- High-level patterns will concern all treatment patterns with frequency ≥5% or ≥30 patients).

Appropriate graphics will be used to illustrate the detail of the medications dispensed/prescribed after valproate discontinuation in these clusters (e.g., Sankey diagram, sunburst).

9.7.4. Secondary outcomes

Each following outcomes will be described during the follow-up period in the overall cohort and by cluster:

- Clinical relapse,
- Occurrence of pregnancy in the subgroup of pregnant women.

This description will include the number of occurrences of each outcome per patient and the proportion of patients with at least one occurrence of each considered outcome. Appropriate graphic will be also used to visualize the counts of events (frequency and cumulative frequency distributions of clinical relapse and pregnancy) and the month of occurrence during the follow-up period.

Other secondary outcomes will be described during follow-up:

- Hospitalization and discharge diagnoses during the pre-index period and the follow-up period (including epilepsy-related diagnosis, psychiatric-related diagnosis, fall, fracture, trauma).
- ER visits and diagnoses, unscheduled care.
- Number of office visits to GP, neurologist or psychiatrist or other medical specialty (gynecologist) or other relevant health care professionals (epilepsy nurse in the UK).
- Death, overall and by age group; If the number of observations is sufficient, the comorbidities by age group will be described.
- Sick-leave days (duration).

(CPRD) The analysis that include variables related to hospitalization will only be conducted in patients with linkage to HES-APC. The analysis in patients with linkage to HES-OP will be proposed separately. In patients with linkage to both HES-APC and HES-OP, only data from HES-APC will be used.
For patients without any linkage, all the descriptive analysis proposed will be performed except if the outcome variable is related to hospitalization or clinical relapse.

Table 3: Summary of analysis for the outcome Clinical Relapse by availability of linkage in the CPRD

	Proxy of clinical relapse (for models)		Primary care codes related to relapse (only descriptive)
Analysis population:	Hospital admission/ discharge specific to each indication	Consultation with specialists (specific to each indication) and related codes	
Sub-population with HES-APC linkage	X		X
Sub-population without HES-APC, with HES-OP		X	X
Sub-population without HES-APC, without HES-OP			X

9.7.5. Identification of covariates

The identification of covariates associated with the most relevant clusters will be performed using a multivariable multinomial logistic regression model, with dependent variable having several possible categories (cluster 1, cluster 2, … cluster n);

Factors associated with the outcome of interest will be identified among a list of potential covariates, which may be related to the prior treatments, or patient's characteristics. The candidate covariates related to treatment patterns will include (non-exhaustive list): MPR of valproate > 60% (as a continuous variable or as categories (60-80% / > 80%), presence of valproate dose-tapering phase (and length if applicable). The candidate covariates related to patients' or disease characteristics will include (non-exhaustive list): age at index date, hospitalizations, long-term disease, medical visit, lab test, medical procedure or drugs related to the disease or to other comorbidities.

Modelling analyses are exploratory conditioned by the size of the clusters and the number of candidate covariates, with at least 10-15 patients expected per covariate modality). Continuous variables with nonlinear effect on the outcome will be categorized according to relevant cut-off or according to their distribution. Modalities of variables with few patients can be grouped with other modalities (if possible).

Clinical experts will identify the most relevant variables that are likely to be related to the outcome (i.e. type of cluster). Some of them can be forced in the model.

Selection of variables in the model will use the "Augmented backward elimination (ABE)" procedure (24,25) which combines the standardized change-in-estimate criterion with significance-based Backward elimination. This method which extends the ideas of 'purposeful variable selection' (*Hosmer, Lemeshow and May*) can address the study objectives to find a parsimonious and valid model that describes the dependency of the outcome on the explanatory variables and increases the stability of the selected model. First a global (full) model will be built including all experts' predefined variables. Non-significant variables (starting by variables with the largest p-value, ie. >0.25 or >0.10) will be eliminated if, after their elimination, the change in estimate of the other variables in the model is less than a specified threshold (commonly <20%), otherwise they will be retained in the model. The final model will be selected after testing all the non-significant variables. The final model goodness-of-fit will be assessed (C statistic, generalized H-L test). Final model stability will be assessed by bootstrap inclusion frequencies and sampling distributions of regression coefficients.

For sensitivity analysis, LASSO method may be used. The shrinkage tuning parameter will be selected through 10 fold-cross validation minimizing the deviance of the model.

Associations will be expressed as odd ratios (OR) for multinomial logistic regression model comparing the clusters 2 by 2 with 95% CI and Wald test p-value. The quality of each model will be estimated using the Akaike criterion.

9.7.6. Missing data

The imputation of missing data will be carried out for the daily dose to estimate the number of days of supplies when information in the prescriptions are not complete, as detailed below:

Method 1: calculation of the days of supply using total quantity supplied divided by daily assumption which is constructed from daily dose.
Method 2: If daily dose is not available in CPRD data, then use Defined Daily Dose (DDD) for VPA. The DDD for VPA is 1500 mg/day.
Method 3: if method 1 and method 2 both are not applicable, we will impute the mode value. The mode value of days of supply from the same patient with same product will be imputed; if that is not available, then the mode value of the same patient regardless products will be imputed but still VPA drug.

No other imputation of missing data will be carried out in these analyses.

Qualitative and ordinal variables will be summarized by frequencies and proportions of each modality, considering missingness as a modality (concerns only CPRD database). Continuous variables will be summarized by size, number of patients with missing data, arithmetic mean, standard deviation, median, interquartile ranges and extreme values. 95% Confidence intervals (CI) will be estimated using Normal approximation for quantitative and qualitative relevant parameters.

SNDS database records all reimbursed claims and hospitalisations without missing values. No imputation of missing data will be carried out in these analyses.

9.8. Quality control

The Clinical Practice Research Data Link Gold plus Hospital Episodes Statistics (HES) Data (CPRD) – UK

CPRD data process/management for this study will be conducted by Cindy Li at Sanofi who has extensive programming and analysis experience with different healthcare databases including CPRD. All programs and result tables will be thoroughly and independently reviewed by another programmer/analyst at each step of the analysis to validate the coding rules, programming and statistical analyses. The guidelines and standard procedures detailed in Sanofi Quality Documents will be followed to ensure the quality of data management and analysis and the storage of analysis programs, results, and other key documents such as the study protocol and reports.

All information is entered by practice staff and is anonymized prior to central collection. The CPRD carries out a series of ongoing checks to ensure that the data are ‘up to standard’; this comprises assessment of patient data (age, gender, registration details and event dates) and the completeness, continuity and plausibility of electronic data recording in key areas at the practice level. Prescription data are well documented since the GP uses the computer to generate prescriptions and these are automatically recorded in the database. The therapy file is therefore virtually complete, except for prescriptions issued in secondary care and for drugs that are purchased over the counter.

The *Système National des Données de Santé* (SNDS) – France

The BPE platform, INSERM CIC1401, has implemented a quality management system for all its activities. CNAM data extraction will be validated using the expected population size estimated using the SNDS. An independent double programming will be performed for main criteria and analyses, and the results compared for validation. All statistical logs are kept and can be provided. In the case of interim analyses, the database for the interim analysis is locked and kept for ulterior validation if needed. The statistical analysis report (SAR) is included in the final study report.

9.9. Limitations of the research methods

This study will use real-life prescription data of valproate in 2 European countries issued from longitudinal electronic medical records in the UK and administrative health care claims in France. Health-record databases can be appropriate tools for drug-use studies due to the recording of prescription data independently of any study purposes and the large panels of prescribers and associated population coverage. Most databases include patient characteristics such as age, pregnancy, major diagnoses, and prescription records and would thus be appropriate to evaluate the patterns of Valproate switch. However, the results may not be generalized to other European countries, due to differences in health care delivery system or clinical practice. Also, only dispensation or prescription data will be available, assumed to reflect the actual patient's intake.

The following limitations that are inherent to each data source may be anticipated:

With CPRD

The CPRD includes prescriptions at primary care so the analysis resulting from this database will be limited to this specific setting and therefore probably not cover the full spectrum of exposure.

Information on pregnancy start and end date / discontinuation is not comprehensively covered. Since April 2017, CPRD has created a new monthly updated pregnancy database containing pre-calculated pregnancy start and end dates, which should be used in addition to the data already used for the analyses to answer the additional questions regarding pregnancies. However, this pregnancy database only contains data on pregnancies with an outcome of live birth.

Use of non-prescribed, non-reimbursed contraceptive methods are not captured.

Information about hospitalization is available for a sample of GP and patients. Therefore, the information needed to define clinical events may be incomplete. However, the selection bias should not be differential, meaning that patients with linkage to hospital data should not have any systematic differences compared to those who have the linkage established. The outcomes related to clinical stability will only be evaluated in a subpopulation and there is a risk of lack of statistical power. Also, the proxy of clinical relapse based on specialist outpatient care data proposed for patients without linkage to inpatient/hospitalization data may indicate clinical relapse or a need for dose adjustment or tolerability issues and therefore is considered as a less robust proxy.

In CPRD, a patient who might be exclusively followed-up in specialist care would not be identified in the study population. Also, the indication for valproate use may be missing in the patient's history. Thus, for both reasons, there is a possibility that the study population from CPRD is not full representative of the overall population of female patients with epilepsy or bipolar disorder in the UK.

With SNDS

The SNDS is a national healthcare claims database linked to the national hospital discharge summaries database that covers about 99% of the French population. It provides a unique opportunity to identify all subject WCBP chronic users of valproate for either epilepsy or bipolar disorder, who have discontinued valproate during the inclusion period, with exhaustive information

about reimbursed treatments out of hospital and use of reimbursed healthcare resources, as well as all hospitalizations. Furthermore, the SNDS has the advantage of any study that use patient records from an existing database that are not impacted by the study, as most of field studies.

This is also the main limit of this claims and hospitalization database that was built for administrative and reimbursement purposes with little clinical data and no biological results, including severity or stage of the disease, or some risk factors such as smoking status, body mass index, blood pressure.

Since all subjects identified will be extracted from a national database, there is no study selection bias, nor attrition bias, except very rare withdrawals for emigrant people.

Women with any of the medical condition of interest (epilepsy or bipolar disorder) who did not have any hospitalization recorded during the 5-year pre-index period would not be included in the study population. Thus, the study sample will not be full representative of the target population.

With CPRD and SNDS

The indication for which valproate is initiated will be determined through records of diagnoses related to drugs prescriptions in the medical history up to 5 years before the index date. This variable can be inaccurate or missing for some patients. This may lead to a misclassification of patients.

Use of non-prescribed, non-reimbursed contraceptive methods are not captured.

Given its descriptive nature, the unsupervised machine learning methods do not require a minimum number of patients to be performed. However, the size of the dataset impacts the clustering process and quality. The PAM algorithm applied in the present analyses is optimized for small-medium sized datasets as it is the case in this study (7,000-10,000 subjects per cohort in SNDS data for this study, 200-500 subjects for CPRD).

9.10. Other aspects

None

10. PROTECTION OF HUMAN SUBJECTS

This study is non-interventional, and analyses are based on pre-existing data.

This study will be conducted in accordance with the guidelines for Good Pharmacoepidemiology Practice (Revision 3 June 2015) (ISPE, 2007), the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology (Revision 9).

In each participating country, all necessary regulatory/ethics submissions are performed in accordance with local regulations including local data protection regulations.

In the **UK**, the protocol must be submitted to the Independent Scientific Advisory Committee (ISAC) using the CPRD ISAC Application form. ISAC approval is required if access to anonymised patient level data is being requested for research purposes. Patient level data used in the generation of aggregated data, intended for publication, also requires approval.

In **France**, as this project is a database analysis with individual anonymous information, subject informed consent is not required. Data extraction from the SNDS is regulated and needs approval from National Institute of Health Data (INDS) and French data protection commission (*Commission Nationale de l'Informatique et des Libertés* - CNIL).

11. MANAGEMENT AND REPORTING OF ADVERSE EVENTS / ADVERSE REACTIONS

This project is a database analysis using anonymous individual information without any spontaneous reporting. Study outcomes will be reported in aggregate in the final study report, and no individual or expedited reporting is required, according to the EMA Guideline on good pharmacovigilance practices (26), as well as the ENCePP Guide on Methodological Standards in Pharmacoepidemiology (27).

The descriptive statistics will include analysis among female patients who discontinued valproate and the proportion of those with some outcomes such as disease relapse or death. Since this study aims at understanding how patterns of discontinuation are associated with various outcomes including those that might meet the definition of Adverse Event /Adverse Drug Reaction, it is proposed not to summarize findings in the safety-specific section.

12. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

An interim report will be submitted 12 months after the PRAC endorsement of the protocol. Since this study is exclusively based on pre-existing data that will be extracted once from the data sources described in section 9.4, it is estimated that two interim reports are needed. The final report will be submitted 6 months after the second interim report.

In accordance with the 2010 EU pharmacovigilance legislation, the protocol of this study will be entered into the publicly available EU PAS register. A completed ENCePP Checklist for study protocols is attached in appendix 2. Updates to the study protocol in case of substantial amendments and the final study report will also be entered in the register.

Study findings will be published in a peer reviewed journal. Any publication will be guided by the Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication of the International Committee of Medical Journal Editors (ICMJE (28)), updated May 2022.

All reporting will be consistent with the STROBE (STrengthening the Reporting of OBservational studies in Epidemiology) Initiative checklist for cohort studies (29).

Still in line with the EMA guideline, and in order to allow competent authorities to review in advance the results and interpretations to be published, the MAHs should communicate to the Agency and the competent authorities of the Member States in which the product is authorised the final manuscript of the article within two weeks after first acceptance for publication.

Any publication has to be disclosed onto the ENCePP site within 2 weeks of acceptation by a Journal.

13. REFERENCES

1. Tomson T, Battino D, Bonizzoni E, Craig J, Lindhout D, Perucca E, et al. Withdrawal of valproic acid treatment during pregnancy and seizure outcome: Observations from EURAP. Epilepsia [Internet]. août 2016 [cité 21 juin 2022];57(8):e173-7.

2. Edey S, Moran N, Nashef L. SUDEP and epilepsy-related mortality in pregnancy. Epilepsia [Internet]. juill 2014 [cité 21 juin 2022];55(7):e72-4.

3. Malhi GS, Bassett D, Boyce P, Bryant R, Fitzgerald PB, Fritz K, et al. Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for mood disorders. Aust N Z J Psychiatry [Internet]. déc 2015 [cité 21 juin 2022];49(12):1087-206.

4. Pacurariu A, Plueschke K, McGettigan P, Morales DR, Slattery J, Vogl D, et al. Electronic healthcare databases in Europe: descriptive analysis of characteristics and potential for use in medicines regulation. BMJ Open [Internet]. sept 2018 [cité 13 déc 2021];8(9):e023090.

5. Aghabozorgi S, Seyed Shirkhorshidi A, Ying Wah T. Time-series clustering – A decade review. Information Systems [Internet]. oct 2015 [cité 16 sept 2022];53:16-38.

6. Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. British Journal of Clinical Pharmacology [Internet]. janv 2010 [cité 21 juin 2022];69(1):4-14.

7. Close H, Reilly J, Mason JM, Kripalani M, Wilson D, Main J, et al. Renal Failure in Lithium-Treated Bipolar Disorder: A Retrospective Cohort Study. van Os J, éditeur. PLoS ONE [Internet]. 26 mars 2014 [cité 21 juin 2022];9(3):e90169.

8. Bloechliger M, Ceschi A, Rüegg S, Kupferschmidt H, Kraehenbuehl S, Jick SS, et al. Risk of Seizures Associated with Antidepressant Use in Patients with Depressive Disorder: Follow-up Study with a Nested Case–Control Analysis Using the Clinical Practice Research Datalink. Drug Saf [Internet]. avr 2016 [cité 21 juin 2022];39(4):307-21.

9. Powell G, Logan J, Kiri V, Borghs S. Trends in antiepileptic drug treatment and effectiveness in clinical practice in England from 2003 to 2016: a retrospective cohort study using electronic medical records. BMJ Open [Internet]. déc 2019 [cité 21 juin 2022];9(12):e032551.

10. Padmanabhan S, Carty L, Cameron E, Ghosh RE, Williams R, Strongman H. Approach to record linkage of primary care data from Clinical Practice Research Datalink to other health-related patient data: overview and implications. Eur J Epidemiol [Internet]. janv 2019 [cité 21 juin 2022];34(1):91-9.

11. Bezin J, Duong M, Lassalle R, Droz C, Pariente A, Blin P, et al. The national healthcare system claims databases in France, SNIIRAM and EGB: Powerful tools for pharmacoepidemiology. Pharmacoepidemiology and Drug Safety [Internet]. 2017 [cité 19 avr 2022];26(8):954-62.

12. Tuppin P, Rudant J, Constantinou P, Gastaldi-Ménager C, Rachas A, de Roquefeuil L, et al. Value of a national administrative database to guide public decisions: From the système national d'information interrégimes de l'Assurance Maladie (SNIIRAM) to the système national des données de santé (SNDS) in France. Revue d'Épidémiologie et de Santé Publique [Internet]. oct 2017 [cité 13 déc 2021];65:S149-67.

13. Blotière PO, Weill A, Dalichampt M, Billionnet C, Mezzarobba M, Raguideau F, et al. Development of an algorithm to identify pregnancy episodes and related outcomes in health care claims databases: An application to antiepileptic drug use in 4.9 million pregnant women in France. Pharmacoepidemiol Drug Saf [Internet]. juill 2018 [cité 21 juin 2022];27(7):763-70

14. Shakespeare FRCGP (1) J, isodiya FRCP (2) SM, (1) On behalf of the Royal College of General Practitioners and (2) Association of British Neurologists and Royal College of Physicians. Guidance Document on Valproate Use in Women and Girls of Childbearing Years [Internet]. 2020.

15. ANSM. [Exposure to valproic acid and its derivatives during pregnancy in France from 2007–2014: an observational study on SNIIRAM data] [Internet]. 2017.

16. Vanoli J, Nava CR, Airoldi C, Ucciero A, Salvi V, Barone-Adesi F. Use of State Sequence Analysis in Pharmacoepidemiology: A Tutorial. IJERPH [Internet]. 20 déc 2021 [cité 16 sept 2022];18(24):13398.

17. Panlilio LV, Stull SW, Bertz JW, Burgess-Hull AJ, Kowalczyk WJ, Phillips KA, et al. Beyond abstinence and relapse: cluster analysis of drug-use patterns during treatment as an outcome measure for clinical trials. Psychopharmacology [Internet]. nov 2020 [cité 16 sept 2022];237(11):3369-81.

18. Abbott A, Forrest J. Optimal Matching Methods for Historical Sequences. Journal of Interdisciplinary History [Internet]. 1986 [cité 16 sept 2022];16(3):471.

19. Dlouhy K, Biemann T. Optimal matching analysis in career research: A review and some best-practice recommendations. Journal of Vocational Behavior [Internet]. oct 2015 [cité 16 sept 2022];90:163-73.

20. Levenshtein VI. Binary codes capable of correcting deletions insertions and reversals. févr 1966;10(8):707-10.

21. Wegmann M, Zipperling D, Hillenbrand J, Fleischer J. A review of systematic selection of clustering algorithms and their evaluation. 2021 [cité 16 sept 2022];

22. Xu D, Tian Y. A Comprehensive Survey of Clustering Algorithms. Ann Data Sci [Internet]. juin 2015 [cité 16 sept 2022];2(2):165-93.

23. Anand N, Vikram P. Comprehensive Analysis & Performance Comparison of Clustering Algorithms for Big Data. Review of Computer Engineering Research [Internet]. 2017 [cité 16 sept 2022];4(2):54-80.

24. Dunkler D, Plischke M, Leffondré K, Heinze G. Augmented Backward Elimination: A Pragmatic and Purposeful Way to Develop Statistical Models. Olivier J, éditeur. PLoS ONE [Internet]. 21 nov 2014 [cité 16 sept 2022];9(11):e113677.

25. Heinze G, Wallisch C, Dunkler D. Variable selection - A review and recommendations for the practicing statistician. Biom J [Internet]. mai 2018 [cité 16 sept 2022];60(3):431-49.

26. Public Policy Committee, International Society of Pharmacoepidemiology. Guidelines for good pharmacoepidemiology practice (GPP). Pharmacoepidemiology and Drug Safety [Internet]. 2016 [cité 28 avr 2022];25(1):2-10.

27. EMA. The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology. 2010.

28. International Committee of Medical Journal Editors (ICMJE). Uniform Requirements for Manuscripts Submitted to Biomedical Journals [Internet].

29. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Journal of Clinical Epidemiology [Internet]. avr 2008 [cité 22 juin 2022];61(4):344-9.

APPENDICES

Annex 1. List of stand-alone documents

Annex 2. ENCePP Checklist for Study Protocols

Annex 3. Additional information

3.1 List of all MAHs with subsidiaries

3.2 List ICD-10 codes for eligibility

3.3 List ICD-10 codes for comorbidities, hospital/ER reason

3.4 List ATC codes for valproate and related substances

3.5 List ATC codes for antiepileptic drugs

3.6 List ATC codes for antidepressants, mood stabilizers and neuroleptics

3.7 List British National Formulary codes for exposure to valproate

3.8 List READ codes for eligibility and for the proxy 'clinical relapse'

Annex 1 List of stand-alone documents

Number	Document reference number	Date	Title
1	version 1.0	08-DEC-2020	UK (CPRD)-specific SAP
2	version 1.0	08-DEC-2020	France (SNDS)- specific SAP
3	version 2.0	17-SEP-2021	UK (CPRD) – specific SAP
4	version 2.0	7 APR 2022	France (SNDS) – specific SAP

Annex 2 ENCePP checklist for study protocols

ENCePP Checklist for Study Protocols (Revision 4) Adopted by the ENCePP Steering Group on 15 October 2018

Study title:

Non-Interventional retrospective longitudinal study in the UK and France to evaluate and identify the best practices for switching of valproate and related substances in clinical practice

EU PAS Register® number:

Study reference number (if applicable): VALSE (VALNAC09344, Sanofi internal ref. system)

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]	☒	☐	☐	6
1.1.2 End of data collection[2]	☒	☐	☐	
1.1.3 Progress report(s)	☐	☐	☒	
1.1.4 Interim report(s)	☒	☐	☐	
1.1.5 Registration in the EU PAS Register®	☒	☐	☐	
1.1.6 Final report of study results.	☒	☐	☐	

Comments:

Section 2: Research question	**Yes**	**No**	**N/A**	**Section Number**
2.1 Does the formulation of the research question and objectives clearly explain:	☒	☐	☐	
2.1.1 Why the study is conducted? (e.g. to address an important public health concern, a risk identified in the risk management plan, an emerging safety issue)	☒	☐	☐	7
2.1.2 The objective(s) of the study?	☒	☐	☐	8
2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	9.1
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)	☒	☐	☐	9.1
3.2	Does the protocol specify whether the study is based on primary, secondary or combined data collection?	☒	☐	☐	9.4
3.3	Does the protocol specify measures of occurrence? (e.g., rate, risk, prevalence)	☒	☐	☐	9.3.2
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))	☒	☐	☐	9.7.5
3.5	Does the protocol describe the approach for the collection and reporting of adverse events/adverse reactions? (e.g. adverse events that will not be collected in case of primary data collection)	☐	☐	☒	

Comments:

Section 4: Source and study populations		**Yes**	**No**	**N/A**	**Section Number**
4.1	Is the source population described?	☒	☐	☐	9.2.1
4.2	Is the planned study population defined in terms of:				
	4.2.1 Study time period	☒	☐	☐	9.1
	4.2.2 Age and sex	☒	☐	☐	9.2.1
	4.2.3 Country of origin	☒	☐	☐	9.1
	4.2.4 Disease/indication	☒	☐	☐	9.2.1
	4.2.5 Duration of follow-up	☒	☐	☐	9.1
4.3	Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	9.2.1

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)	☒	☐	☐	9.3.1
5.2	Does the protocol address the validity of the exposure measurement? (e.g. precision, accuracy, use of validation sub-study)	☐	☒	☐	

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
5.3 Is exposure categorised according to time windows?	☐	☒	☐	
5.4 Is intensity of exposure addressed? (e.g. dose, duration)	☒	☐	☐	9.3.1
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:

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?	☒	☐	☐	9.3.2
6.2 Does the protocol describe how the outcomes are defined and measured?	☒	☐	☐	9.3.2
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:

SEVERAL OUTCOMES ARE PROPOSED

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)	☒	☐	☐	9.2
7.3 Does the protocol address information bias? (e.g. misclassification of exposure and outcomes, time-related bias)	☐	☒	☐	

Comments:

Section 8: Effect measure modification	**Yes**	**No**	**N/A**	**Section Number**
8.1 Does the protocol address effect modifiers? (e.g. collection of data on known effect modifiers, sub-group analyses, anticipated direction of effect)	☐	☒	☐	

Comments:

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				9.4
9.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview)	☒	☐	☐	
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?	☒	☐	☐	
9.2 Does the protocol describe the information available from the data source(s) on:				9.4
9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	
9.2.2 Outcomes? (e.g. date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	
9.2.3 Covariates and other characteristics? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	☒	☐	☐	
9.3 Is a coding system described for:				9.4
9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☒	☐	☐	
9.3.2 Outcomes? (e.g. International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	☒	☐	☐	
9.3.3 Covariates and other characteristics?	☒	☐	☐	
9.4 Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☐	☒	☐	

Comments:

Coding system and linkage method for CPRD will be provided in the CPRD SAP.

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Section Number**
10.1 Are the statistical methods and the reason for their choice described?	☒	☐	☐	9.7
10.2 Is study size and/or statistical precision estimated?	☒	☐	☐	9.5

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Section Number**
10.3 Are descriptive analyses included?	☒	☐	☐	9.7.1
10.4 Are stratified analyses included?	☐	☒	☐	
10.5 Does the plan describe methods for analytic control of confounding?	☒	☐	☐	9.7.5
10.6 Does the plan describe methods for analytic control of outcome misclassification?	☐	☒	☐	
10.7 Does the plan describe methods for handling missing data?	☒	☐	☐	9.7.6
10.8 Are relevant sensitivity analyses described?	☐	☒	☐	

Comments:

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Section Number**
11.1 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	☒	☐	☐	9.8
11.2 Are methods of quality assurance described?	☐	☒	☐	
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?	☒	☐	☐	9.9
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)	☒	☐	☐	7.2

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?	☒	☐	☐	10
13.2 Has any outcome of an ethical review procedure been addressed?	☐	☒	☐	
13.3 Have data protection requirements been described?	☒	☐	☐	10

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?	☒	☐	☐	5

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)?	☒	☐	☐	12
15.2 Are plans described for disseminating study results externally, including publication?	☒	☐	☐	12

Comments:

Name of the main author of the protocol: Pauline Bosco-Lévy

Date: 08/November/2022

Signature: _____

Annex 3 Additional information

3.1 List of all MAHs with subsidiaries

Of note, Alfasigma and Glenmark left the Consortium on 31 December 2019.

MAH identified as contact for referral	Represented MAHs
APOTEX EUROPE B.V. Archimedesweg 2 2333 CN Leiden The Netherlands	APOTEX EUROPE B.V. Archimedesweg 2 2333 CN Leiden The Netherlands
ARISTO PHARMA GMBH Wallenroder Str. 8-10 D-13435 Berlin Germany	ARISTO PHARMA GMBH Wallenroder Str. 8-10 D-13435 Berlin Germany
ARROW GENERIQUES 26 Avenue Tony Garnier 69007 Lyon France	ARROW GENERIQUES 26 Avenue Tony Garnier 69007 Lyon France
BETAPHARM ARZNEIMITTEL GMBH Kobelweg 95 86156 Augsburg Germany	BETAPHARM ARZNEIMITTEL GMBH Kobelweg 95 86156 Augsburg Germany
BIOGARAN 15, Boulevard Charles de Gaulle 92707 Colombes Cedex France	BIOGARAN 15, Boulevard Charles de Gaulle 92707 Colombes Cedex France
BIOMO PHARMA GMBH Josef-Dietzgen-Str 3 53773 Hennef Germany	BIOMO PHARMA GMBH Josef-Dietzgen-Str 3 53773 Hennef Germany
CONSILIENT HEALTH LIMITED	CONSILIENT HEALTH 5th Floor, Beaux Lane House Mercer Street Lower Dublin 2, Ireland

CRESCENT PHARMA	CRESCENT PHARMA Sarum Hill Basingstoke RG21 8SR Royaume-Uni
DESITIN ARZNEIMITTEL GMBH Weg beim Jäger 214 22335 Hamburg Germany	DESITIN ARZNEIMITTEL GMBH Weg beim Jäger 214 22335 Hamburg Germany
GENERIS FARMACÊUTICA S.A. Rua João de Deus, 19 2700-487 Amadora Portugal	GENERIS FARMACÊUTICA S.A. Rua João de Deus, 19 2700-487 Amadora Portugal
G.L. PHARMA GMBH Schlossplatz 1 8502 Lannach Austria	Gerot PHARMAZEUTIKA GES.M.B.H. Arnethgasse 3 1160 Vienna Austria LANNACHER HEILMITTEL GES.M.B.H. Schlossplatz 1 8502 Lannach Austria
LUPIN HEALTHCARE Uk Limited The Urban Building Albert St Slough UK SL1 2BE	LUPIN HEALTHCARE Uk Limited The Urban Building Albert St Slough UK SL1 2BE
SANDOZ/HEXAL AG Industriestrasse 25 83607 Holzkirchen Germany	1 A PHARMA GMBH Keltenring 1 + 3 82041 Oberhaching Germany HEXAL AG Industriestrasse 25 83607 Holzkirchen Germany

	SANDOZ N.V. Telecom Gardens, Medialaan 40 1800 Vilvoorde Belgium SANDOZ GMBH Biochemiestrasse 10 6250 Kundl Austria SANDOZ A/S Edvard Thomsens Vej 14 2300 Copenhagen Denmark SANDOZ PHARMACEUTICALS D.D. Verovškova 57 1000 Ljubljana Slovenia SANDOZ S.P.A. Largo U. Boccioni, 1 21040 Origgio (VA) Italy SANDOZ B.V. Veluwezoom 22 1327 AH Almere The Netherlands
NEURAXPHARM ARZNEIMITTEL GMBH Elisabeth-Selbert-Str. 23 40764 Langenfeld Germany	NEURAXPHARM ARZNEIMITTEL GMBH Elisabeth-Selbert-Str. 23 40764 Langenfeld Germany

ORION CORPORATION, ORION PHARMA Orionintie 1, P.O.Box 65 FI-02101 Espoo Finland	ORION CORPORATION, ORION PHARMA Orionintie 1, P.O.Box 65 FI-02101 Espoo Finland
PHARMASWISS ČESKÁ REPUBLIKA S.R.O Jankovcova 1569/2c 170 00 Prague 7 Czech Republic	PHARMASWISS ČESKÁ REPUBLIKA S.R.O Jankovcova 1569/2c 170 00 Prague 7 Czech Republic
SANOFI AVENTIS R&D 1 Avenue Pierre Brossolette 91385 Chilly-Mazarin France	Sanofi-Aventis GmbH Austria Leonard-Bernstein-Strasse 10 Vienna 1220 Austria SANOFI BELGIUM Leonardo Da Vincilaan 19 1831 Diegem Belgium SANOFI-AVENTIS, S.R.O. Evropská 846/176a 160 00 Prague 6 Czech Republic Sanofi A/S Lyngbyvej 2 2100 København Ø Denmark Sanofi Oy Revontulenkuja 1 02100 ESPOO Finland

	Sanofi-Aventis France 82, avenue Raspail 94250 Gentilly France SANOFI-AVENTIS DEUTSCHLAND GMBH Industriepark Höchst Gebäude K703 65926 Frankfurt am Main Germany Sanofi-Aventis A.E.B.E. 348, Syngrou Ave. - Building A Athens 17674 Greece SANOFI-AVENTIS ZRT. Tó u. 5. H-1045 Budapest Hungary Sanofi Aventis Ireland Ltd. T/A SANOFI 18 Riverwalk Citywest Business Campus Dublin 24 Ireland SANOFI-AVENTIS SPA Viale Luigi Bodio, 37/B 20158 Milano Italy

	Sanofi Malta Ltd, Level 2, Fort Business Centre, Mriehel Bypass, Mriehel BKR3000 Malta SANOFI-AVENTIS SP. Z O.O. 17 Bonifraterska Str. 00-203 Warsaw Poland SANOFI – PRODUTOS FARMACÊUTICOS, LDA. Empreendimento Lagoas Park Edifício 7, 3º Piso, 2740-244 Porto Salvo Portugal SANOFI-AVENTIS SLOVAKIA S R.O. Einsteinova 24 85101 Bratislava Slovak Republic SANOFI AVENTIS, S.A. Josep Plá, 2 08019 Barcelona Spain Sanofi Lindhagensgatan 120 Stockholm 112 51 Sweden

	Sanofi-aventis Netherlands B.V. Paasheuvelweg 25 1105 BP Amsterdam The Netherlands AVENTIS PHARMA LIMITED One Onslow Street Guildford, Surrey GU1 4SY United Kingdom For Croatia, Cyprus, Bulgaria, Slovenia, Estonia, Lithuania and Latvia: Sanofi-aventis groupe 54, rue La Boétie, F-75008 Paris, France
STADA ARZNEIMITTEL AG Stadastraße 2-18 61118 Bad Vilbel Germany	EUROGENERICS N.V./S.A. Esplanade Heysel b 22 1020 Bruxelles Belgium EG LABO LABORATOIRES EUROGENERICS Central Park 9-15 Rue Maurice Mallet 92130 Issy-les-Moulineaux France ALIUD PHARMA GMBH Gottlieb-Daimler-Straße 19 89150 Laichingen Germany STADAPHARM GMBH Stadastraße 2-18 61118 Bad Vilbel Germany

	EG S.p.A. Via Pavia, 6 20136 Milano Italy CENTRAFARM B.V. Van de Reijtstraat 31-E 4814 NE Breda Netherlands
TECNIFAR S.A Rua Jose da Costa Pedreira, n° 11 B, Torre Sul 1750-130 Lisboa Portugal	TECNIFAR S.A Rua Jose da Costa Pedreira, n° 11 B, Torre Sul 1750-130 Lisboa Portugal
TEVA PHARMACEUTICALS EUROPE B.V. Piet Heinkade 107, P.O. Box 16416, 1001 RM Amsterdam The Netherlands	ABZ-PHARMA GMBH Graf-Arco-Str. 3 89079 Ulm Germany PHARMACHEMIE B.V. Swensweg 5, Box 552 2031 GA HAARLEM The Netherlands RATIOPHARM-COMERCIO E INDUSTRIA DE PRODUTOS FARMACEUTICOS LDA Lagoas Park 5-A, Piso 2, Porto Salvo 2740-245 Pt Portugal RATIOPHARM GMBH Graf-Arco-Str. 3 89079 Ulm Germany

	TEVA NEDERLAND B.V. Swensweg 5 2031 GA HAARLEM The Netherlands
VIATRIS HEALTHCARE SAS (FORMERLY MYLAN EMEA SAS) 1 rue de Turin 69007 Lyon France	MYLAN BVBA/SPRL Terhulpsesteenweg 6A B-1560 Hoeilaart Belgium MYLAN DURA GMBH Wittichstr. 6 64295, Darmstadt Germany VIATRIS HEALTHCARE SAS (FORMERLY MYLAN EMEA SAS) 1 rue de Turin 69007 Lyon France
WOCKHARDT UK LIMITED Ash Road North Wrexham Industrial Estate Wrexham LL13 9UF United Kingdom	WOCKHARDT UK LIMITED Ash Road North Wrexham Industrial Estate Wrexham LL13 9UF United Kingdom

3.2 List ICD-10 codes for eligibility

For the cohort of female patients with epilepsy	
G40	Epilepsy
G41	Status epilepticus
R56	Convulsions, not elsewhere classified

For the cohort of female patients with bipolar disorder	
F25	Schizoaffective disorders (all codes)
F31	Bipolar affective disorder (all codes)
F30	Manic episode (all codes)

3.3 List ICD-10 codes for comorbidities, hospital/ER reason

For the cohort of female patients with epilepsy	
G40	Epilepsy
G41	Status epilepticus
R56	Convulsions, not elsewhere classified

For the cohort of female patients with bipolar disorder	
F10	Mental and behavioural disorders due to use of alcohol
F11	Mental and behavioural disorders due to use of opioids
F12	Mental and behavioural disorders due to use of cannabinoids
F13	Mental and behavioural disorders due to use of sedatives or hypnotics
F14	Mental and behavioural disorders due to use of cocaine
F15	Mental and behavioural disorders due to use of other stimulants, including caffeine
F16	Mental and behavioural disorders due to use of hallucinogens
F17	Mental and behavioural disorders due to use of tobacco
F18	Mental and behavioural disorders due to use of volatile solvents
F19	Mental and behavioural disorders due to multiple drug use and use of other psychoactive substances
F20	Schizophrenia (all codes)
F25	Schizoaffective disorders (all codes)
F30	Manic episode (all codes)
F31	Bipolar affective disorder (all codes)
F32	Depressive episode (all codes)
F33	Recurrent depressive disorder
F34	Persistent mood [affective] disorders

For the cohort of female patients with bipolar disorder	
F38	Other mood [affective] disorders
F40	Phobic anxiety disorders (all codes)
F41	Other anxiety disorders (all codes)
F42	Obsessive-compulsive disorder (all codes)
F43	Reaction to severe stress, and adjustment disorders
F50	Eating disorders (all codes)
F51	Nonorganic sleep disorders (all codes)
F60	Specific personality disorders

3.4 List ATC codes for valproate and related substances

Drug class (Seizure control) / Substance	ATC WHO code
Valproic acid, magnesium valproate, sodium valproate, sodium valproate/ valproic acid, valproate semisodium	N03AG01
Valpromide	N03AG02

3.5 List ATC codes for antiepileptic drugs

Drug class (Seizure control) / Substance	ATC WHO code
Carbamazepine	N03AF01
Gabapentin	N03AX12
Lacosamide	N03AX18
Lamotrigine	N03AX09
Levetiracetam	N03AX14
Oxcarbazepine	N03AF02
Phenobarbital	N03AA02
Phenytoin	N03AB02
	N03AB52
Pregabalin	N03AX16
Topiramate	N03AX11
Zonisamide	N03AX15
Clobazam	N05BA09
Clonazepam	N03AE01

Drug class (Seizure control) / Substance	ATC WHO code
Eslicarbazepine acetate	N03AF04
Ethosuximide	N03AD01
	N03AD51
Felbamate	N03AX10
Fosphenytoin	N03AB05
Mesuximide	N03AD03
Perampanel	N03AX22
Primidone	N03AA03
Retigabine	N03AX21
Rufinamide	N03AF03
Stiripentol	N03AX17
Sultiame	N03AX03
Tiagabine	N03AG06
Vigabatrin	N03AG04

3.6 List ATC codes for antidepressants, mood stabilizers and neuroleptics

Drug class	Substance	ATC WHO code
Antidepressants	Agomelatine	N06AX22
	Amitriptyline	N06AA09
		N06CA01
	Bupropion	N06AX12
		A08AA62
	Citalopram	N06AB04
	Clomipramine	N06AA04
	Doxepin	N06AA12
	Duloxetine	N06AX21
	Escitalopram	N06AB10
	Fluoxetine	N06AB03
		N06CA03
	Imipramine	N06AA02
	Maprotiline	N06AA21
	Mianserin	N06AX03
	Mirtazapine	N06AX11
	Nortriptyline	N06AA10

Drug class	Substance	ATC WHO code
	Paroxetine	N06AB05
	Reboxetine	N06AX18
	Sertraline	N06AB06
	Tranylcypromine	N06AF04
	Trimipramine	N06AA06
	Venlafaxine	N06AX16
Mood stabilizers	Carbamazepine	N03AF01
	Lamotrigine	N03AX09
	Lithium	N05AN (N05AN01)
Typical neuroleptics	Haloperidol	N05AD01
Atypical neuroleptics	Amisulpride	N05AL05
	Aripiprazole	N05AX12
	Asenapine	N05AH05
	Clozapine	N05AH02
	Olanzapine	N05AH03
	Paliperidone	N05AX13
	Quetiapine	N05AH04
	Risperidone	N05AX08
	Ziprasidone	N05AE04

3.7 List British National Formulary codes for exposure to valproate

Drug class (Seizure control) / Substance	BNF codes
Sodium valproate	04020300/0408100
Valproate semisodium	04020300/04070402/04080100
Valproic acid	04070402/04080100

3.8 List of READ codes

List of READ codes for epilepsy diagnosis

READ code	Read term	Selected for Diagnosis (Y/N)
F13z500	Benign neonatal sleep myoclonus	Y
6674	Epilepsy associated problems	Y
667L.00	Epilepsy does not limit activities	Y
667J.00	Epilepsy impairs education	Y
667K.00	Epilepsy limits activities	Y
667M.00	Epilepsy management plan given	Y
667H.00	Epilepsy prevents employment	Y
667G.00	Epilepsy restricts employment	Y
667N.00	Epilepsy severity	Y
F256000	Hypsarrhythmia	Y
F257.00	Kojevnikov's epilepsy	Y
667P.00	No seizures on treatment	Y
F258.00	Post-ictal state	Y
667F.00	Seizure free >12 months	Y
667C.00	Epilepsy control good	Y
Fyu5200	[X]Other status epilepticus	Y
Fyu5900	[X]Status epilepticus, unspecified	Y
667Q.00	1 to 12 seizures a year	Y
667S.00	1 to 7 seizures a week	Y
667R.00	2 to 4 seizures a month	Y
F254500	Complex partial epileptic seizure	Y
F25y300	Complex partial status epilepticus	Y
667T.00	Daily seizures	Y
667W.00	Emergency epilepsy treatment since last appointment	Y
F254400	Epileptic automatism	Y
F250300	Epileptic seizures - akinetic	Y
F250200	Epileptic seizures - atonic	Y
F251200	Epileptic seizures - clonic	Y
F251300	Epileptic seizures - myoclonic	Y
F251400	Epileptic seizures - tonic	Y
F251600	Grand mal seizure	Y
F253.00	Grand mal status	Y
F253.00	Grand mal status	Y
667V.00	Many seizures a day	Y
F252.00	Petit mal status	Y
F252.00	Petit mal status	Y
F255600	Simple partial epileptic seizure	Y
F25X.00	Status epilepticus, unspecified	Y
667D.00	Epilepsy control poor	Y
F253.11	Status epilepticus	Y
F25z.11	Fit (in known epileptic) NOS	Y
Fyu5000	[X]Other generalized epilepsy and epileptic syndromes	Y
Fyu5100	[X]Other epilepsy	Y
F25B.00	Alcohol-induced epilepsy	Y
F25y400	Benign Rolandic epilepsy	Y
F25y000	Cursive (running) epilepsy	Y
F25C.00	Drug-induced epilepsy	Y
F25..00	Epilepsy	Y
1O30.00	Epilepsy confirmed	Y
F25z.00	Epilepsy NOS	Y
F25y100	Gelastic epilepsy	Y
F250z00	Generalised nonconvulsive epilepsy NOS	Y
F251z00	Generalised convulsive epilepsy NOS	Y
F251.00	Generalised convulsive epilepsy	Y
F250.00	Generalised nonconvulsive epilepsy	Y
F251000	Grand mal (major) epilepsy	Y

READ code	Read term	Selected for Diagnosis (Y/N)
F250400	Juvenile absence epilepsy	Y
F25A.00	Juvenile myoclonic epilepsy	Y
F250500	Lennox-Gastaut syndrome	Y
F254300	Limbic system epilepsy	Y
F25y200	Locl-rlt(foc)(part)idiop epilep&epilptic syn seiz locl onset	Y
F25D.00	Menstrual epilepsy	Y
F251100	Neonatal myoclonic epilepsy	Y
667B.00	Nocturnal epilepsy	Y
F25y.00	Other forms of epilepsy	Y
F25yz00	Other forms of epilepsy NOS	Y
F25y500	Panayiotopoulos syndrome	Y
F254z00	Partial epilepsy with impairment of consciousness NOS	Y
F254.00	Partial epilepsy with impairment of consciousness	Y
F255z00	Partial epilepsy without impairment of consciousness NOS	Y
F255y00	Partial epilepsy without impairment of consciousness OS	Y
F255.00	Partial epilepsy without impairment of consciousness	Y
F255000	Jacksonian, focal or motor epilepsy	Y
F250000	Petit mal (minor) epilepsy	Y
F25F.00	Photosensitive epilepsy	Y
F254100	Psychomotor epilepsy	Y
F254200	Psychosensory epilepsy	Y
F250100	Pykno-epilepsy	Y
F255100	Sensory induced epilepsy	Y
F25G.00	Severe myoclonic epilepsy in infancy	Y
F255200	Somatosensory epilepsy	Y
F25E.00	Stress-induced epilepsy	Y
F254000	Temporal lobe epilepsy	Y
F251500	Tonic-clonic epilepsy	Y
F255500	Unilateral epilepsy	Y
F255300	Visceral reflex epilepsy	Y
F255400	Visual reflex epilepsy	Y
F256.12	West syndrome	Y
F255011	Focal epilepsy	Y
F255012	Motor epilepsy	Y
F255311	Partial epilepsy with autonomic symptoms	Y
F259.11	Ohtahara syndrome	Y
F251111	Otohara syndrome	Y
F251y00	Other specified generalised convulsive epilepsy	Y
F250y00	Other specified generalised nonconvulsive epilepsy	Y
F25G.11	Dravet syndrome	Y
F259.00	Early infant epileptic encephalopathy wth suppression bursts	Y
F256.11	Lightning spasms	Y
Eu84200	[X]Rett's syndrome generalised	Y
F035200	Rasmussen syndrome partial	Y
F130600	Aicardi Goutieres syndrome mixed	Y
F132111	Unverricht - Lundborg disease generalised	Y
P228300	Aicardi syndrome mixed	Y
PK61.00	Sturge-Weber syndrome partial	Y
PKyz511	Angelman syndrome generalised	Y
PKyz700	Angelman's syndrome generalised	Y
PKyz711	Angelman syndrome generalised	Y
ZS82.11	Landau-Kleffner syndrome mixed	Y
Eu80300	[X]Acquired aphasia with epilepsy [Landau - Kleffner] mixed	Y
F132100	Progressive myoclonic epilepsy generalised	Y
F251011	Tonic-clonic epilepsy generalised	Y
SC20000	Traumatic epilepsy partial	Y

List of READ codes for epilepsy outcome

READ code	Read term	Selected for Outcome (Y/N)
1B1W.00	Transient epileptic amnesia	Y
1B27.00	Seizures in response to acute event	Y
1B64.00	Had a convulsion	Y
1B64.11	Convulsion - symptom	Y
282..00	O/E - fit/convulsion	Y
282..11	O/E - a convulsion	Y
2828	Absence seizure	Y
282Z.00	O/E - fit/convulsion NOS	Y
F132z12	Myoclonic seizure	Y
F250011	Epileptic absences	Y
F25H.00	Generalised seizure	Y
R003.00	[D]Convulsions	Y
R003100	[D]Convulsions, infantile	Y
R003400	[D]Nocturnal seizure	Y
R003y00	[D]Other specified convulsion	Y
R003z00	[D]Convulsion NOS	Y
R003z11	[D]Seizure NOS	Y
Ryu7100	[X]Other and unspecified convulsions	Y
1B63.00	Had a fit	Y
1B63.11	Fit - had one, symptom	Y
282..12	O/E - a fit	Y
2824,11	O/E - Jacksonian fit	Y
2824,12	O/E - focal fit	Y
R003200	[D]Fit	Y
282..13	O/E - a seizure	Y
Fyu5200	[X]Other status epilepticus	Y
Fyu5900	[X]Status epilepticus, unspecified	Y
667Q.00	1 to 12 seizures a year	Y
667S.00	1 to 7 seizures a week	Y
667R.00	2 to 4 seizures a month	Y
F254500	Complex partial epileptic seizure	Y
F25y300	Complex partial status epilepticus	Y
667T.00	Daily seizures	Y
667W.00	Emergency epilepsy treatment since last appointment	Y
667D.00	Epilepsy control poor	Y
F254400	Epileptic automatism	Y
F250300	Epileptic seizures - akinetic	Y
F250200	Epileptic seizures - atonic	Y
F251200	Epileptic seizures - clonic	Y
F251300	Epileptic seizures - myoclonic	Y
F251400	Epileptic seizures - tonic	Y
F25z.11	Fit (in known epileptic) NOS	Y
F251600	Grand mal seizure	Y
F253.00	Grand mal status	Y
F256.11	Lightning spasms	Y
667V.00	Many seizures a day	Y
2824	O/E - focal (Jacksonian) fit	Y
2822	O/E - grand mal fit	Y
2823	O/E - petit mal fit	Y
2825	O/E - psychomotor fit	Y
F252.00	Petit mal status	Y
F255600	Simple partial epileptic seizure	Y
F253.11	Status epilepticus	Y
F25X.00	Status epilepticus, unspecified	Y

List of READ codes for bipolar disorder

Read code	Read term	Selected for Diagnosis (Y/N)	Selected for Outcome* (Y/N)
Eu31400	[X]Bipol aff disord, curr epis sev depress, no psychot symp	Y	Y
Eu31300	[X]Bipolar affect disorder cur epi mild or moderate depressn	Y	Y
Eu31500	[X]Bipolar affect dis cur epi severe depres with psyc symp	Y	N
Eu31100	[X]Bipolar affect disorder cur epi manic wout psychotic symp	Y	N
Eu31000	[X]Bipolar affective disorder, current episode hypomanic	Y	Y
Eu31700	[X]Bipolar affective disorder, currently in remission	Y	N
Eu31800	[X]Bipolar affective disorder type I	Y	N
Eu31900	[X]Bipolar affective disorder type II	Y	N
Eu31200	[X]Bipolar affect disorder cur epi manic with psychotic symp	Y	N
Eu31600	[X]Bipolar affective disorder, current episode mixed	Y	Y
Eu31.00	[X]Bipolar affective disorder	Y	N
Eu31z00	[X]Bipolar affective disorder, unspecified	Y	N
Eu31y00	[X]Other bipolar affective disorders	Y	N
E11y100	Atypical manic disorder	Y	N
E115.00	Bipolar affective disorder, currently depressed	Y	Y
E114.00	Bipolar affective disorder, currently manic	Y	Y
E115600	Bipolar affective disorder, now depressed, in full remission	Y	N
E115500	Bipolar affect disord, now depressed, part/unspec remission	Y	Y
E115400	Bipolar affect disord, now depressed, severe with psychosis	Y	Y
E115200	Bipolar affective disorder, currently depressed, moderate	Y	Y
E115300	Bipolar affect disord, now depressed, severe, no psychosis	Y	Y
E115100	Bipolar affective disorder, currently depressed, mild	Y	Y
E115z00	Bipolar affective disorder, currently depressed, NOS	Y	Y
E115000	Bipolar affective disorder, currently depressed, unspecified	Y	Y
E114600	Bipolar affective disorder, currently manic, full remission	Y	N
E114500	Bipolar affect disord,currently manic, part/unspec remission	Y	Y
E114400	Bipolar affect disord, currently manic,severe with psychosis	Y	Y
E114200	Bipolar affective disorder, currently manic, moderate	Y	Y
E114300	Bipolar affect disord, currently manic, severe, no psychosis	Y	Y
E114z00	Bipolar affective disorder, currently manic, NOS	Y	Y
E114100	Bipolar affective disorder, currently manic, mild	Y	Y
E114000	Bipolar affective disorder, currently manic, unspecified	Y	Y
E116600	Mixed bipolar affective disorder, in full remission	Y	N
E116500	Mixed bipolar affective disorder, partial/unspec remission	Y	N
E116400	Mixed bipolar affective disorder, severe, with psychosis	Y	N
E116200	Mixed bipolar affective disorder, moderate	Y	N
E116300	Mixed bipolar affective disorder, severe, without psychosis	Y	N
E116.00	Mixed bipolar affective disorder	Y	N
E116z00	Mixed bipolar affective disorder, NOS	Y	N
E116100	Mixed bipolar affective disorder, mild	Y	N
E116000	Mixed bipolar affective disorder, unspecified	Y	N
E11yz00	Other and unspecified manic-depressive psychoses NOS	Y	N
E11y.00	Other and unspecified manic-depressive psychoses	Y	N

Read code	Read term	Selected for Diagnosis (Y/N)	Selected for Outcome* (Y/N)
E11y300	Other mixed manic-depressive psychoses	Y	N
E111600	Recurrent manic episodes, in full remission	Y	N
E111500	Recurrent manic episodes, partial or unspecified remission	Y	Y
E111400	Recurrent manic episodes, severe, with psychosis	Y	Y
E111200	Recurrent manic episodes, moderate	Y	Y
E111300	Recurrent manic episodes, severe without mention psychosis	Y	Y
E111z00	Recurrent manic episode NOS	Y	Y
E111100	Recurrent manic episodes, mild	Y	Y
E111.00	Recurrent manic episodes	Y	Y
E111000	Recurrent manic episodes, unspecified	Y	Y
E117600	Unspecified bipolar affective disorder, in full remission	Y	N
E117500	Unspecified bipolar affect disord, partial/unspec remission	Y	N
E117400	Unspecified bipolar affective disorder,severe with psychosis	Y	N
E117200	Unspecified bipolar affective disorder, moderate	Y	N
E117300	Unspecified bipolar affective disorder, severe, no psychosis	Y	N
E117100	Unspecified bipolar affective disorder, mild	Y	N
E117.00	Unspecified bipolar affective disorder	Y	N
E11y000	Unspecified manic-depressive psychoses	Y	N
E117z00	Unspecified bipolar affective disorder, NOS	Y	N
E117000	Unspecified bipolar affective disorder, unspecified	Y	N
Eu30.11	[X]Bipolar disorder, single manic episode	Y	Y
Eu31911	[X]Bipolar II disorder	Y	N
Eu31y11	[X]Bipolar II disorder	Y	N
Eu34012	[X]Cycloid personality	Y	N
Eu34000	[X]Cyclothymia	Y	N
Eu34013	[X]Cyclothymic personality	Y	N
E11..11	Bipolar psychoses	Y	N
E11..13	Manic psychoses	Y	N
E110.00	Manic disorder, single episode	Y	Y
E110.11	Hypomanic psychoses	Y	N
E110000	Single manic episode, unspecified	Y	Y
E110100	Single manic episode, mild	Y	Y
E110200	Single manic episode, moderate	Y	Y
E110300	Single manic episode, severe without mention of psychosis	Y	Y
E110400	Single manic episode, severe, with psychosis	Y	Y
E110500	Single manic episode in partial or unspecified remission	Y	Y
E110600	Single manic episode in full remission	Y	N
E110z00	Manic disorder, single episode NOS	Y	Y
E114.11	Manic-depressive - now manic	Y	Y
E115.11	Manic-depressive - now depressed	Y	Y
E211100	Hypomanic personality disorder	Y	N
E211300	Cyclothymic personality disorder	Y	N
Eu30.00	[X]Manic episode	Y	Y

Read code	Read term	Selected for Diagnosis (Y/N)	Selected for Outcome* (Y/N)
Eu30000	[X]Hypomania	Y	N
Eu30100	[X]Mania without psychotic symptoms	Y	N
Eu30200	[X]Mania with psychotic symptoms	Y	N
Eu30211	[X]Mania with mood-congruent psychotic symptoms	Y	N
Eu30212	[X]Mania with mood-incongruent psychotic symptoms	Y	N
Eu30y00	[X]Other manic episodes	Y	Y
Eu30z00	[X]Manic episode, unspecified	Y	Y
Eu30z11	[X]Mania NOS	Y	N
Eu31.11	[X]Manic-depressive illness	Y	N
Eu31.12	[X]Manic-depressive psychosis	Y	N
Eu31.13	[X]Manic-depressive reaction	Y	N
Eu31y12	[X]Recurrent manic episodes	Y	Y
Eu33213	[X]Manic-depress psychosis,depressd,no psychotic symptoms	Y	N
Eu33312	[X]Manic-depress psychosis,depressed type+psychotic symptoms	Y	N
1BD5.00	High suicide risk	N	Y
TK...00	Suicide and selfinflicted injury	N	Y
TK...14	Suicide and self harm	N	Y
TK...15	Attempted suicide	N	Y
U2...13	[X]Suicide	N	Y
U2...14	[X]Attempted suicide	N	Y
ZQ54.00	Suicide risk assessment	N	Y

*Records indicating '**current**' OR '**curr**' OR '**episode**' OR '**now**' OR '**suicide**' are retained as potential codes to consider for the outcome definition and will be confirmed in the SAP, after consultation with the CPRD support staff.