Pharmacologie médicale

ALCAPONE

Alert generation using the case-population approach in the French claims databases

Protocol

Version 1.1, April 5th 2016

This study has received funding from the French Ministry of Health.
(PREPS, 2014, 14-0635)

Code gestionnaire: CHUBX 2014/39

Bordeaux PharmacoEpi
Plateforme de recherche en pharmaco-épidémiologie
Service de Pharmacologie médicale, CIC Bordeaux CIC1401
INSERM - Université de BORDEAUX - CHU de Bordeaux - Adera
Bâtiment Le Tondu - case 41
146 rue Léo Saignat - 33076 Bordeaux Cedex

STUDY INFORMATION

Title	Alert generation using the case-population approach in the French claims databases
Protocol version identifier	Version 1.1
Date of last version of protocol	April 5th 2016
IMPACT study number	N/A
EU PAS register number	Study to be registered
Active substance	Positive and negative controls from OMOP reference set
Medicinal product	N/A
Product reference	Negative controls from OMOP reference set
Procedure number	N/A
Marketing authorisation holder(s)	N/A
Joint PASS	No
Research question and objectives	To assess the suitability of the French nationwide healthcare insurance system database (SNIIRAM and EGB) for drug safety signal generation based on the OMOP reference set and the case-population approach.
Country of study	France
Author	Pr. Nicholas Moore, Head of CIC1401 Bordeaux PharmacoEpi (BPE), Service de Pharmacologie médicale, INSERM CIC1401 Université de Bordeaux – CHU de Bordeaux – Adera Bâtiment Le Tondu – case 41 146 rue Léo Saignat – 33076 Bordeaux Cedex, France ☎ +33 (0)557 571 560- Fax: +33 (0)557 574 740 nicholas.moore@u-bordeaux.fr

1 TABLE OF CONTENTS

2 LIST OF ABBREVIATIONS

ADR	Adverse drug reaction
ALCAPONE	Alert generation using the case-population approach in the French claims databses
ALD	*Affection longue durée* (Long Term Disease)
ALI	Acute Liver Injury
ATC	Anatomical Therapeutic Chemical
BPE	Bordeaux PharmacoEpi
CNAMTS	*Caisse National de l'Assurance Maladie des Travailleurs Salariés* (National healthcare insurance system for salaried workers)
CNIL	*Commission Nationale de l'Informatique et des Libertés* (French data protection commission)
CPR	Case population ratio
DDD	Defined Daily Dose
DGOS	Direction générale de l'Offre de Soins
EGB	*Echantillon généraliste de bénéficiaires* (1/97 random sample of the national health insurance database)
EHR	Electronic health records
EMA	European Medecines Agency
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EU-ADR	Exploring and Understanding Adverse Drug Reactions
GVP	Guideline on good pharmacovigilance practices
ICD10	International Classification of Diseases 10th revision
ICD9CM	International Classification of Diseases, 9th revision, Clinical Modification
IDS	*Institut des Données de Santé* (Institute of health data)
INSERM	*Institut National de la Santé et de la Recherche Médicale* (National medical research institute)
ISPE	International Society of Pharmacoepidemiology
KI	Acute Kidney Injury
LEOPARD	Longitudinal Evaluation of Observational rofiles of Adverse events Related to Drugs
LGPS	Longitudinal Gamma Poisson Shrinker
LTD	Long-term disease (registration for major chronic diseases with full insurance coverage of all claims related to disease)
MI	Myocardial infarction
NSAIDs	Non-steroidal anti-inflammatory drugs
OMOP	Observational Medical Outcome Partnership
OR	Odds ratio
PMSI	*Programme de Médicalisation des Systèmes d'Information* (National hospital discharge summary database)
PREPS	Programme de recherche sur la performance du système des soins
RI	Relative incidence
ROC	Receptor operating characteristics
RR	Relative risk

SAP	Statistical Analytical Plan
SAR	Statistical Analysis Report
SNIIRAM	*Système national d'information inter-régimes de l'Assurance maladie* (National healthcare insurance system database)
SR	Spontaneous reporting
UGIB	Upper Gastrointestinal bleeding
US	United States

3 RESPONSIBLE PARTIES

SCIENTIFIC COMMITTEE	
Pr. Jacques Benichou Biostatistician	CHU de Rouen Université de Rouen
Dr. Martijn Schuemie Epidemiologist	Janssen Research and Development Erasmus University Medical Center of Rotterdam
Dr. Jeremy Rassen Epidemiologist	Aetion, Inc. Brigham and Women's Hospital
Dr. Joshua Gagne Epidemiologist	Brigham and Women's Hospital Harvard Medical School
Dr. Alain Weil Public Health specialist	Caisse nationale de l'assurance maladie des travailleurs salaries (CNAMTS)
COORDINATING CENTRE	
Bordeaux PharmacoEpi (BPE) **INSERM CIC1401** **Service de Pharmacologie médicale**	**Université de Bordeaux – CHU de Bordeaux – Adera** Bâtiment 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
Pr Nicholas Moore Head of CIC1401	☎ +33 (0)557 571 560 nicholas.moore@u-bordeaux.fr
Dr Patrick Blin Scientific and medical director	☎ +33 (0)557 579 563 patrick.blin@u-bordeaux.fr
Cécile Droz-Perroteau Chief operating officer	☎ +33 (0)557 574 737 cecile.droz@u-bordeaux.fr
Nicolas Thurin Scientific advisor	☎ +33 (0)557 579 209 nicolas.thurin@u-bordeaux.fr
Régis Lassalle Statistics & data management manager	☎ +33 (0)557 574 764 regis.lassalle@u-bordeaux.fr
SPONSOR	
Direction Générale de l'Offre de Soins (DGOS)	14 avenue Duquesne 75007 Paris SP 07 ☎ +33 (0)140 566 000 – Fax: +33 (0)140 566 066 http://www.sante.gouv.fr
MANAGER	
Centre Hospitalier Universitaire de Bordeaux	12 rue Dubernat 33400 Talence ☎ +33 (0)556 795 679 http://www.chu-bordeaux.fr
Joaquin Martinez *Directeur de la recherche Clinique et de l'Innovation*	☎ +33 (0)556 795 346 joaquin.martinez@chu-bordeaux.fr
Sylvie Blazejewski *Responsable de la « Promotion CHU de Bordeaux »*	☎ +33 (0)557 820 313 sylvie.blazejewski@chu-bordeaux.fr
Céline Bairras-Martin *Responsable d'Etudes Cliniques*	☎ +33 (0)557 820 329 celine.bairras-martin@chu-bordeaux.fr

4 ABSTRACT

TITLE

Alert generation using the case-population approach in the French claims databases

RATIONAL AND BACKGROUND

Post-marketing surveillance is crucial for the identification of previously unknown ADRs and in particular rare and severe ADRs. Longitudinal observational healthcare databases, such as administrative claims databases, represent a huge source of information for safety signal management. Some research projects, such as Observational Medical Outcome Partnership (OMOP) in the United States (US), have tried to perform safety signal detection across claims databases. To this intent, seven analytic methods for identifying risk in observational healthcare data were evaluated. All methods were applied to a reference set composed of 165 positive and 234 negative control drug-event pairs across four outcomes: acute liver injury (ALI), acute myocardial infarction (MI), acute kidney injury (KI), and upper gastrointestinal bleeding (UGIB).

However the OMOP group did not test the case-population approach. To our knowledge, only one study has explored the performance of this method as tool for signal detection. Yet, the case-population design is ideally suited to situations of rare exposures and rare events, which is typical of the alert generation environment.

France has a nationwide healthcare insurance system database: the *Système National d'Information Inter-régimes de l'Assurance Maladie* (SNIIRAM). SNIIRAM currently covers about 98,8% of the French population (66.6 millions persons). It is on of the largest homogeneous claims database in the world. A 1/97 permanent representative sample of SNIIRAM, the *Echantillon Généraliste de Bénéficiaires* (EGB) is also available. SNIIRAM has not been tested for alert generation so far. It is large enough that rare events, such as those in the OMOP reference set or others that are common reasons for removing drugs from the market, can be identified and previous exposures determined. In these circumstances, the combination of OMOP analytic methods and the case-population approach applied to SNIIRAM could provide a very simple and effective method to identify emerging risks, including rare but very serious events.

RESEARCH QUESTION AND OBJECTIVES

The research question is to assess the suitability of the French nationwide healthcare insurance system database (SNIIRAM and EGB) for drug safety signal generation based on the OMOP reference set and methodologies, and the case-population approach.

The main objectives are:

- To assess the performance of SNIIRAM for the detection of drug safety signals based on the OMOP reference set and methodologies
- To develop on SNIIRAM the case-population approach and assess its performance for safety signal generation based on the OMOP reference set.

Specific objectives deriving from the main objectives are:

- To describe qualitative changes required for the adaptation of the OMOP reference set to SNIIRAM
- To assess the feasibility of the project through EGB
- To apply OMOP methodologies and the case-population approach on a combination of SNIIRAM and EGB, and compare their performance.

Secondary objective: To test the application of the method on incident alerts and drugs not described in the reference set.

STUDY DESIGN

This study will be a several-stage study using data from the national claims and hospital databases (SNIIRAM and EGB), based on the validation of the database according to an adapted version of the OMOP reference set.

OMOP reference set consists of four main events of interest: (i) ALI, (ii) MI, (iii) KI, and (iv) UGIB. For each of these events, the OMOP reference set lists a series of

molecules that have (positive controls) or have not (negative controls) been associated with the events of interest. Cases were reviewed and linked to the relevant ICD10 codes. Reference set, especially positive controls, will be adapted to the French market.

The specificities of access to SNIIRAM require that for each access to the database, a single extraction process is defined and implemented. Because of this, multiple exposure cohorts are rather complex compared to case-based extraction. ALCAPONE will therefore rely on case-based approaches only. Thus, for each sub-study, three case-based approaches will be used and compared to measure the association between positive and negative controls, and events of interest:

- **Case-control design**

These compare the rate of exposure in a group of subjects presenting a given event (the cases) to that in a group not presenting the event (the controls).

- **Self-controlled case series**

These consist in comparing each individual to his/herself.

- **Case-population approach**

In this method, exposure in cases is compared to the general population exposure to the drug of interest.

POPULATION

Study population is the whole French population as identified from the French nationwide healthcare system claim databases SNIIRAM and EGB.

- **Cases**

Cases will be defined according to the four main events of interest: (i) ALI, (ii) MI, (iii) KI, and (iv) UGIB. For each event of interest cases will be extracted separately from SNIIRAM.

- **Controls**
  - **Case-control design:** controls will be extracted from EGB.
  - **Self-controlled case series:** extraction of controls is not required; control periods and risk periods are both defined based on cases observation period.
  - **Case-population approach: "**controls" consist of the whole population. Reference data will be extracted from EGB.

VARIABLES

Exposures to the drugs of interest (positive and negative controls) for the case-control design and the case-population approach will be defined initially as any exposure within 30 days before event onset. For the self-controlled case series, exposure matches with the risk period (drug dispensation period followed by a 30 days risk window). Sensitivity analysis will be done regarding exposure windows, risk windows, and several exclusions periods.

Principal **index date** will be considered as the date of hospital admission for the event of interest. Secondary index date will be used for sensitivity analysis.

For the case-control approach, age, gender as well as a disease risk score will be used as **matching variables**.

Evaluation criteria will consist in the number of true positive and true negative pairs identified in SNIIRAM compared to all positive and negative pairs tested, for each of the four events and each of the methodological approach.

DATA SOURCES

SNIIRAM 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. It currently includes more than 98.8% of the French population of 66.6 million persons. To the extent that SNIIRAM is a national database, all subjects are

followed-up from birth (or immigration) to death (or emigration), even if a subject changes occupations or retires. EGB is a 1/97 permanent representative sample of SNIIRAM, with the same linked hospital discharge database and death registry. Both of them contain individual anonymised information on, among others: demographics, outpatient reimbursed healthcare expenditures including drugs, and hospital-discharge summaries from PMSI including ICD10 diagnosis codes.

Access to SNIIRAM is regulated and needs approval from IDS (Institute of health data) and CNIL (French data protection commission). EGB is however readily available: a direct access is possible through the French National Institute of Health and Medical Research (*Institut national de la santé et de la recherche médicale* - INSERM).

STUDY SIZE

Number of cases extracted from SNIIRAM will vary according to the event of interest frequency. A preliminary EGB extraction will enable the estimation of the annual number of cases by event, and facilitate the determination of SNIIRAM length of extraction in order to reach enough power.

DATA ANALYSIS

A statistical analysis plan (SAP) will be developed and validated by the scientific committee before analysis. The statistical analysis will be performed using SAS® software (latest current version), following the SAP.

Three different designs are envisaged:

- **Case-control design**

Odds ratios (OR) will be calculated using a conditional logistic regression. Several degrees of matching will be considered, going from non-matched approach to disease risk score matching including loose matching on simply age and sex.

- **Self-controlled case series**

Risk periods will be determined based on exposition and presumed biological mechanisms. Relative incidences (RI) for the risk periods will be computed.

- **Case-population approach**

Case population ratio (CPR) will be calculated according to two distinct exposure approaches: (i) a person-time approach and (ii) a *per-user* approach.

The following **evaluation criteria** will be used to compare designs performance

The receptor operating characteristics (ROC) will be used to choose the best compromise between sensitivity and specificity among the different designs and their variants according to drug-event pair characteristics.

MILESTONES

Study protocol	Feb-Mar 2016
IDS authorization	Q2 2016
CNIL authorization	Q3-Q4 2016
Development of the Statistical Analysis Plan	Q2-Q3 2016
EGB data extraction	Q3 2016
Description of EGB cases, power calculation and sensitivity analysis	Q3-Q4 2016
SNIIRAM data extraction	Q1-Q2 2017
Data management and statistical analysis	Q2-Q4 2017
Final report	Q1 2018
Manuscript	Q2-Q3 2018

5 AMENDMENTS AND UPDATES

Number	Date	Section of study protocol	Amendment or update	Reason

6 MILESTONES

Milestones	Planned Date*
Study protocol	Feb-Mar 2016
IDS authorization	Q2 2016
CNIL authorization	Q3-Q4 2016
Development of the Statistical Analysis Plan	Q2-Q3 2016
EGB data extraction	Q3 2016
Description of EGB cases, power calculation and sensitivity analysis	Q3-Q4 2016
SNIIRAM data extraction	Q1-Q2 2017
Data management and statistical analysis	Q2-Q4 2017
Final report	Q1 2018
Manuscript	Q2-Q3 2018

*Trimesters after funding

7 RATIONALE AND BACKGROUND

7.1 GENERALITIES

During the pre-approval phase, clinical trials may fail to identify adverse drug reactions (ADRs) as they are limited in size (and lack power to identify rare ADRs), often restricted to patients without co-morbidities or co-medications and reflect a carefully controlled setting that corresponds poorly to real-word clinical practice. Therefore, post-marketing surveillance is crucial for the identification of previously unknown ADRs and in particular rare and severe ones. Currently, spontaneous reporting is the general standard for ADRs detection (1-4). It is invaluable for signal generation. Although mandatory in France (5), spontaneous reporting is based on the motivation of the reporter; therefore, under-reporting, biases due to selective reporting and incomplete information lead to fears of missing some signals, especially for very rare and very serious events. Moreover the lack of a denominator, *i.e.* the inability to determine what percent of the population was exposed and what percent experienced the adverse events, seriously limit the value of spontaneous reporting for relevant safety signal assessment (4, 6, 7).

Unlike spontaneous report data, longitudinal observational healthcare databases, such as administrative claims data, or electronic health records, or hospital data, are typically accrued automatically and prospectively. Thus, information about events and exposures are collected independently and are therefore largely unaffected by selection biases (8). Electronic health records (EHRs) may also contain a large numbers of time-stamped medical records from routine clinical practice (9). As a result, such data represent a valuable source of information for safety signal strengthening and validation. Furthermore, because they catch the very first prescriptions of new drugs prospectively, these databases may have an enormous potential for early detection of drug safety signals. In broad outline, the methods used to explore longitudinal observational data can be divided in a few main categories based on the entry in the study through exposure (cohort-based designs) or events (case-based designs). Other general design options can be considered, especially concerning the control groups that range from the self-controlled methods to population-wide controls (10).

7.2 PREVIOUS WORKS

7.2.1 Analytics methods for identifying risk in observational healthcare data

Some research projects such as Exploring and Understanding Adverse Drug Reactions project (EU-ADR) (11-13) in Europe or the Sentinel Initiative (previously Mini-Sentinel) in the United States (US) (14) have tried to combine data from several databases to increase the power to detect new signals. Alongside these projects, the Observational Medical Outcomes Partnership (OMOP), a public-private partnership, was created aiming at identifying the most reliable methods for analyzing huge volumes of data drawn from heterogeneous sources (15).

In 2013, Drug Safety devoted a supplement entitled *Studying the Science of Observational Research: Empirical Finding from the Observational Medical Outcomes Partnership (OMOP)*. In this supplement, the performance of seven analytic methods for identifying risk in observational healthcare data was evaluated (16-22) and compared (23): disproportionality safety signaling (16), new user cohort method (19), case-control method (17), self-controlled case series design (22), self-controlled cohort method (20),

calibrated self-controlled cohort analysis within temporal pattern discovery (18), and Longitudinal Gamma Poisson Shrinker (LGPS) (21). The performance of each method was evaluated in five large-scale US observational databases (four claims databases of respectively 1.2, 4.6, 10.8 and 46.5 million persons, and one EHR of 11.2 million persons) through a reference set composed of 165 positive and 234 negative control drug-event pairs across four outcomes: acute liver injury (ALI), acute myocardial infarction (MI), acute kidney injury (KI), and upper gastrointestinal bleeding (UGIB). These same methods were also applied to six European EHRs or claims databases from EU-ADR (11, 12) covering nine million persons, by using the same test cases. In both experiments, self-controlled methods appear to provide better discrimination between positive and negative controls and lesser bias than other study designs (23, 24). However, it has been shown that all methods require calibration to ensure proper interpretation of study results.

A simulated observational dataset of 10 million persons was also used to evaluate the performance of these methods (25). By construction in the simulation, datasets with no effect (Relative Risk (RR) =1) and injected RRs of 1.25, 1.5, 2,4 and 10 were created and applied to the 165 positive and 234 negative control situations. When simulated data were injected with RR >2, all designs had good predictive accuracy, but when RR<2, no methods achieved 100% predictions.

7.2.2 Case population approach

Yet, the OMOP group did not test the case-population approach (26-28). In the case-population approach (also called population-based case-cohort study or case-only study), the exposure to a suspect drug in patients with a particular event (the cases) is compared to the overall exposure to this drug in the entire population from which the cases are extracted, using drug sales or actual number of users in population databases (10, 26-31).

Studies have compared results provided by the case population approach to other study designs (28, 32-35). For all drug-event pairs identified, the case population ratio (CPR) computed by means of the case-population approach and sales data are of the same order of magnitude as the corresponding relative risks estimated with a standard control group. Besides, Bordeaux PharmacoEpi team has investigated the case-population design to look into risk for very rare events, *i.e.* acute liver failure leading to transplantation (10, 26, 27, 31, 36). Results (*cf.* Appendix 1) strengthened by a theoretical demonstration showed that the rarer are the exposure and the event rate, the better the CPR approximates the actual RR (Appendix 2) (10). To our knowledge, only one study has explored the performance of the case-population design as tool for signal detection (37). In this study, authors concluded that the method was suitable for detecting signals of possible teratogenicity as long as databases are larger enough to identify cases for drugs not commonly used.

In circumstances where exposure in the population and cases are either rare or very rare, which is typical of the alert generation environment, cohort and case-control studies may be ineffective because of power issues or because of confounding bias (the low level of exposition makes matching or adjustment impossible). Conversely, case-population approach seems to be ideally suited to these circumstances (10, 27); once the method characteristics and performances have been established, case-population method implementation could lead to a very simple, rapid and effective way to conduct systematic surveillance of emerging risks in preselected rare but very serious events.

7.3 FRANCE NATIONWIDE HEALTHCARE DATABASE

France has a nationwide claims and hospital database: the *Système National d'Information Inter-régimes de l'Assurance Maladie* (SNIIRAM). SNIIRAM covers currently about 98,8% of the French population (66.6 millions persons) and comprises all reimbursed medical expenses. It is also linked to the national hospital-discharge summaries database system (PMSI) and the national death registry. A 1/97 permanent representative sample of SNIIRAM with the same linked hospital discharge database and death registry, the *Echantillon Généraliste de Bénéficiaires* (EGB), is also available. A direct access to EGB is possible through the French National Institute of Health and Medical Research (*Institut national de la santé et de la recherche médicale* - INSERM).

SNIIRAM is one of the largest homogeneous claims database in the world; this should ensure the identification of rare events, such as those in the OMOP reference set or others that are common reasons for removing drugs from the market. The determination of previous exposures is also possible. Nevertheless, to fully tap these resources, a previous calibration of the database is required, firstly through OMOP methodologies and secondly regarding the case population approach.

8 RESEARCH QUESTION AND OBJECTIVES

The research question is to assess the suitability of the French nationwide healthcare insurance system database (SNIIRAM and EGB) for drug safety signal generation based on the OMOP reference set and methodologies, and the case-population approach.

The main objectives are:

- To assess the performance of SNIIRAM for the detection of drug safety signals based on the OMOP reference set and methodologies
- To develop on SNIIRAM the case-population approach and assess its performance for safety signal generation based on the OMOP reference set.

Specific objectives deriving from the main objectives are:

- To describe qualitative changes required for the adaptation of the OMOP reference set to SNIIRAM
- To assess the feasibility of the project through EGB
- To apply OMOP methodologies and the case-population approach on a combination of EGB and SNIIRAM, and compare their performance.

Secondary objective: To test the application of the method on incident alerts and drugs not described in the reference set.

9 RESEARCH METHODS

9.1 STUDY DESIGN

This study will be a several-stage study using historical data in the national claims databases SNIIRAM and EGB, based on the validation of the database according to an adapted version of the OMOP reference set.

9.1.1 OMOP reference set

OMOP reference set consists of four main events of interest: (i) ALI, (ii) MI, (iii) KI, and (iv) UGIB. For each of these events, the OMOP reference set lists a series of molecules that have (positive controls) or have not (negative controls) been associated with the events of interest. In order to take into account specificities relative to each event dataset, ALCAPONE study will be divided in four sub-studies that will be treated in parallel:

- ALCAPONE acute Liver Injury (ALCAPONE-ALI),
- ALCAPONE Myocardial Infarction (ALCAPONE-MI),
- ALCAPONE acute Kidney Injury (ALCAPONE-KI),
- ALCAPONE upper Gastro-Intestinal bleeding (ALCAPONE-UGIB).

9.1.2 Methodological approaches

The specificities of access to SNIIRAM require that for each access to the database, a single extraction process is defined and implemented. Because of this, multiple exposure cohorts are rather complex compared to case-based extraction. ALCAPONE will therefore rely on case-based approaches only. Thus, for each sub-study, three case-based approaches will be used and compared to measure the association between positive and negative controls, and events of interest.

9.1.2.1 Case-control design

These compare the rate of exposure in a group of subjects presenting a given event (the cases) to that in a group not presenting the event (the controls). The overall design of the study is presented in **Figure 1**.

Data will be extracted from 1 January 2009 to 31 December 2014 from SNIIRAM for the cases and from EGB for the controls. Principal index date will be considered for each sub-study as the date of hospital admission for the first occurrence of the event of interest during the extraction period. In ALCAPONE-ALI, a secondary index date will be used for sensitivity analysis 10 days prior to hospitalization to take into account the clinical evolution of liver disease before hospital admission (prodrome). Drug dispensation periods will be used to approximate drug exposure periods. A sensitivity analysis will exclude all first dispensations within a given time previous to index date (exclusion periods). A drug contributes to the exposure to drug count for a particular patient in the case group or control group if the whole or a part of the drug dispensation period is within the exposure window (*cf. 9.3.2*. section).

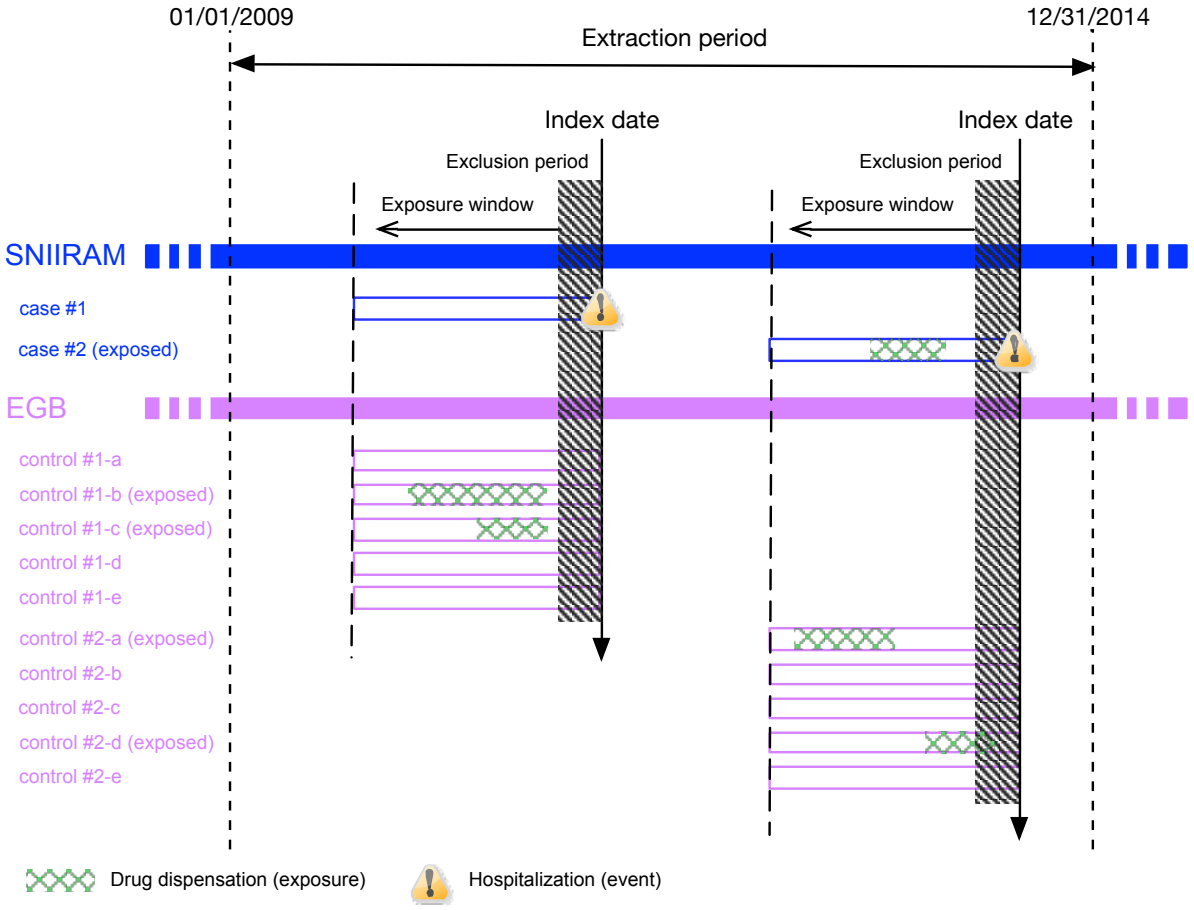

Figure 1: Diagram of the case-control design for 2 cases and the corresponding controls

9.1.2.2 Self-controlled case series

These consist in comparing each individual to his/herself: the event rate during the determined risk periods is compared to the event rate during the baseline risk periods. The overall design of the study is presented in **Figure 2**.

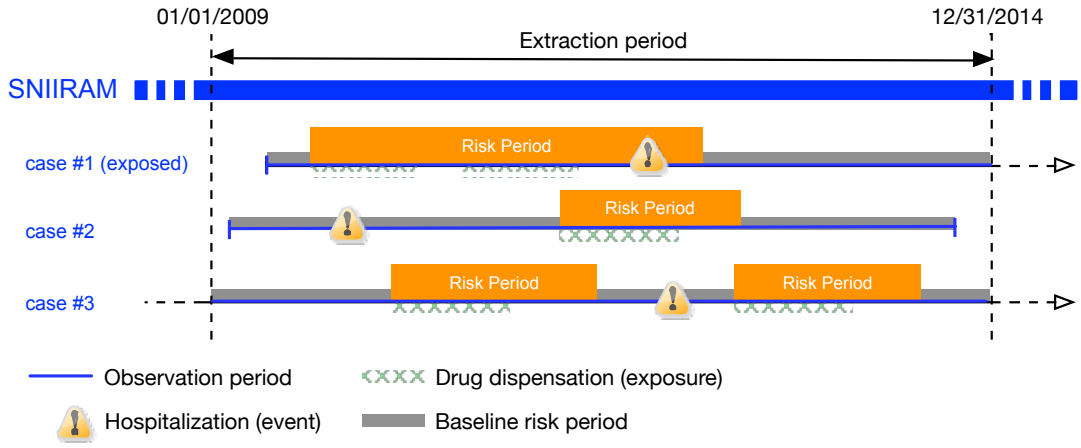

Figure 2: Diagram of the self-controlled case series design

Data will be extracted from 1 January 2009 to 31 December 2014 from SNIIRAM. Observation period will depend on how long the patient record lasts in the extraction period. Risk periods will be determined according to drug dispensation periods and presumed biological mechanisms (*cf. 9.3.2.* section). For each case, only the first occurrence of the event of interest in the observation period will be considered. A drug

contributes to the exposure to drug for the patient in the “case group” if the event is within the risk period.

9.1.2.3 Case-population approach

In this method (**Figure 3**), exposure in cases is compared to the general population exposure to the drug of interest. Two approaches will be envisaged: (i) the *per-user* approach and (ii) the person-time approach. Index dates, exclusion periods and exposure windows are defined in the same way as for case-control design (9.1.2.1 section).

Data will be extracted from 1 January 2009 to 31 December 2014 from SNIIRAM. The corresponding number of exposed and non-exposed users will be extracted from EGB for the same date.

Exposed and non-exposed person-time in the reference population will be estimated from Medic'AM[1] through the determination of the Defined Daily Dose[2] (DDD) per year for each one of the drugs of interest.

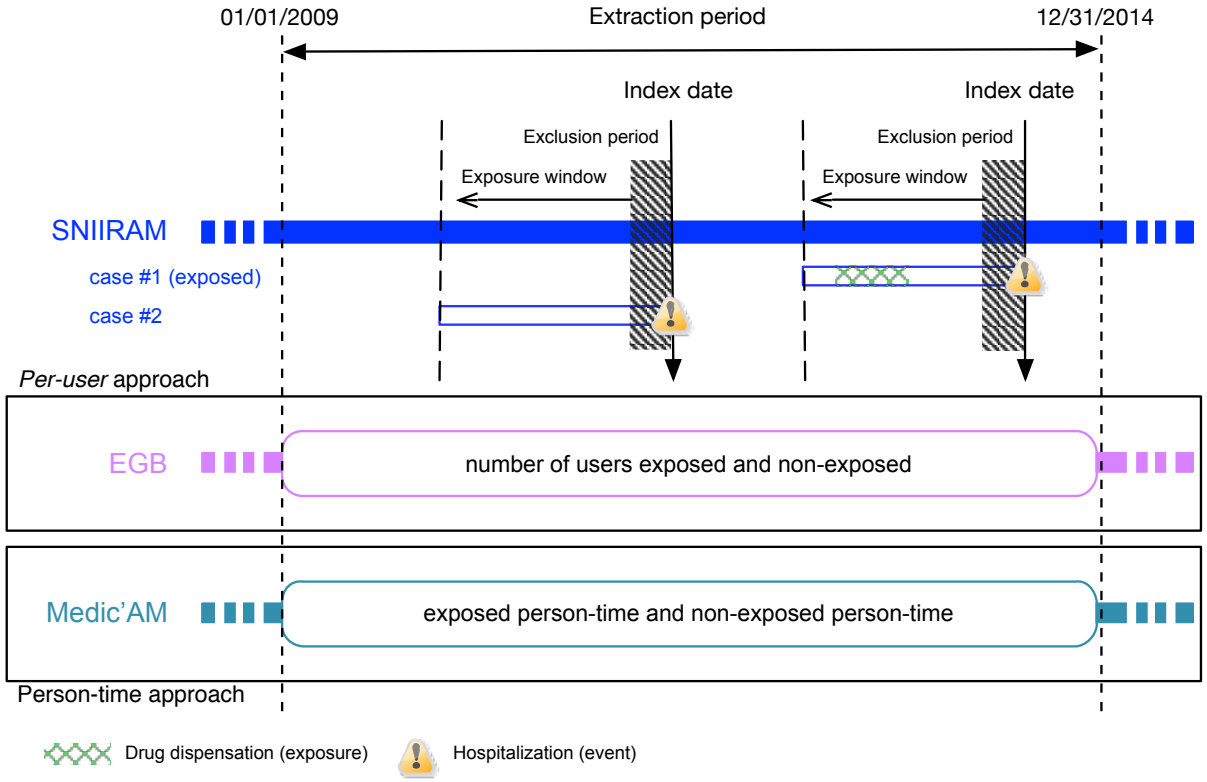

Figure 3: Diagram of the case-population approach

9.2 SETTING

Study population is the whole French population as identified from the French nationwide healthcare system claim databases SNIIRAM and EGB.

[1] Medic’AM is a database, based on years, presenting detailed information regarding rembursed drugs (including the number of rembursed boxes sold for each speciality).
[2] The DDD is the assumed average maintenance dose per day for a drug used for its main indication in adults.

9.2.1 Cases

Cases will be defined according to the four main events of interest: (i) ALI, (ii) MI, (iii) KI, and (iv) UGIB.

For each event of interest, International Classification of Diseases, 9th revision, Clinical Modification (ICD9CM) event codes used in the original OMOP exercise have been mapped to the relevant ICD10 code(s) present in the French databases. Then, obtained codes have been compared with ICD10 codes used in other studies on the same events in the same databases (26, 27, 31, 38, 39), or defined by *ad-hoc* study groups set up to that purpose. The resulting list includes ICD10 codes retained for more or less specific case selection (Appendix 3) (40).

Subsequently, cases (defined by their ICD10 codes) corresponding to each sub-study will be subjected to a specific extraction:

- From EGB in a first step in order to estimate the annual number of cases
- From SNIIRAM afterwards.

For each case, the extraction will include:

- Date of hospitalization (index date), and data concerning the index hospital admission.
- Case demographics.
- All dispensation data for the year previous to index date
- All medical resource utilization data
- Registration for chronic diseases (ALD) at index date and previously
- All hospital admission data and diagnoses for the year previous to index date.

9.2.1.1 ALCAPONE-ALI

Acute liver injury inclusion and exclusion ICD10 codes have been validated in the EPIHAM study (26, 27, 31). Acute liver injury will be defined as “toxic liver disease” and “acute and subacute hepatic failure”. Besides, patients presenting one of the following associated diagnoses will be excluded:

- Chronic viral hepatitis
- Malignant neoplasms
- Mental and behavioural disorders due to use of alcohol
- Other degenerative diseases of nervous system
- Heart failure
- Portal vein thrombosis
- Oesophageal varices
- Alcoholic liver disease
- Fibrosis and cirrhosis of liver
- Other diseases of liver
- Cholelithiasis
- Other diseases of biliary tract
- Ascites
- Presence of cardiac and vascular implants and grafts.

Corresponding inclusion and exclusion ICD10 codes are available in Appendix 3.

A sensitivity analysis will be conducted on “Toxic liver disease with hepatitis, not elsewhere classified”, and “Toxic liver disease, unspecified” codes.

9.2.1.2 ALCAPONE-MI

ALCAPONE-MI will cover, MI, acute transmural myocardial infarction and acute subendocardial myocardial infarction. Due to the recent enlargement of the MI definition, “Unstable angina” code will be included in a sensitivity analysis.

Codes were already investigated in MILO (occurrence of MI in users of low-dose and prescription strength NSAIDs) (39), B2PAC (impact of beta-blockers post-MI in secondary prevention) (38), HORUS (Health Outcomes, Resource Use, costs in patients with Stable coronary artery disease a cohort study in the EGB database) (41) and ENGEL-AVK (Real-life anticoagulants benefit-risk in atrial fibrillation in France) (42). Corresponding ICD10 codes are available in Appendix 3.

9.2.1.3 ALCAPONE-KI

Acute kidney injury will be defined as acute kidney failure. Corresponding codes are presented in Appendix 3.

9.2.1.4 ALCAPONE-UGIB

Upper gastrointestinal bleeding will be defined as gastric, duodenal, peptic or gastrojejunal ulcer. Acute haemorrhagic gastritrits, haematemesis and melaena will also be included. Oesophageal varices with bleeding, Gastro-oesophageal laceration-haemorrhage syndrome, and Gastrointestinal haemorrhage will be included in the sensitivity analysis. Corresponding codes are presented in Appendix 3.

9.2.2 Controls

The need for controls depends on study design:

- **Case-control design:** controls will be extracted from EGB patients that do not have the PMSI codes defined in Appendix 3.
- **Self-controlled case series:** extraction of controls is not required; baseline risk periods (control periods) and risk periods are both defined based on cases observation period.
- **Case-population approach:** “controls” consist of the whole population. Reference data will be extracted from EGB.

9.3 VARIABLES

9.3.1 Definition of drugs of interest

For each of the previously cited events, the OMOP reference set lists a series of molecules that have (positive controls) or have not (negative controls) been associated with the events of interest. The original set is available in Appendix 4. However this list was initially devised for the OMOP initiative in the US and not all are available on the French Market. In addition, some are not reimbursed by the French healthcare system and do not therefore appear in the database, even if they are available on the market. As a consequence, the presence in the National database of each of the positive and negative control molecules will be checked. Moreover some drugs specific to the French/European market and well known for their association with the events of interest will be added to the initial reference set.

In a further stage, once the SNIIRAM is calibrated, the method will be tested on drugs suspected of causing one of the four events of interest.

9.3.2 Definition of Exposure

Exposure definition varies according to the study design.

- **For the case-control design and the case-population approach**, exposure to the drugs of interest will be defined initially as any exposure within 30 days before event onset (exposure window). This is the definition of exposure for instance for hepatic injury, (43-45). Different exposure windows will be tested and compared to identify the one(s) that give the best performance for each event, and possibly event and exposure typology.

  To avoid protopathic bias[3], a sensitivity analysis will exclude all first dispensations within a given time previous to index date. Several exclusion periods will be tested (*e.g.* 7, 14, 21, 30 days). This does not concern repeat dispensations of previously dispensed drugs.
- **For the self-controlled case series**, exposure matches with the risk period. Risk periods are composed of a drug dispensation period followed by a risk window. Risk windows will be defined initially as a 30 days period following the drug dispensation period. For example a chronic treatment usually dispensed in 30-day or 28-day boxes will generate a 60 days risk period whereas a 10 days antibiotic treatment will generate a 40 days risk period (10 dispensation days + 30 days risk window). Risk periods may also vary according to presumed biological mechanisms. Different risk windows will be tested and compared to identify the one(s) that give the best performance for each event, and possibly event and exposure typology.

9.4 DATA SOURCES

9.4.1 SNIIRAM

SNIIRAM database is the nationwide healthcare insurance system database that contains individual anonymous information on all reimbursed outpatient claims and is linked to the national hospital-discharge summaries database system (PMSI) and the national death registry, using a unique national pseudonymised identifier (46). It currently includes the 3 main healthcare insurance systems, and several other smaller ones, which represent more than 98.8% of the French population (66.6 million subjects) from birth (or immigration) to death (or emigration), even if a subject changes occupations or retires. SNIIRAM contains individual information on (47, 48):

- General characteristics: gender, year of birth, affiliation scheme, area of residence, date of death,
- Long-term disease (LTD, or ALD in French, and associated ICD10 codes) with starting and ending date. There is a list of 30 LTD for a total of 3448 available ICD10 codes, which includes most of chronic diseases with long term and/or expensive treatment; *e.g.* a disease such as atrial fibrillation is specified by the ICD10 code within LTD. Registration with an LTD is obtained at the request of a

[3] Prescription because of early pre-hospital symptoms of disease *e.g.* NSAIDs prescription for deferred pain from angina or for early symptoms of hepatitis.

patient's general practitioner and must be validated by the health insurance system physician. Once registered, patients receive full (*i.e.* 100%) reimbursement for expenditure related to the LTD, as defined by the health authorities. 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 an LTD registration;

- Outpatient reimbursed healthcare expenditures: visits, medical procedures, medical imaging, lab tests, drugs, medical devices, transport, sick leaves *etc.* with prescriber and other health professionals information (medical or paramedical specialty, private/public practice), dates (prescription and dispensation), and codes (but not the medical indication nor result);
- Hospital-discharge summaries from the PMSI: ICD10 diagnosis codes (primary, linked and associated diagnosis) for all medical, obstetric and surgery hospitalizations, with the date and duration, medical procedures and cost coding system. 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 to 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 ICD10 classification (*e.g.* chemotherapy session) for a chronic or LTD disease. It indicates the pathology at the origin of the care procedure. Linked 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).
    - 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).

Non-hospital data are updated monthly with a lag of 6 months to reach about 99% of the information uploaded and hospital-discharge summaries yearly during the third quarter for the data of the previous year. Access to SNIIRAM is regulated and needs approval from institute of health data (*Institut des Données de Santé* - IDS) and the French data protection commission (*Commission Nationale de l'Informatique et des Libertés* - CNIL).

9.4.2 EGB

EGB (*Echantillon Généraliste des Bénéficiaires*) is a permanent 1/97th representative sample of SNIIRAM and currently includes about 700 000 subjects with more than 10 years history. Unlike SNIIRAM, EGB is readily available: access is free and full for certain entities fixed by ministerial order, including accredited INSERM units such as INSERM CIC1401 Bordeaux PharmacoEpi platform.

To the extent that the access to EGB is easier and faster than SNIIRAM, the use of EGB is favored. It allows the description of disease or drug utilization that is relatively frequent, as well as to prepare the protocol and the analysis of a SNIIRAM study. SNIIRAM is required when the size of EGB does not have the statistical power to

answer the question, and generally when the question includes “real-life” benefit and/or risk of a drug or class, as well as comparative risk or effectiveness.

9.5 STUDY SIZE

A preliminary EGB extraction will enable the estimation of the annual number of cases by event of interest, and facilitate the determination of SNIIRAM length of extraction in order to reach enough power.

Number of cases to extract from SNIIRAM will vary according to the event of interest frequency. For example, ALIs were already investigated in EPIHAM study (26, 27, 31); 63 corresponding cases were identified in EGB between 2009 and 2013. By extrapolation around 6,100 are expected in SNIIRAM on the same period, being 1,220 ALIs/year. On the other hand, about 1,000 MIs are yearly identifiable in EGB, *i.e.*, 97,000 by year.

9.6 DATA MANAGEMENT

EGB extractions will be done in April, after the study protocol has been sent to INSERM.

SNIIRAM extraction criteria will be described in a data extraction plan approved prior to initiating extraction. Data extraction will be done by the CNAMTS. The Bordeaux PharmacoEpi data manager in charge of the project will validate the population extracted by the CNAMTS from SNIIRAM using EGB data extraction. In SNIIRAM the extraction period will depend on the number of cases identified in EGB for each event of interest: the rarer is the event, the longer the extraction period would be.

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

Data will be stored following a strict standardized operating procedure described under the CNIL.

9.7 DATA ANALYSIS

Statistical analysis will be performed using SAS® software (SAS Institute, latest current version, North Carolina, USA), and followed the SAP developed and validated by the scientific committee before analysis.

The following three different designs are envisaged:

9.7.1 Case-control design

The case-control design compares the rate of exposure in a group of subjects presenting a given event (the cases) to that in a group not presenting the event (the controls).

Odds ratios (OR) will be calculated according to standard procedures using conditional logistic regression models (49). Several degrees of matching will be considered, going from non-matched approach to disease risk score matching including loose matching on simply age and sex. Disease risk scores are the probability in a given population of having the disease of interest, independent of exposure to drugs. Thus, disease risk scores will be computed for all cases in SNIIRAM population, and random controls matched on age, sex and disease risk scores will be identified in EGB (*cf.* 9.2.2 section).

9.7.2 Self-controlled case series

This approach consists in comparing each individual to his/herself. As a general principle, such designs compare the occurrence of the event and exposure within risk periods to the baseline risk (*i.e.* control periods) (50, 51).

Risk periods will be determined based on exposition and presumed biological mechanisms. Seasonal variations of events and exposures will be studied in order to select for each case one or more control period(s) that would have the same probability of exposure. Relative incidence (RI) for the risk periods will then be computed.

9.7.3 Case-population approach

In this method, exposure in cases is compared to the general population exposure to the drug of interest (10, 27, 28, 31). As explicated in 9.2.2 section, reference data will be extracted from EGB, which is a random sample of the case source population.

Case population ratio (CPR) will be calculated according to two distinct exposure approaches: (i) a person-time approach and (ii) a per-user approach (27).

9.7.3.1 Person-time approach

In the person-time approach, the exposure to the drug of interest in the reference population can be estimated from statistics of drug sales (*e.g.* Medic'AM) in the general population where cases arose. Consequently, the cells of the contingency table in the control group are expressed not in number of individuals but in person-time units (*e.g.* person-months), the latter representing the sum of the durations of exposure (PTE) and not-exposed time (PTNE).

	Cases	Population (person-time or persons)
Exposed	a	PTE
Not exposed	c	PTNE
Total	n	PTPOP

PTE=sum of the durations of exposure; PTNE= sum of the durations of non-exposed time; PTPOP=total person-time.

Figure 4: Comparison of exposure in cases with aggregated exposure data from the entire population from which the cases were identified

Under the null hypothesis (*i.e.* no association between exposure and event), and in the absence of selection bias, the ratio of exposed to not-exposed cases (a/c) is not expected to differ from the ratio of exposed person-time to not-exposed person-time (PTE/PTNE) in the source-population. If the exposure in events is rare and the exposure in the population is small, one can also use the (a/n) / (PTE/ PTPOP) ratio. The confidence interval for CPR can be computed by using the formula used for the odds of exposure in cases a/c and from the assumption that the ratio PTE/PTNE does not contribute to the variability because it is based on very large numbers (28). The person-time approach is especially relevant if the event is thought to be dose/duration dependent (type A reaction) (27).

9.7.3.2 Per-user approach

An alternative is to use as denominator the number of users of the drug of interest over the study period, especially when the event mechanism is patient-dependent (*e.g.* immuno-allergic or genetically determined). This can be obtained from the full SNIIRAM database for very low exposures, since this includes all users in France. For drugs where the utilization data shows more common exposure, even if the events of interest are very rare, exposure may be obtained from EGB. Denominator will then be the number of users over the study period.

	Cases	Population (persons)
Exposed	a	e
Not exposed	c	f
Total	n	N

Figure 5: Comparison of exposure in cases with the number of users exposed and non exposed from the entire population from which the cases were identified

In this case the CPR is (a/n) / (e/N), with the usual confidence interval (10).

If the number of exposed events is great enough to identify drug usage patterns among cases, this distribution can be compared to the drug usage patterns among population users, to build hazard functions. This cannot be done from sales data unless usage pattern is described elsewhere, but it is easily done in the claims databases. These hazard functions can be used to identify putative drug mechanisms or event typology as A, B, or late type reactions (27, 52). Knowing the type of adverse reaction will also contribute to the choice of the best denominator for the case-population approach: person-years for type A reactions that are dose- and duration-dependent, persons for type B reactions.

9.7.4 Evaluation criteria

The evaluation criteria will consist in the number of true positive and true negative pairs identified in SNIIRAM/EGB for the four events of interest, compared to all positive and negative pairs tested, for each of the four events and each of the three methods and their variants.

From this, the receptor operating characteristics (ROC) will be used to choose the best compromise between sensitivity and specificity among the different designs and their variants according to drug-event pair characteristics.

9.8 QUALITY CONTROL

The Bordeaux PharmacoEpi - INSERM CIC1401 team, has implemented a quality management system for all its activities. CNAMTS data extraction will be validated using the expected population size estimated using EGB. 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

9.9.1 Selection bias

Since all subjects identified with extraction criteria will be extracted from a national database, there is no study selection bias, nor attrition bias, except very rare withdrawals from one of the healthcare insurance systems including and covering more than 95% of the French population.

9.9.2 Information bias

- **Event of interest:** Clinical outcomes will be defined using the ICD10 discharge primary diagnosis. Miscoding cannot be excluded but should be sparse for the clearly defined events studied.
- **Drug exposure:** Positive control and negative control drugs exposure will be assessed using exhaustive non-hospital drug claims. Drugs prescribed during hospital stays are not recorded and could represent a potential risk of exposure underestimation. However, it should concern few subjects for a very short period of time, and the impact over a study period of several years should be negligible.

9.9.3 Counfouding bias

SNIIRAM is one of the largest homogeneous claims database in the world. Nevertheless, it was built for administrative and reimbursement purposes. The lack of clinical data and biological results, including some risk factors (*e.g.* smoking status, body mass index, blood pressure, and cholesterol values), which may or not be confounders, form its main limit.

Statistical methods such as disease risk scores were developed to improve control of confounding using a large number of covariates ascertained from patients' healthcare claims data, as these variables may collectively be proxies for unmeasured confounders. Other study design, such as self-controlled case series also allowed to address these potential confounders as soon as they are not time depending.

9.9.4 Specific limitations of the methods

9.9.4.1 Case-control design

The main limitations with this method are (i) the difficulty to identify controls that must be as similar as possible to the cases, except for the presence of the event of interest, (ii) its limited statistical power when the prevalence of exposure amongst controls is low, which is usually the case for medicines recently introduced to the market. In this case, however, the size of the population available for the selection of controls is very large, so that there is essentially no limit to the number of controls, which may indeed be selected randomly.

9.9.4.2 Self-controlled case series

Self-controlled case series do not control for the time- or age-dependent risk since at the time of risk periods patients are not the same age as for baseline risk period. In the same way, the occurrence of an event may lead to reduced future exposure: for instance the occurrence of an MI may reduce the probability of future use of NSAIDs, and therefore spuriously increase the apparent association between NSAIDs and MI (the Casablanca effect).

Hence, the main limitations with these methods are (i) the applicability of the design to the drug-event pair studied, (ii) the difficulty to precisely define adequate risk and control periods, and (iii) the limitations with prolonged exposures or chronic events. By requiring exchangeability between exposed and unexposed time periods, the use of self-controlled designs has been restricted in epidemiology to specific scenarios, such as those with intermittent exposure and acute outcome onset.

However, self-controlled case series are a very powerful ways of controlling for confounding variables and selection bias that do not vary over time, *e.g.:* baseline health status, genetic predispositions, diets *etc.*

9.9.4.3 Case-population approach

Because of its selection process of "controls" (*i.e.* the whole source population), the case-population approach does not take into account the individual characteristics and exposure patterns of the controls. These gives rise to at least two majors limitation (28): (i) the difficulty of controlling for potential confounders and (ii) the necessity to include the totality or a representative sample of cases that have occurred in the considered area during the study period (CPR is correlated to the respective number of exposed and non-exposed cases. Any bias in the selection process would have a major impact on it).

Concerning this last aspect, meeting such a requirement would be close to impossible for non-serious adverse reactions; but the four selected events of interest seem to be ideally suited for case–population studies: (i) case definition are rather simple, (ii) there is exhaustive case ascertainment given that patients suffering ALI, KI or MI have a strong probability of being hospitalized, and that only UGIB leading to hospitalization are targeted. Thus, only patients who die before hospitalization will be missed.

9.9.5 Perspective

This method will calibrate the French nationwide healthcare insurance system database (SNIIRAM and EGB) for the four events of interest, thus making related signal generation and management possible. Exploration of other events using the same methodology will therefore be facilitated. This could provide a simple system to identify or confirm drug related alert regarding some determined or suspected events of interest.

9.10 OTHER ASPECTS

ALCAPONE project is structured in successive tasks corresponding to several livrables:

Livrable	Timeline*
Livrable 1.1: ICD10 code list corresponding to the events of interest (principal and sensitivity analysis) with the number of events identificated in EGB.	Q1
Livrable 1.2: List of the drugs of the OMOP reference set available in France and identifiable in SNIIRAM, with their corresponding number of users in EGB.	Q1
Livrable 1.3: Power computation for each drug-event pair	Q1
Livrable 2.1a: Case description including demographics, LTD and main drug classes for each event of interest in EGB.	Q3
Livrable 2.1b: Case description including demographics, LTD and main drug classes for each event of interest in SNIIRAM.	Q6
Livrable 2-2 a: Number of cases exposed to the reference drugs for each event of interest in EGB and corresponding sensitivity analysis.	Q4
Livrable 2-2 b: Number of cases exposed to the reference drugs for each event of interest in SNIIRAM and corresponding sensitivity analysis.	Q6
Livrable 3.1: Statistical analysis plan (SAP)	Q2
Livrable 3.2a: Results of the three statistical approaches in EGB.	Q6

Livrable 3.2b: Results of the three statistical approaches in SNIIRAM.	Q7
Livrable 4: Final report	Q8
Livrable 5: Papers	Q10.

*Trimesters after funding

10 PROTECTION OF HUMAN SUBJECTS

This project is a database analysis with individual anonymous information for which subject informed consent is not required. Data extraction from SNIIRAM is regulated and needs approval from Institute of Health Data (*Institut des Données de Santé* - IDS) and French data protection commission (*Commission Nationale de l'Informatique et des Libertés* - CNIL).

EGB is fully anonymised and has been released access without need of CNIL authorization. However a previous declaration to INSERM is required to trace its proper use. EGB utilization is subject to a convention aiming to a secure and effective data use as part of projects helping to improve health policies and healthcare quality.

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 European Medicines Agency (EMA) Guideline on good pharmacovigilance practices cited above (GVP VI*), as well as the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology.

* The latest revision of the Guideline on good pharmacovigilance practices (GVP) Module VI – Management and reporting of adverse reactions to medicinal products (Rev 1) from EMA (coming into effect 16 Sept 2014) specifies: *For Non-interventional post-authorisation studies based on secondary use of data (VI.C.1.2.1.b): "The design of such studies is characterised by the secondary use of data previously collected from consumers or healthcare professionals for other purposes. Examples include medical chart reviews (including following-up on data with healthcare professionals), analysis of electronic healthcare records, systematic reviews, meta-analyses. For these studies, the reporting of suspected adverse reactions in the form of ICSRs is not required. Reports of adverse events/reactions should be summarised as part of any interim safety analysis and in the final study report unless the protocol provides for different reporting*".

12 PLANS FOR DISSEMINATING AND COMMUNATING STUDY RESULTS

This database analysis will be performed by the Bordeaux PharmacoEpi, INSERM CIC1401, an academic research organization, for which scientific communication and publication is a major component of its activities. All results will be made public, as a report to the French Ministry of Health (*via* the DGOS), the CNAMTS and to French regulatory authorities.

In addition, results will be presented at the annual meetings of the International Society of Pharmacoepidemiology (ISPE), and proposed for publication in either Pharmacoepidemiology and Drug Safety, or in Drug Safety. Drug-specific or event-specific results may be submitted in addition to specialist journals.

13 REFERENCES

1. Meyboom RH, Lindquist M, Egberts AC, Edwards IR. Signal selection and follow-up in pharmacovigilance. Drug Saf. 2002;25(6):459-65. PubMed PMID: 12071784. Epub 2002/06/20. eng.

2. Paludetto MN, Olivier-Abbal P, Montastruc JL. Is spontaneous reporting always the most important information supporting drug withdrawals for pharmacovigilance reasons in France? Pharmacoepidemiol Drug Saf. 2012 Dec;21(12):1289-94. PubMed PMID: 22961933. Epub 2012/09/11. eng.

3. Rodriguez EM, Staffa JA, Graham DJ. The role of databases in drug postmarketing surveillance. Pharmacoepidemiol Drug Saf. 2001 Aug-Sep;10(5):407-10. PubMed PMID: 11802586. Epub 2002/01/23. eng.

4. Moore N. The past, present and perhaps future of pharmacovigilance: homage to Folke Sjoqvist. European journal of clinical pharmacology. 2013 May;69 Suppl 1:33-41. PubMed PMID: 23640186. Epub 2013/05/10. eng.

5. Arrêté du 28 avril 2005 relatif aux bonnes pratiques de pharmacovigilance, (2005).

6. Begaud B, Moride Y, Tubert-Bitter P, Chaslerie A, Haramburu F. False-positives in spontaneous reporting: should we worry about them? Br J Clin Pharmacol. 1994 Nov;38(5):401-4. PubMed PMID: 7893579. Pubmed Central PMCID: 1364871. Epub 1994/11/01. eng.

7. Goldman SA. Limitations and strengths of spontaneous reports data. Clin Ther. 1998;20 Suppl C:C40-4. PubMed PMID: 9915089. Epub 1999/01/23. eng.

8. Martin-Latry K, Begaud B. Pharmacoepidemiological research using French reimbursement databases: yes we can! Pharmacoepidemiol Drug Saf. 2010 Mar;19(3):256-65. PubMed PMID: 20128015. Epub 2010/02/04. eng.

9. Schuemie MJ, Coloma PM, Straatman H, Herings RM, Trifiro G, Matthews JN, et al. Using electronic health care records for drug safety signal detection: a comparative evaluation of statistical methods. Medical care. 2012 Oct;50(10):890-7. PubMed PMID: 22929992. Epub 2012/08/30.

10. Moore N, Gulmez SE, Blin P, Lassalle R, Jove J, Theophile H, et al. Relative risks from case-population data. Epidemiology. 2013 Nov;24(6):935-6. PubMed PMID: 24076997. Epub 2013/10/01. eng.

11. Trifiro G, Pariente A, Coloma PM, Kors JA, Polimeni G, Miremont-Salame G, et al. Data mining on electronic health record databases for signal detection in pharmacovigilance: which events to monitor? Pharmacoepidemiol Drug Saf. 2009 Dec;18(12):1176-84. PubMed PMID: 19757412. Epub 2009/09/17.

12. Trifiro G, Fourrier-Reglat A, Sturkenboom MC, Diaz Acedo C, Van Der Lei J, Group E-A. The EU-ADR project: preliminary results and perspective. Studies in health technology and informatics. 2009;148:43-9. PubMed PMID: 19745234.

13. Trifiro G, Patadia V, Schuemie MJ, Coloma PM, Gini R, Herings R, et al. EU-ADR healthcare database network vs. spontaneous reporting system database: preliminary comparison of signal detection. Studies in health technology and informatics. 2011;166:25-30. PubMed PMID: 21685607. Epub 2011/06/21.

14. US Food and Drug Administration. FDA's Sentinel Initiative 2016 [cited 2016 Mar]. Available from: http://www.fda.gov/Safety/FDAsSentinelInitiative/default.htm.

15. Observational Medical Outcomes Partnership. About Us [cited 2016 Mar]. Available from: http://omop.org/.

16. DuMouchel W, Ryan PB, Schuemie MJ, Madigan D. Evaluation of disproportionality safety signaling applied to healthcare databases. Drug Saf. 2013 Oct;36 Suppl 1:S123-32. PubMed PMID: 24166229. Epub 2013/11/06. eng.

17. Madigan D, Schuemie MJ, Ryan PB. Empirical performance of the case-control method: lessons for developing a risk identification and analysis system. Drug Saf. 2013 Oct;36 Suppl 1:S73-82. PubMed PMID: 24166225. Epub 2013/11/06. eng.

18. Noren GN, Bergvall T, Ryan PB, Juhlin K, Schuemie MJ, Madigan D. Empirical performance of the calibrated self-controlled cohort analysis within temporal pattern discovery: lessons for developing a risk identification and analysis system. Drug Saf. 2013 Oct;36 Suppl 1:S107-21. PubMed PMID: 24166228. Epub 2013/11/06. eng.

19. Ryan PB, Schuemie MJ, Gruber S, Zorych I, Madigan D. Empirical performance of a new user cohort method: lessons for developing a risk identification and analysis system. Drug Saf. 2013 Oct;36 Suppl 1:S59-72. PubMed PMID: 24166224. Epub 2013/11/06. eng.

20. Ryan PB, Schuemie MJ, Madigan D. Empirical performance of a self-controlled cohort method: lessons for developing a risk identification and analysis system. Drug Saf. 2013 Oct;36 Suppl 1:S95-106. PubMed PMID: 24166227. Epub 2013/11/06. eng.

21. Schuemie MJ, Madigan D, Ryan PB. Empirical performance of LGPS and LEOPARD: lessons for developing a risk identification and analysis system. Drug Saf. 2013 Oct;36 Suppl 1:S133-42. PubMed PMID: 24166230. Epub 2013/11/06. eng.

22. Suchard MA, Zorych I, Simpson SE, Schuemie MJ, Ryan PB, Madigan D. Empirical performance of the self-controlled case series design: lessons for developing a risk identification and analysis system. Drug Saf. 2013 Oct;36 Suppl 1:S83-93. PubMed PMID: 24166226. Epub 2013/11/06. eng.

23. Ryan PB, Stang PE, Overhage JM, Suchard MA, Hartzema AG, DuMouchel W, et al. A comparison of the empirical performance of methods for a risk identification system. Drug Saf. 2013 Oct;36 Suppl 1:S143-58. PubMed PMID: 24166231. Epub 2013/11/06. eng.

24. Schuemie MJ, Gini R, Coloma PM, Straatman H, Herings RM, Pedersen L, et al. Replication of the OMOP experiment in Europe: evaluating methods for risk identification in electronic health record databases. Drug Saf. 2013 Oct;36 Suppl 1:S159-69. PubMed PMID: 24166232. Epub 2013/11/06. eng.

25. Ryan PB, Schuemie MJ. Evaluating performance of risk identification methods through a large-scale simulation of observational data. Drug Saf. 2013 Oct;36 Suppl 1:S171-80. PubMed PMID: 24166233. Epub 2013/11/06. eng.

26. Gulmez SE, Larrey D, Pageaux GP, Lignot S, Lassalle R, Jove J, et al. Transplantation for acute liver failure in patients exposed to NSAIDs or paracetamol (acetaminophen): the multinational case-population SALT study. Drug Saf. 2013 Feb;36(2):135-44. PubMed PMID: 23325533. Pubmed Central PMCID: 3568201. Epub 2013/01/18. eng.

27. Moore N, Gulmez SE, Larrey D, Pageaux GP, Lignot S, Lassalle R, et al. Choice of the denominator in case population studies: event rates for registration for liver transplantation after exposure to NSAIDs in the SALT study in France. Pharmacoepidemiol Drug Saf. 2013 Feb;22(2):160-7. PubMed PMID: 23166053. Epub 2012/11/21. eng.

28. Theophile H, Laporte JR, Moore N, Martin KL, Begaud B. The case-population study design: an analysis of its application in pharmacovigilance. Drug Saf. 2011 Oct 1;34(10):861-8. PubMed PMID: 21879780. Epub 2011/09/02. eng.

29. Capella D, Laporte JR, Vidal X, Wiholm BE, Begaud B, Langman MJ, et al. European network for the case-population surveillance of rare diseases (Euronet). A prospective feasibility study. European journal of clinical pharmacology. 1998 Jan;53(5):299-302. PubMed PMID: 9516026. Epub 1998/03/27. eng.

30. Capella D, Pedros C, Vidal X, Laporte JR. Case-population studies in pharmacoepidemiology. Drug Saf. 2002;25(1):7-19. PubMed PMID: 11820909. Epub 2002/02/01. eng.

31. Gulmez SE, Larrey D, Pageaux GP, Lignot-Maleyran S, de Vries C, Sturkenboom M, et al. Methodology for a multinational case-population study on liver toxicity risks with NSAIDs: the Study of Acute Liver Transplant (SALT). European journal of clinical pharmacology. 2013 Mar;69(3):605-16. PubMed PMID: 22878690. Epub 2012/08/11. eng.

32. Ibanez L, Ballarin E, Vidal X, Laporte JR. Agranulocytosis associated with calcium dobesilate clinical course and risk estimation with the case-control and the case-population approaches. European journal of clinical pharmacology. 2000 Dec;56(9-10):763-7. PubMed PMID: 11214789. Epub 2001/02/24. eng.

33. Keisu M, Ekman E, Wiholm BE. Comparing risk estimates of sulphonamide-induced agranulocytosis from the Swedish Drug Monitoring System and a case-control study. European journal of clinical pharmacology. 1992;43(3):211-4. PubMed PMID: 1358621. Epub 1992/01/01. eng.

34. Laporte JR, Ibanez L, Ballarin E, Perez E, Vidal X. Fatal aplastic anaemia associated with nifedipine. Lancet. 1998 Aug 22;352(9128):619-20. PubMed PMID: 9746025. Epub 1998/09/24. eng.

35. Laporte JR, Vidal X, Ballarin E, Ibanez L. Possible association between ocular chloramphenicol and aplastic anaemia--the absolute risk is very low. Br J Clin Pharmacol. 1998 Aug;46(2):181-4. PubMed PMID: 9723830. Pubmed Central PMCID: 1873671. Epub 1998/09/02. eng.

36. Moore ND, Gulmez E, Blin P. Is paracetamol hepatotoxic at normal doses? Bmj. 2013;346:f1519. PubMed PMID: 23482978. Epub 2013/03/14. eng.

37. de Jonge L, Zetstra-van der Woude PA, Bos HJ, de Jong-van den Berg LT, Bakker MK. Identifying associations between maternal medication use and birth defects using a case-population approach: an exploratory study on signal detection. Drug Saf. 2013 Nov;36(11):1069-78. PubMed PMID: 23828658. Epub 2013/07/06. eng.

38. Bezin J, Pariente A, Lassalle R, Dureau-Pournin C, Abouelfath A, Robinson P, et al. Use of the recommended drug combination for secondary prevention after a first occurrence of acute coronary syndrome in France. European journal of clinical pharmacology. 2013 Nov 24. PubMed PMID: 24271648. Epub 2013/11/26. Eng.

39. Duong M, Salvo F, Pariente A, Abouelfath A, Lassalle R, Droz C, et al. Usage patterns of 'over-the-counter' vs. prescription-strength nonsteroidal anti-inflammatory drugs in France. Br J Clin Pharmacol. 2014 May;77(5):887-95. PubMed PMID: 24102791.

40. Reich CG, Ryan PB, Schuemie MJ. Alternative outcome definitions and their effect on the performance of methods for observational outcome studies. Drug Saf. 2013 Oct;36 Suppl 1:S181-93. PubMed PMID: 24166234. Epub 2013/11/06. eng.

41. Blin P, Dureau-Pournin C, Lassalle R, Jove J, Thomas-Delecourt F, Droz-Perroteau C, et al. Outcomes in patients with stable coronary artery disease post-myocardial infarction: the HORUS cohort study in the French health insurance and hospital-discharge database,. 31st ICPE; August 23-26; Boston, MA, USA2015.

42. Blin P, Dureau-Pournin C, Lassalle R, Abouelfath A, Droz-Perroteau C, Moore N. A population database study of outcomes associated with vitamin K antagonists in atrial fibrillation before DOAC. Br J Clin Pharmacol. 2016 Mar;81(3):569-78. Epub 2015 Dec 28.

43. Danan G, Benichou C. Causality assessment of adverse reactions to drugs--I. A novel method based on the conclusions of international consensus meetings: application to drug-induced liver injuries. Journal of clinical epidemiology. 1993 Nov;46(11):1323-30. PubMed PMID: 8229110. Epub 1993/11/01. eng.

44. Benichou C. Criteria of drug-induced liver disorders. Report of an international consensus meeting. J Hepatol. 1990 Sep;11(2):272-6. PubMed PMID: 2254635. Epub 1990/09/01. eng.

45. Benichou C, Danan G, Flahault A. Causality assessment of adverse reactions to drugs--II. An original model for validation of drug causality assessment methods: case reports with positive rechallenge. Journal of clinical epidemiology. 1993 Nov;46(11):1331-6. PubMed PMID: 8229111. Epub 1993/11/01. eng.

46. Caisse nationale de l'assurance maladie des travailleurs salariés. Sniiram 2015. Available from: http://www.ameli.fr/l-assurance-maladie/statistiques-et-publications/sniiram/finalites-du-sniiram.php.

47. Moulis G, Lapeyre-Mestre M, Palmaro A, Pugnet G, Montastruc J, Sailler L. French health insurance databases: What interest for medical research? Rev Med Interne. 2015 Jun;36(6):411-7.

48. Tuppin P, de Roquefeuil L, Weill A, Ricordeau P, Merliere Y. French national health insurance information system and the permanent beneficiaries sample. Rev epidemiol sante publique. 2010;58:286-90.

49. Pierfitte C, Macouillard G, Thicoipe M, Chaslerie A, Pehourcq F, Aissou M, et al. Benzodiazepines and hip fractures in elderly people: case-control study. Bmj. 2001 March 24, 2001;322(7288):704-8.

50. Farrington C. Relative incidence estimation from case series for vaccine safety evaluation. Biometrics. 1995 Mar;51(1):228-35. PubMed PMID: 7766778. Epub 1995/03/01. eng.

51. Whitaker H, Farrington C, Spiessens B, Musonda P. Tutorail in biostatistics: The self-controlled case series method. Stat Med. 2006 May;25(10):1768-97.

52. Abenhaim L, Moore N, Begaud B. The role of pharmacoepidemiology in pharmacovigilance: A conference at the 6th ESOP meeting, Budapest, September 28, 1998. Pharmacoepidemiol Drug Saf. 1999;8:S1-S7.

APPENDIX 1 - CASE-POPULATION RATIO (CPR)

Event rates in SALT* and population exposure rates in EGB**, case-population ratio (CPR) with confidence intervals for the association of selected medication with acute liver failure leading to transplantation from Moore *et al.* (10):

	Cases (n)	Exposed population (n)	Exposure in cases (%)	Exposure in population (%)	Case-population ratio (95%CI)
All	72	351792	100	100	1
M01A NSAIDs	10	201195	13.89	57.19	0.24 (0.13-0.47)
Ibuprofen (M01AE01)	4	98602	5.56	28.03	0.20 (0.07-0.54)
Diclofenac (M01AB05)	2	37622	2.78	10.69	0.26 (0.06-1.06)
Ketoprofen (M01AE03)	3	61731	4.17	17.55	0.24 (0.07-0.75)
Nimesulide (M01AX17)	1	20573	1.39	5.85	0.24 (0.03-1.71)
Paracetamol	49	251569	68.06	71.51	0.95 (0.66-1.37
N05B Anxiolytic drugs	13	89157	18.06	25.34	0.71 (0.39-1.29)
N03A Antiepileptic drugs	11	20318	15.28	5.78	2.65 (1.40-4.99)
Clonazepam (N03AE01)	3	11778	4.17	3.35	1.24 (0.39-3.95)
Valproic acid (N03AG01)	2	3122	2.78	0.89	3.13 (0.77-12.8)
Carbamazepine (N03AF01)	1	1657	1.39	0.47	2.95 (0.41-21.2)
Levetiracetam (N03AX14)	1	321	1.39	0.09	15.2 (2.12-109.5)
Phenobarbital (N03AA02)	1	828	1.39	0.24	5.90 (0.82-42.4)
Phenytoin (N03AB02)	1	135	1.39	0.04	36.2 (5.03-260)
Pregabalin (N03AX16)	1	2715	1.39	0.77	1.80 (0.25-12.9)
Valpromide (N03AG02)	1	1436	1.39	0.41	3.40 (0.47-24.5)
R06A Antihistamines	8	117867	11.11	33.50	0.33 (0.16-0.69)
Alimemazine (R06AD01)	2	2632	2.78	0.75	3.71 (0.91-15.1)
Desloratadine (R06AX27)	2	44396	2.78	12.62	0.22 (0.05-0.90)
Oxomemazine (R06AD08)	2	40022	2.78	11.38	0.24 (0.06-1.00)
A02B Gastric Antisecretory drugs	7	112384	9.72	31.95	0.30 0.14-0.69)
N06A Antidepressant drugs	6	50636	8.33	14.39	0.58 (0.25-1.33)

* SALT: Study of Acute Liver Transplantation, French sample,

** EGB: Echantillon Généraliste de Bénéficiaires, a 1/97 representative sample of French national health-care database covering 85% of the population.

APPENDIX 2 - CASE-POPULATION RATIOS (CPR) FOR VARIOUS VALUES OF RELATIVE RISK (RR) AND POPULATION EXPOSURE (PEXP)

Pexp	0.00001	0.0001	0.001	0.01	0.02	0.05	0.1	0.2	0.3
RR									
0.2	0.200	0.200	0.200	0.202	0.203	0.208	0.218	0.240	0.269
0.4	0.400	0.400	0.400	0.402	0.405	0.413	0.427	0.460	0.503
0.6	0.600	0.600	0.600	0.602	0.605	0.613	0.627	0.660	0.703
0.8	0.800	0.800	0.800	0.802	0.803	0.808	0.818	0.840	0.869
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1.2	1.200	1.200	1.200	1.198	1.195	1.187	1.173	1.140	1.097
1.4	1.400	1.400	1.399	1.394	1.389	1.371	1.338	1.260	1.160
1.6	1.600	1.600	1.599	1.590	1.580	1.549	1.493	1.360	1.189
1.8	1.800	1.800	1.799	1.785	1.771	1.724	1.640	1.440	1.183
2	2.000	2.000	1.998	1.980	1.959	1.895	1.778	1.500	1.143
2.2	2.200	2.200	2.197	2.173	2.146	2.061	1.907	1.540	1.069
2.4	2.400	2.400	2.397	2.366	2.331	2.223	2.027	1.560	0.960
2.6	2.600	2.600	2.596	2.558	2.515	2.381	2.138	1.560	0.817
2.8	2.800	2.799	2.795	2.749	2.697	2.535	2.240	1.540	0.640
3	3.000	2.999	2.994	2.939	2.878	2.684	2.333	1.500	0.429
5	5.000	4.998	4.980	4.798	4.592	3.947	2.778	/	/
10	9.999	9.991	9.910	9.091	8.163	5.263	/		
15	14.998	14.979	14.790	12.879	10.714	/			
20	19.996	19.962	19.620	16.162	12.245				
25	24.994	24.940	24.399	18.939	12.755				

From Moore et *al.* (10).

APPENDIX 3 - ALCAPONE EVENT OF INTEREST CODES

ALCAPONE-ALI

ICD10 codes for acute liver injury

ICD10	Definition
K71.1	Toxic liver disease with hepatic necrosis
K71.2	Toxic liver disease with acute hepatitis
K71.6	Toxic liver disease with hepatitis, not elsewhere classified
K71.9	Toxic liver disease, unspecified
K72.0	Acute and subacute hepatic failure

K71.6, Toxic liver disease with hepatitis, not elsewhere classified, and K71.9, Toxic liver disease, unspecified, will be included in the sensitivity analysis.

ICD10 exclusion codes for acute liver injury

ICD10	Definition
B18	Chronic viral hepatitis
C	Malignant neoplasms
F10	Mental and behavioural disorders due to use of alcohol
G31	Other degenerative diseases of nervous system, not elsewhere classified
I50	Heart failure
I81	Portal vein thrombosis
I85	Oesophageal varices
K70	Alcoholic liver disease
K74	Fibrosis and cirrhosis of liver
K76	Other diseases of liver
K80	Cholelithiasis
K83	Other diseases of biliary tract
R18	Ascites
Z95	Presence of cardiac and vascular implants and grafts

ALCAPONE-MI

ICD10 codes for acute myocardial infarction

ICD10	Definition
I21.0	Acute transmural myocardial infarction of anterior wall
I21.1	Acute transmural myocardial infarction of inferior wall
I21.2	Acute transmural myocardial infarction of other sites
I21.3	Acute transmural myocardial infarction of unspecified site
I21.4	Acute subendocardial myocardial infarction
I21.9	Acute myocardial infarction, unspecified
I20.0	Unstable angina

I20.0, Unstable angina will be included in the sensitivity analysis.

ALCAPONE-KI

ICD10 codes for acute kidney injury

ICD10	Definition
N17.0	Acute kidney failure with tubular necrosis
N17.1	Acute kidney failure with acute cortical necrosis
N17.2	Acute kidney failure with medullary necrosis
N17.8	Other acute kidney failure
N17.9	Acute kidney failure, unspecified

ALCAPONE-UGIB

ICD10 codes for upper gastrointestinal bleeding

ICD10	Definition
I85.0	Oesophageal varices with bleeding
I98.3	Oesophageal varices with bleeding in diseases classified elsewhere
K22.6	Gastro-oesophageal laceration-haemorrhage syndrome
K25.0	Gastric ulcer; Acute with haemorrhage
K25.2	Gastric ulcer; Acute with both haemorrhage and perforation
K25.4	Gastric ulcer; Chronic or unspecified with haemorrhage
K25.6	Gastric ulcer; Chronic or unspecified with both haemorrhage and perforation
K26.0	Duodenal ulcer; Acute with haemorrhage
K26.2	Duodenal ulcer; Acute with both haemorrhage and perforation
K26.4	Duodenal ulcer; Chronic or unspecified with haemorrhage
K26.6	Duodenal ulcer; Chronic or unspecified with both haemorrhage and perforation
K27.0	Peptic ulcer, site unspecified; Acute with haemorrhage
K27.2	Peptic ulcer, site unspecified; Acute with both haemorrhage and perforation
K27.4	Peptic ulcer, site unspecified; Chronic or unspecified with haemorrhage
K27.6	Peptic ulcer, site unspecified; Chronic or unspecified with both haemorrhage and perforation
K28.0	Gastrojejunal ulcer; Acute with haemorrhage
K28.2	Gastrojejunal ulcer; Acute with both haemorrhage and perforation
K28.4	Gastrojejunal ulcer; Chronic or unspecified with haemorrhage
K28.6	Gastrojejunal ulcer; Chronic or unspecified with both haemorrhage and perforation
K29.0	Acute haemorrhagic gastritis
K92.0	Haematemesis
K92.1	Melaena
K92.2	Gastrointestinal haemorrhage, unspecified

I85.0, Oesophageal varices with bleeding, I98.3, Oesophageal varices with bleeding in diseases classified elsewhere, K22.6 Gastro-oesophageal laceration-haemorrhage syndrome, and K92.2, Gastrointestinal haemorrhage, unspecified, will be included in the sensitivity analysis.

APPENDIX 4 - OMOP REFERENCE SET

Acute kidney injury

Positive controls

Acyclovir
Hydrochlorothiazide
Ibuprofen
Lisinopril Meloxicam
Naproxen
Olmesartan medoxomil
Allopurinol
Candesartan
Capreomycin
Captopril
Chlorothiazide
Cyclosporine
Diflunisal
Enalaprilat
Etodolac
Fenoprofen
Ketoprofen
Ketorolac
Mefenamate
Moexipril
Oxaprozin
Piroxicam
Telmisartan

Negative controls

Benzonatate
Ketoconazole
Loratadine
Metaxalone
Temazepam
Acarbose
Adenosine
Almotriptan
Amylases
Benzocaine
Bromfenac
Chlorambucil
Clorazepate
Clozapine
Cosyntropin
Dacarbazine
Darbepoetin alfa
Darifenacin
Darunavir
Dicyclomine
Disulfiram
Eletriptan
Endopeptidases
Entecavir
Ergotamine
Ferrous gluconate
Flavoxate
Flutamide
Frovatriptan
Gatifloxacin
Griseofulvin
Hyoscyamine
Imipramine
Infliximab
Ketotifen
Lactulose
Lipase
Mebendazole
Methenamine
Methocarbamol
Miconazole
Nelfinavir
Neostigmine
Nortriptyline
Orlistat
Paromomycin
Penicillin V
Phentermine
Phentolamine
Prilocaine
Primidone
Prochlorperazine
Ramelteon
Rizatriptan
Scopolamine
Simethicone
Sodiumphosphate, monobasic
Tetrahydrocannabinol
Thiabendazole
Thiothixene
Tinidazole
Urea
Vitamin A
Zafirlukast

Acute liver injury

Positive controls

Allopurinol
Carbamazepine
Celecoxib
Ciprofloxacin
Cyclosporine
Diltiazem
Erythromycin
Etodolac
Fluconazole
Ibuprofen
Indomethacin
Ketorolac
Lamotrigine
Levofloxacin
Lisinopril
Methotrexate
Naproxen
Niacin
Nifedipine
Nitrofurantoin
Nortriptyline
Ofloxacin
Oxaprozin
Pioglitazone
Piroxicam
Quinapril
Ramipril
Sulindac
Tamoxifen
Terbinafine
Trandolapril
Valproate
Acetazolamide
Abacavir
Alatrofloxacin
Bortezomib
Bosentan
Busulfan
Captopril
Caspofungin
Clozapine
Dacarbazine
Darunavir
Didanosine
Disulfiram
Efavirenz
Enalaprilat
Felbamate
Flutamide
Gemcitabine
Gemifloxacin
Imatinib
Infliximab
Interferon beta-1a
Isoniazid
Itraconazole
Lamivudine
Methimazole
Methyldopa
Moexipril
Nefazodone
Nevirapine
Norfloxacin
Orlistat
Penicillamine
Posaconazole
Propylthiouracil
Rifampin
Stavudine
Sulfisoxazole
Tenofovir
Thiabendazole
Thioguanine
Tipranavir
Tolcapone
Tolmetin
Trovafloxacin
Voriconazole
Zafirlukast
Zalcitabine
Zidovudine

Negative controls

Adenosine
Benzocaine
Benzonatate
Dicyclomine
Fluticasone
Gatifloxacin
Griseofulvin
Hyoscyamine
Lactulose
Miconazole
Oxybutynin
Penicillin V
Salmeterol
Scopolamine
Sitagliptin
Sucralfate
Almotriptan
Amylases
Cosyntropin
Droperidol
Endopeptidases
Ergotamine
Ferrous gluconate
Flavoxate
Ketotifen
Lipase
Lithium citrate
Methenamine
Neostigmine
Paromomycin
Phentermine
Phentolamine
Primidone
Propantheline
Sodium phosphate, monobasic
Tetrahydrocannabinol
Tinidazole

Acute myocardial infarction

Positive controls

Amlodipine	Piroxicam	Factor VIIa
Darbepoetin alfa	Sulindac	Fenoprofen
Dipyridamole	Sumatriptan	Flurbiprofen
Epoetin Alfa Estradiol	Almotriptan	Frovatriptan
Estrogens, conjugated	Amoxapine	Imipramine
Etodolac	Bromocriptine	Ketoprofen
Indomethacin	Desipramine	Moexipril
Ketorolac	Diflunisal	Naratriptan
Nabumetone	Eletriptan	Rizatriptan Salsalate
Nifedipine Nortriptyline	Enalaprilat	Tolmetin
Oxaprozin	Estropipate	Zolmitriptan

Negative controls

Acarbose	Benzonatate	Nevirapine
Acetazolamide	Clindamycin	Paromomycin
Amylases	Dicyclomine	Pemoline
Bromfenac	Fluticasone	Penicillamine
Chlorambucil	Gatifloxacin	Posaconazole
Clorazepate	Hyoscyamine	Prilocaine
Chlorothiazide	Ketoconazole	Primidone
Cosyntropin	Lactulose	Propantheline
Darifenacin	Loratadine	Simethicone
Didanosine	Metaxalone	Sodium phosphate,
Droperidol	Methocarbamol	monobasic
Endopeptidases	Penicillin V	Stavudine
Entecavir	Prochlorperazine	Sulfasalazine
Ferrous gluconate	Oxybutynin	Sulfisoxazole
Flavoxate	Ramelteon	Tetrahydrocannabinol
Flutamide	Salmeterol	Thiabendazole
Ketotifen	Scopolamine	Thiothixene
Lipase	Sitagliptin	Tinidazole
Lithium citrate	Sucralfate	Tipranavir
Mebendazole	Temazepam	Vitamin A
Methenamine	Terbinafine	Zafirlukast
Methimazole	Urea	
Miconazole	Nelfinavir	

Upper gastrointestinal bleeding

Positive controls

Citalopram	Ketorolac	Diflunisal
Clindamycin	Meloxicam	Fenoprofen
Clopidogrel	Nabumetone	Flurbiprofen
Escitalopram	Naproxen	Ketoprofen
Etodolac	Piroxicam	Mefenamate
Fluoxetine	Potassium chloride	Sulindac
Ibuprofen	Sertraline	Tolmetin
Indomethacin	Oxaprozin	Valdecoxib

Negative controls

Abacavir	Hyoscyamine	Paromomycin
Acarbose	Itraconazole	Pemoline
Adenosine	Ketotifen	Phentermine
Amylases	Ketoconazole	Phentolamine
Benzocaine	Lactulose	Prilocaine
Benzonatate	Lamivudine	Propantheline
Bromfenac	Lipase	Rosiglitazone Salmeterol
Chlorambucil	Lithium citrate	Scopolamine
Clorazepate	Loratadine	Simethicone
Cosyntropin	Metaxalone	Sitagliptin
Dacarbazine	Methocarbamol	Stavudine
Darifenacin	Mebendazole	Sucralfate
Dicyclomine	Miconazole	Temazepam
Disulfiram	Moexipril	Terbinafine
Droperidol	Neostigmine	Tetrahydrocannabinol
Endopeptidases	Nevirapine	Thiabendazole
Entecavir	Nitrofurantoin	Thiothixene
Epoetin alfa	Orlistat	Tinidazole
Ergotamine	Oxybutynin	Urea
Ferrous gluconate	Penicillin V	Vitamin A
Fluticasone	Pioglitazone	Zidovudine
Griseofulvin	Prochlorperazine	

APPENDIX 5- ENCEPP CHEKLIST FOR STUDY PROTOCOLS

European Network of Centres for Pharmacoepidemiology and Pharmacovigilance

Doc.Ref. EMA/540136/2009

ENCePP Checklist for Study Protocols (Revision 2, amended)

Adopted by the ENCePP Steering Group on 14/01/2013

The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) welcomes innovative designs and new methods of research. This Checklist has been developed by ENCePP to stimulate consideration of important principles when designing and writing a pharmacoepidemiological or pharmacovigilance study protocol. The Checklist is intended to promote the quality of such studies, not their uniformity. The user is also referred to the ENCePP Guide on Methodological Standards in Pharmacoepidemiology which reviews and gives direct electronic access to guidance for research in pharmacoepidemiology and pharmacovigilance.

For each question of the Checklist, the investigator should indicate whether or not it has been addressed in the study protocol. If the answer is "Yes", the page number(s) of the protocol where this issue has been discussed should be specified. It is possible that some questions do not apply to a particular study (for example in the case of an innovative study design). In this case, the answer 'N/A' (Not Applicable) can be checked and the "Comments" field included for each section should be used to explain why. The "Comments" field can also be used to elaborate on a "No" answer.

This Checklist should be included as an Annex by marketing authorisation holders when submitting the protocol of a non-interventional post-authorisation safety study (PASS) to a regulatory authority (see the Guidance on the format and content of the protocol of non-interventional post-authorisation safety studies). Note, the Checklist is a supporting document and does not replace the format of the protocol for PASS as recommended in the Guidance and Module VIII of the Good pharmacovigilance practices (GVP).

Study title:
Alert generation using the case-population approach in the French claims databases (ALCAPONE)

Study reference number:

Section 1: Milestones	**Yes**	**No**	**N/A**	**Page Number(s)**
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[1]	☒	☐	☐	12
1.1.2 End of data collection[2]	☒	☐	☐	12
1.1.3 Study progress report(s)	☐	☐	☒	
1.1.4 Interim progress report(s)	☐	☐	☒	
1.1.5 Registration in the EU PAS register	☒	☐	☐	3
1.1.6 Final report of study results.	☒	☐	☐	12

Comments:
1.1.3; 1.1.4 These items are not planned for this study.

[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.

ENCePP Checklist for Study Protocols (Revision 2)

Section 2: Research question	**Yes**	**No**	**N/A**	**Page Number(s)**
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)	☒	☐	☐	13-14-15
2.1.2 The objective(s) of the study?	☒	☐	☐	15
2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	15
2.1.4 Which formal hypothesis(-es) is (are) to be tested?	☒	☐	☐	15-16
2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☐	☒	

Comments:

2.1.5 There is a hypothesis: Positive controls sould be associated with events of interest; negative controls should not.

Section 3: Study design	**Yes**	**No**	**N/A**	**Page Number(s)**
3.1 Is the study design described? (e.g. cohort, case-control, randomised controlled trial, new or alternative design)	☒	☐	☐	15-16-17
3.2 Does the protocol specify the primary and secondary (if applicable) endpoint(s) to be investigated?	☒	☐	☐	19
3.3 Does the protocol describe the measure(s) of effect? (e.g. relative risk, odds ratio, deaths per 1000 person-years, absolute risk, excess risk, incidence rate ratio, hazard ratio, number needed to harm (NNH) per year)	☒	☐	☐	23-24-25

Comments:

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Page Number(s)**
4.1 Is the source population described?	☒	☐	☐	18
4.2 Is the planned study population defined in terms of:				
4.2.1 Study time period?	☒	☐	☐	16-17-18
4.2.2 Age and sex?	☒	☐	☐	18
4.2.3 Country of origin?	☒	☐	☐	18
4.2.4 Disease/indication?	☒	☐	☐	19-20
4.2.5 Co-morbidity?	☐	☒	☐	
4.2.6 Seasonality?	☐	☒	☐	
4.3 Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	19-20

Comments:

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Page Number(s)**
5.1 Does the protocol describe how exposure is defined and measured? (e.g. operational details for defining and categorising exposure)	☒	☐	☐	20-21
5.2 Does the protocol discuss the validity of exposure				

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Page Number(s)**
measurement? (e.g. precision, accuracy, prospective ascertainment, exposure information recorded before the outcome occurred, use of validation sub-study)	☒	☐	☐	20-21-26
5.3 Is exposure classified according to time windows? (e.g. current user, former user, non-use)	☒	☐	☐	20-21
5.4 Is exposure classified based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☐	☐	☒	
5.5 Does the protocol specify whether a dose-dependent or duration-dependent response is measured?	☐	☐	☒	

Comments:

5.4; 5.5 Drug exposure will be assessed using exhaustive non-hospital drug claims.

Section 6: Endpoint definition and measurement	**Yes**	**No**	**N/A**	**Page Number(s)**
6.1 Does the protocol describe how the endpoints are defined and measured?	☒	☐	☐	19-20
6.2 Does the protocol discuss the validity of endpoint measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, prospective or retrospective ascertainment, use of validation sub-study)	☒	☐	☐	22-26-27

Comments:

Section 7: Confounders and effect modifiers	**Yes**	**No**	**N/A**	**Page Number(s)**
7.1 Does the protocol address known confounders? (e.g. collection of data on known confounders, methods of controlling for known confounders)	☒	☐	☐	26-27
7.2 Does the protocol address known effect modifiers? (e.g. collection of data on known effect modifiers, anticipated direction of effect)	☐	☐	☒	

Comments:

7.2 No effect modifiers known.

Section 8: Data sources	**Yes**	**No**	**N/A**	**Page Number(s)**
8.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				
8.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview, etc.)	☒	☐	☐	21-22
8.1.2 Endpoints? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics, etc.)	☒	☐	☐	19 to 22
8.1.3 Covariates?	☒	☐	☐	21-22
8.2 Does the protocol describe the information available from the data source(s) on:				
8.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	21-22
8.2.2 Endpoints? (e.g. date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	21-22
8.2.3 Covariates? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, life style, etc.)	☒	☐	☐	21-22

Section 8: Data sources	**Yes**	**No**	**N/A**	**Page Number(s)**
8.3 Is a coding system described for:				
8.3.1 Diseases? (e.g. International Classification of Diseases (ICD)-10)	☒	☐	☐	19-21-22
8.3.2 Endpoints? (e.g. Medical Dictionary for Regulatory Activities (MedDRA) for adverse events)	☒	☐	☐	21-22
8.3.3 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC)Classification System)	☒	☐	☐	20-21-22
8.4 Is the linkage method between data sources described? (e.g. based on a unique identifier or other)	☒	☐	☐	21-22

Comments:

Section 9: Study size and power	**Yes**	**No**	**N/A**	**Page Number(s)**
9.1 Is sample size and/or statistical power calculated?	☒	☐	☐	23

Comments:

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Page Number(s)**
10.1 Does the plan include measurement of excess risks?	☐	☒	☐	
10.2 Is the choice of statistical techniques described?	☒	☐	☐	13-14-16
10.3 Are descriptive analyses included?	☒	☐	☐	23-24
10.4 Are stratified analyses included?	☐	☒	☐	
10.5 Does the plan describe methods for adjusting for confounding?	☒	☐	☐	23-24
10.6 Does the plan describe methods addressing effect modification?	☐	☐	☒	

Comments:

10.6 See item 7.2

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Page Number(s)**
11.1 Is information provided on the management of missing data?	☐	☐	☒	
11.2 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	☒	☐	☐	23
11.3 Are methods of quality assurance described?	☒	☐	☐	25
11.4 Does the protocol describe possible quality issues related to the data source(s)?	☒	☐	☐	25-26
11.5 Is there a system in place for independent review of study results?	☒	☐	☐	28

Comments:

11.1 The SNIIRAM (data source) contains exhaustive information about reimbursmed treatments out of hospital and use reimbursed healthcare resources.

Section 12: Limitations	**Yes**	**No**	**N/A**	**Page Number(s)**
12.1 Does the protocol discuss:				
12.1.1 Selection biases?	☒	☐	☐	26
12.1.2 Information biases? (e.g. anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods)	☒	☐	☐	26
12.2 Does the protocol discuss study feasibility? (e.g. sample size, anticipated exposure, duration of follow-up in a cohort study, patient recruitment)	☒	☐	☐	19-21-22
12.3 Does the protocol address other limitations?	☒	☐	☐	26-27

Comments:

Section 13: Ethical issues	**Yes**	**No**	**N/A**	**Page Number(s)**
13.1 Have requirements of Ethics Committee/Institutional Review Board approval been described?	☒	☐	☐	1-28
13.2 Has any outcome of an ethical review procedure been addressed?	☐	☒	☐	
13.3 Have data protection requirements been described?	☒	☐	☐	28

Comments:

Section 14: Amendments and deviations	**Yes**	**No**	**N/A**	**Page Number(s)**
14.1 Does the protocol include a section to document future amendments and deviations?	☒	☐	☐	12

Comments:

Section 15: Plans for communication of study results	**Yes**	**No**	**N/A**	**Page Number(s)**
15.1 Are plans described for communicating study results (e.g. to regulatory authorities)?	☒	☐	☐	28
15.2 Are plans described for disseminating study results externally, including publication?	☒	☐	☐	28

Comments:

Name of the main author of the protocol: Nicholas Moore

Date: 24/3/2016

Signature: _____

ENCePP Checklist for Study Protocols (Revision 2) 5

APPENDIX 6 - SIGNATURE PAGES