



HAL
open science

Empirical assessment of case-based designs for drug safety alert generation in the French National Healthcare System database (SNDS)

Nicolas H. Thurin

► **To cite this version:**

Nicolas H. Thurin. Empirical assessment of case-based designs for drug safety alert generation in the French National Healthcare System database (SNDS). Human health and pathology. Université de Bordeaux, 2019. English. NNT: 2019BORD0408 . tel-03008822

HAL Id: tel-03008822

<https://theses.hal.science/tel-03008822>

Submitted on 17 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE PRÉSENTÉE
POUR OBTENIR LE GRADE DE

**DOCTEUR DE
L'UNIVERSITÉ DE BORDEAUX**

École Doctorale, Sociétés, Politique, Santé Publique
Pharmacologie, option Pharmaco-épidémiologie, Pharmacovigilance

Par Nicolas H. THURIN

**Évaluation empirique d'approches basées sur les cas pour
la génération d'alertes de pharmacovigilance à partir du
Système National des Données de Santé (SNDS)**

**Empirical assessment of case-based designs for drug safety alert generation in
the French National Healthcare System database (SNDS)**

Sous la direction de
Nicholas MOORE et Martijn SCHUEMIE

Soutenue le 17 décembre 2019

Membres du jury :

M. BEGAUD, Bernard	Professeur, Université de Bordeaux	Président
M. HERINGS, Ron	Professor, Amsterdam UMC	Rapporteur
M. SCHNEEWEISS, Sebastian	Professor, Harvard Medical School	Rapporteur
M. TZOURIO, Christophe	Professeur, Université de Bordeaux	Examinateur
M. RYAN, Patrick	Assistant Professor, Columbia University	Examinateur
M. MOORE, Nicholas	Professeur, Université de Bordeaux	Directeur
M. SCHUEMIE, Martijn	Senior Director, Janssen R&D	Co-directeur

ACKNOWLEDGMENTS

À Monsieur le Professeur Nicholas Moore,

Quelques fois le destin tient à peu de choses... Où serais-je aujourd'hui si je n'avais pas croisé votre chemin en cette après-midi du mois d'octobre 2015, dans le hall du Clarion Congress Hotel, à Prague ! Je ne vous remercierai jamais assez de m'avoir accordé votre confiance pour mener ce projet à bien. Merci pour votre constante bonne humeur, votre bienveillance et vos idées aussi pertinentes que profuses. Tout au long de ces quatre années vous n'avez eu de cesse de m'encourager et de me pousser à aller jusqu'au bout de mes idées. *"They did not know it was impossible so they did it"* ; Mark Twain l'a dit le premier mais cette phrase aurait pu être de vous.

À Monsieur le Docteur Martijn Schuemie,

Dear Martijn, I cannot express how grateful I am for having accepted to co-supervise this work. From inception to completion and in spite of the geographical distance, you have always been present to help us overcome scientific and technical issues; and God knows how numerous they were. Many thanks for your dedication.

*À Messieurs les Docteurs et Professeurs Bernard Bégaud, Ron Herings,
Sebastian Schneeweiss, Christophe Tzourio et Patrick Ryan*

Thank you for being part of the very few who will ever read these lines. You have accepted to spend some of your precious time to judge this work, and for this I am very much obliged. It is a great honor to have you as part of this Jury.

*À Messieurs les Docteurs et Professeurs Joshua Gagne, Jeremy Rassen,
Jacques Benichou et Alain Weill,*

The scientific quality and achievement of this project owe much to your invaluable advice and your careful reviews. Once more, thank you for everything.

À Monsieur le Professeur Adrián Puello Guerrero,

Mis primeros pasos en el hermoso mundo de la investigación y de la epidemiología fueron contigo. Me enseñaste a confiar en mí y a superar retos en un entorno desconocido. Por eso quería agradecerte sinceramente.

À Madame le Docteur Françoise Haramburu,

Vos conseils m'ont permis de trouver ma voie et aujourd'hui de m'épanouir pleinement dans mon domaine d'activité. C'est pourquoi je tenais à vous témoigner toute ma gratitude.

À Madame le Professeur Yola Moride,

Pendant plus de 7 mois vous m'avez accueilli dans votre équipe. Malgré un planning surchargé, vous avez pris le temps de me faire partager vos valeurs : la rigueur scientifique et la recherche de l'excellence qui sont aujourd'hui mon pain quotidien. Pour cela, je vous témoigne toute ma reconnaissance.

À Madame le Docteur Cécile Droz-Perroteau,

Même si je pense rester à jamais « le petit Nicolas », ces dernières années m'ont énormément apporté, et tu y es pour beaucoup. Merci de ton accueil, ton écoute, ta confiance et ton diagnostic tardif de dyslexie. Ton goût du challenge et ta bienveillance font du quotidien à la BPE une aventure humaine et professionnelle sans commune mesure. "L'impossible c'est la routine, l'amusant c'est l'impensable." Un bel adage, qui, je l'espère, nous mènera encore vers de beaux défis à relever.

À Monsieur le Docteur Patrick Blin et Monsieur Régis Lassalle,

Aujourd'hui encore vous prenez le temps de me rassurer, d'écouter avec patience mes questions plus ou moins pertinentes et surtout d'y répondre. Une vraie prouesse quand on mesure la charge de travail que vous abattez. Ce projet et moi-même vous devons beaucoup.

À Mademoiselle Marine Pénichon,

Marine, ta patience et ta persévérance sont les fils conducteurs de ces travaux. Merci pour la mise en forme de toutes ces belles idées.

À Monsieur le Docteur Patrick Bouex,

Merci pour toutes ces heures passées à tenter de résoudre les nombreux mystères informatiques que nous avons rencontrés. Ce projet a été un défi scientifique et technique duquel nous ne serions jamais venus à bout sans ton aide.

À toute l'équipe de la BPE,

Merci pour votre accueil, votre soutien et votre bonne humeur au quotidien. Votre enthousiasme et votre professionnalisme savent venir à bout de tous les défis.

À Monsieur le futur Docteur Iman Boudraa,

Merci pour ta patience, ta bienveillance, et ton inaltérable joie de vivre desquels même ces centaines de pages n'ont pu venir à bout.

À mes parents, grands-parents, mon frère et ma sœur,

C'est donc la deuxième fois que je vous adresse mes remerciements dans les premières pages d'un ouvrage. Il y a cinq ans, je vous remerciai pour m'avoir entouré, soutenu, guidé et laissé libre de mes choix. Cela reste encore vrai aujourd'hui. Mais avec ces quelques années en plus, j'ajouterais : merci de supporter mes sautes d'humeur, mes taquineries, mes remarques parfois un peu incisives et mon côté peu expansif. Faute de le dire, je vais l'écrire : je vous aime.

RESUME

La France possède une large base de données nationale regroupant les données de liquidation de l'Assurance Maladie, de mortalité et des données hospitalières : le Système National des Données de Santé (SNDS). Celui-ci couvre actuellement la quasi-totalité de la population française de la naissance (ou immigration), au décès (ou émigration), en incluant tous les remboursements de frais médicaux ou paramédicaux. En recueillant de manière systématique et prospective les dispensations médicamenteuses, les événements hospitaliers et les décès, le SNDS est doté d'un fort potentiel pour l'évaluation du médicament en vie réelle. Suite au retrait mondial du rofecoxib en 2004, de nombreuses initiatives visant au développement et à l'évaluation de méthodologies adaptées aux bases de données populationnelles pour la surveillance des risques liés à l'usage du médicament ont vu le jour, en particulier le réseau EU-ADR en Europe (*Exploring and Understanding Adverse Drug Reactions by integrative mining of clinical records and biomedical knowledge*) et OMOP (*Observational Outcomes Partnership*) aux États-Unis. Ces travaux ont démontré l'utilité des approches pharmaco-épidémiologiques pour la détection de signaux de pharmacovigilance. Cependant, le SNDS n'a jamais été testé dans cette optique.

L'objectif de cette thèse était d'évaluer de manière empirique, 3 approches pharmaco-épidémiologiques basées sur les cas pour la génération d'alerte(s) de pharmacovigilance dans le SNDS : étude cas-population, étude cas-témoins et séries de cas autocontrôlés. Ces approches ont été appliquées à deux événements médicaux d'intérêt récurrents en pharmacovigilance : l'hémorragie digestive haute (UGIB) et l'hépatite aigue (ALI).

Le projet a été composé de 4 principales étapes : (1) le formatage des données selon les spécifications du modèle commun de données d'OMOP et la sélection des médicaments témoins positifs et négatifs pour chaque événement d'intérêt ; (2) l'analyse des médicaments témoins sélectionnés en utilisant les 3 approches basées sur les cas, en déclinant chaque approche selon plusieurs variantes (par exemple, en testant différentes fenêtres de risque, stratégies d'ajustement, etc.) ; (3) la comparaison des performances des variantes selon leur aire sous la courbe ROC (AUC), leur erreur quadratique moyenne (MSE) et leur probabilité de couverture ;

(4) la sélection de la meilleure variante pour chaque événement d'intérêt et son étalonnage.

Sur les 3 approches étudiées, c'est la série de cas autocontrôlés qui a montré les meilleures performances dans UGIB et ALI avec des AUC respectifs de 0,80 et 0,94 et des MSE de 0,07 et 0,12. Pour UGIB, les performances optimales ont été observées lorsque l'ajustement tenait compte des traitements concomitants et lorsque les 30 premiers jours d'exposition au médicament d'intérêt étaient utilisés comme fenêtre de risque. Pour ALI, les performances optimales ont été également obtenues lors de l'ajustement en fonction des traitements concomitants, mais en utilisant une fenêtre de risque correspondant à l'ensemble de la période couverte par les dispensations de médicament d'intérêt. L'utilisation de médicaments témoins négatifs a montré que l'erreur systématique résultant de l'application de l'approche et des paramètres optimaux dans le SNDS semblait faible, mais que les biais protopathiques et de confusion restaient présents.

Au total, ces travaux ont montré que les séries de cas autocontrôlées sont à considérer comme une approche adaptée à la détection d'alertes de pharmacovigilance associées à ALI et à UGIB dans le SNDS. Un point de vue clinique demeure toutefois nécessaire pour écarter tout risque de faux positif résultant de potentiels biais résiduels. L'application d'une telle approche à d'autres événements d'intérêt et son utilisation en routine constitueraient des progrès substantiels en matière de pharmacovigilance en France.

Titre

Évaluation empirique d'approches basées sur les cas pour la génération d'alertes de pharmacovigilance à partir du Système National des Données de Santé (SNDS)

Mots-clefs

Pharmaco-épidémiologie ; SNDS ; Pharmacovigilance ; Cas-témoins ; Cas-population ; Série de cas autocontrôlés ; base de données ; Étalonnage ; Hépatite aigüe ; Hémorragie digestive haute

Intitulé et adresse du laboratoire

Bordeaux Population Health Research Center INSERM U1219, Equipe «Pharmaco-épidémiologie et impact des médicaments sur les populations »

Université de Bordeaux

Site de Carreire, Bâtiment Pharmacie, quatrième tranche, troisième étage. Case 121

146 rue Léo Saignat

33076 Bordeaux cedex, France

Laboratoire d'accueil

Bordeaux PharmacoEpi, INSERM CIC1401

Université de Bordeaux

Bâtiment du Tondu – case 41

146 rue Léo Saignat

33076 Bordeaux cedex, France

ABSTRACT

France has a large nationwide longitudinal database with claims and hospital data, the *Système National des Données de Santé* (French National healthcare database – SNDS), which currently covers almost the complete French population, from birth or immigration to death or emigration, and includes all reimbursed medical and paramedical encounters. Since SNDS systematically and prospectively captures drug dispensings, death and events leading to hospital stays, it has a strong potential for drug assessment in real life settings. Following the worldwide withdrawal of rofecoxib in 2004, several initiatives aiming to develop and evaluate methodologies for drug safety monitoring on healthcare databases emerged. The EU-ADR alliance (Exploring and Understanding Adverse Drug Reactions by integrative mining of clinical records and biomedical knowledge) and OMOP (Observational Outcomes Partnership) were respectively launched in Europe and in the Unites-States. These experiments demonstrated the usefulness of pharmacoepidemiological approaches in drug safety signal detection. However, the SNDS had never been tested in this scope. The objective of this thesis was to empirically assess 3 case-based designs – case-population, case-control, and self-controlled case series – for drug-safety alert generation in the SNDS, taking as examples two health outcome of interest: upper-gastrointestinal bleeding (UGIB) and acute liver injury (ALI).

The overall project consisted of 4 main stages: (1) preparation of the data to fit the OMOP common data model and the selection of positive and negative drug controls for each outcome of interest; (2) analysis of the selected drug controls using the 3 case-based designs, testing several design variants (*e.g.* testing different risk windows, adjustment strategies, *etc.*); (3) comparison of design variant performances through the calculation of the area under the receiver operating characteristics curve (AUC), the mean square error (MSE) and the coverage probability; (4) the selection of the best design variant and its calibration for each health outcome of interest.

Self-controlled case series showed the best performances in both outcomes, ALI and UGIB, with AUCs reaching respectively 0.80 and 0.94 and MSEs 0.07 and 0.12. For UGIB optimal performances were observed when adjusting for multiple drugs and using a risk window corresponding to the 30 first days of exposure. For ALI, optimal performances were also observed when adjusting for multiple drugs but using a risk window corresponding to the overall period covered by drug dispensings. Negative

drug control implementation highlighted that a low systematic error seemed to be generated by the optimum variants in the SNDS but that protopathic bias and confounding by indication remained unaddressed issues.

These results showed that self-controlled case series were well suited to detect drug safety alerts associated with UGIB and ALI in the SNDS in an accurate manner. A clinical perspective remains necessary to rule out potential false positive signals from residual confounding. The application in routine of such approaches extended to other outcomes of interest could result in substantial progress in pharmacovigilance in France.

Title

Empirical assessment of case-based designs for drug safety alert generation in the French National Healthcare System database (SNDS)

Keywords

Pharmacoepidemiology; SNDS; Pharmacovigilance; Drug safety; Case-control; Case-population; Self-controlled case series; Calibration; Acute Liver Injury; Upper-gastrointestinal bleeding

RESUME SUBSTANTIEL DES TRAVAUX EN LANGUE FRANÇAISE

▪ Introduction

Les données de sécurité fournies lors de la mise sur le marché d'un médicament se basent essentiellement sur les essais cliniques réalisés lors de leur développement. Cependant, ces essais peuvent faillir à identifier certains effets indésirables. Ils manquent souvent de puissance, sont restreints à des patients spécifiques, sans comorbidités, et sont conduits selon des paramètres fixes et préalablement établis ne reflétant que très peu la pratique clinique habituelle. La surveillance du médicament après sa commercialisation à travers la pharmacovigilance et les études de pharmaco-épidémiologie demeure donc cruciale pour pallier à ces manques et identifier les effets indésirables jusque-là inconnus, en particulier ceux graves et de faible fréquence de survenue. Actuellement, la notification spontanée reste l'outil privilégié à travers le monde pour l'identification de ces effets indésirables. Bien qu'ayant fait ses preuves, cette approche ne permet cependant pas de quantifier le risque potentiellement identifié : il est impossible de déterminer quelle proportion de la population a été exposée au médicament et quelle proportion a effectivement présenté un événement indésirable.

Alors que la notification spontanée fait état de données concernant un patient à un moment donné, les grandes bases de données populationnelles regroupent au cours du temps l'ensemble des informations relatives à l'historique médical des individus constitutifs d'une population. Elles peuvent se présenter sous la forme de bases de données de remboursement, de dossiers médicaux électroniques ou encore de bases de données hospitalières. Les informations relatives aux événements et aux expositions sont collectées indépendamment, de manière systématique et prospective. Elles sont donc peu affectées par les biais de sélection, et ne sont pas sujettes aux biais de mémorisation ni à l'effet Hawthorne. En étudiant l'impact de l'utilisation des médicaments en situation réelle de soins, la pharmaco-épidémiologie rend possible l'exploitation de ces bases, notamment la caractérisation des effets du médicament au niveau populationnel, y compris ceux indésirables. En outre, la présence d'un dénominateur permet de quantifier ces effets et donc leur impact potentiel en termes de santé publique. Il en résulte que les bases de données populationnelles représentent une source précieuse de données pour l'identification et la validation des signaux de pharmacovigilance.

La France dispose d'un bel exemple de base de données populationnelle, le Système National des Données de Santé (SNDS). Il couvre 99% de la population française de la naissance (ou immigration) jusqu'au décès (ou émigration). Le SNDS est une base de données qui utilise un identifiant pseudonymisé unique pour chaîner les informations de remboursement des soins de ville (dispensations de médicaments, actes de biologie, visites médicales, *etc.*) contenu dans Système National d'Information Inter Régimes de l'Assurance Maladie (SNIIRAM), avec les données d'hospitalisation (diagnostics hospitaliers, actes médicaux, *etc.*) provenant du Programme de médicalisation des systèmes d'information (PMSI) et la Base de Causes Médicales de Décès géré par le CépiDc.

En 2004, suite au retrait mondial du rofecoxib du marché, une réflexion internationale a été menée sur la manière de mieux évaluer les bénéfices et anticiper les risques au cours du cycle de vie des médicaments. Le désir permanent d'étudier des événements et des expositions toujours plus rares a conduit les acteurs de l'évaluation des produits de santé à considérer l'intérêt des bases de données populationnelles, celles-ci permettant d'inclure toujours plus de patients. Ainsi, au cours des dernières années, différents réseaux de recherche et consortia incluant de multiples bases de données ont vu le jour à l'international. A l'initiative de la Commission Européenne, en 2007, le réseau EU-ADR, *Exploring and Understanding Adverse Drug Reactions by integrative mining of clinical records and biomedical knowledge*, a été lancé avec pour objectif de développer de nouvelles méthodologies pour la surveillance du médicament en s'appuyant sur 8 bases de données de pays membres. Aux États-Unis, entre 2009 et 2013, le consortium OMOP, *Observational Outcomes Partnership*, a conduit une évaluation de différentes méthodes pour l'identification de signaux de sécurité liés aux médicaments sur 10 bases de données. Dans la plupart de ces projets, les analyses étaient conduites à travers les bases partenaires selon un modèle de réseaux distribué (*distributed network approaches*). Cependant, l'hétérogénéité résultant de la mise en commun de ces multiples sources demeurait un problème substantiel malgré l'utilisation de modèles communs de données ou l'utilisation de méta-analyses.

En France, le SNDS permet d'étudier la population nationale, soit plus de 66,6 millions de personnes, sans problèmes d'hétérogénéité. Lors de la mise en place des réseaux EU-ADR et OMOP, les données du SNDS n'étaient pas encore disponibles.

De fait, l'intérêt du SNDS pour la génération de signaux de pharmacovigilance par de telles approches n'a pas été évalué.

En 2014, la plateforme Bordeaux PharmacoEpi, a soumis une lettre d'intention dans le cadre d'un appel à projets national sur les performances des soins de santé organisé par le Ministère de la santé français (PREPS), le projet ALCAPONE : *Alert generation using the case-population approach in the French National healthcare databases*. ALCAPONE a été approuvé (PREPS 2014, 0635) et lancé en 2016, avec pour objectif de déterminer dans quelle mesure le SNDS était un outil adapté à la génération de signaux de pharmacovigilance. L'idée principale était de s'appuyer sur la méthodologie développée par OMOP pour comparer, à partir de médicament témoins, plusieurs méthodes de génération de signaux afin de trouver la plus appropriée pour quatre événements d'intérêt : l'hépatite aigue (*acute liver injury – ALI*), l'insuffisance rénale aiguë (*acute kidney injury – AKI*), l'infarctus du myocarde (*myocardial infarction – MI*) et l'hémorragie digestive haute (*upper-gastrointestinal bleeding – UGIB*). Parce qu'elles quantifient le risque en même temps qu'elles le détectent, les approches habituellement employées en pharmaco-épidémiologie ont été retenues. Les signaux ainsi détectés étant directement validés au niveau populationnel, ils constituent de fait une « alerte ».

L'objectif de cette thèse basée sur le projet ALCAPONE, était d'évaluer de manière empirique 3 approches pharmaco-épidémiologiques basées sur les cas pour la génération d'alerte(s) de pharmacovigilance dans le SNDS : la série de cas autocontrôlées (*self-controlled case series – SCCS*), l'étude cas-témoins (*case-control – CC*) et l'étude cas-population (*case-population – CP*). Deux exemples d'application ont été plus largement développés : UGIB et ALI. Les travaux présentés vont de la rédaction du protocole et du plan d'analyses statistiques à l'interprétation des résultats et leur valorisation, en passant par la supervision des étapes de gestion des données et la conduite des analyses statistiques réalisées avec le support des équipes de la plateforme Bordeaux PharmacoEpi et les outils développés par la communauté OHDSI, *Observational Health Data Sciences and Informatics*.

▪ **Méthodologie**

Le projet se composait de 4 principales étapes présentées ci-après.

1. **Extraction** des populations d'intérêt du SNDS, **mise en forme des données** selon le modèle commun de données OMOP (OMOP CDM), et **sélection des médicaments témoins positifs et négatifs** présentant suffisamment de puissance pour pouvoir être détectés ;
2. **Analyse des médicaments témoins** sélectionnés en utilisant les 3 approches basées sur les cas, en déclinant chaque approche selon plusieurs variantes (par exemple, en testant différentes fenêtres de risque, stratégies d'ajustement, *etc.*) ;
3. **Évaluation des performances des variantes de chaque approche** selon leur aire sous la courbe ROC (AUC), leur erreur quadratique moyenne (MSE) et leur probabilité de couverture ;
4. **Sélection et étalonnage de l'approche et des paramètres optimaux** selon la variante retenue pour chaque événement d'intérêt.

- *Mise en place et sélection des médicaments témoins*

La loi française rend disponible les extractions du SNDS uniquement à des fins de recherche. Afin de limiter le nombre d'extractions à conduire (une extraction par événement d'intérêt), seules les approches basées sur les cas ont été retenues : SCCS, CC et CP.

Les patients ayant présentés un des événements d'intérêt – ALI, AKI, MI et UGIB – entre le 01/01/2009 et le 31/12/2014 ont été extraits du SNDS à partir des codes diagnostics présents dans les résumés de sortie d'hospitalisation. Pour chaque événement une définition spécifique et une définition sensible ont été utilisées. Les 612 tables extraites ont été synthétisées en 14 tables transitoires puis formatées selon le OMOP CDM.

A partir des médicaments utilisés comme référence dans les projets OMOP et EU-ADR et commercialisés en France, un set de médicaments témoins positifs et négatifs a été définis pour chacun des événements d'intérêt. Les témoins positifs correspondaient à des médicaments présentant une association connue avec l'événement étudié. Les témoins négatifs correspondaient à des médicaments pour lesquels il n'existait aucune preuve permettant d'établir une relation de cause à effet

avec l'événement. En fonction du nombre de patients présents dans chaque population d'intérêt, seuls les médicaments témoins ayant suffisamment de puissance pour être détectés au sein du SNDS ont été conservés (*i.e.* ceux avec un risque relatif minimum détectable $< 1,30$).

- *Analyse des médicaments témoins*

Trois approches basées sur les cas ont été utilisées : SCCS, CC et CP. Le SCCS consiste à comparer chez un même sujet la survenue d'événements sur une période où il est exposé au médicament d'intérêt par rapport à une période où il ne l'est pas. Le CC compare la distribution de l'exposition au médicament avant la survenue de l'événement chez un cas par rapport à celle-ci chez un témoin susceptible de présenter ce même événement. Tout comme le CC, le CP s'intéresse à la distribution de l'exposition du médicament chez les cas mais utilise comme comparateur la distribution de l'exposition dans l'ensemble de la population source d'où les cas sont extraits. Chaque approche a été déclinée suivant plusieurs variantes qui différaient les unes des autres par leurs paramètres (ex : stratégie d'ajustement, fenêtre de risque, nombre de témoins appariés par cas, *etc.*). Un total de 96 variantes de SCCS, de 20 variantes de CC et de 80 variantes de CP ont été appliquées dans les différentes populations pour évaluer l'association entre les médicaments témoins et les événements d'intérêt considérés. Pour optimiser les temps de calcul, les populations MI, UGIB et AKI ont été échantillonnées à des proportions respectives de 1/20, 1/10 et 1/3 avant l'exécution des analyses.

- *Évaluation des performances*

Pour chaque variante exécutée, les mesures d'associations générées pour chaque médicament témoin ont servi à calculer des indicateurs de performance : sensibilité, spécificité, aire sous la courbe ROC (AUC), erreur quadratique moyenne (MSE) et probabilité de couverture.

- *Sélection et étalonnage de l'approche et des paramètres optimaux*

Afin de différencier certaines variantes d'une même approche aux performances parfois très proches, des régressions logistiques univariées ont été réalisées pour identifier les paramètres ayant un impact significatif sur le pouvoir discriminant du

modèle concerné, c'est à dire sa capacité à différencier les témoins positifs des témoins négatifs pour un événement d'intérêt donné. La variable dépendante à expliquer était la probabilité qu'une variante ait une AUC supérieure au 70ème centile de la distribution des AUC. Les covariables indépendantes explicatives testées correspondaient aux paramètres qui différaient d'une variante à une autre. Par exemple, pour le SCCS, les covariables indépendantes incluaient l'ajustement en fonction des traitements concomitants (codée en oui / non), l'utilisation d'une fenêtre de pré-exposition à 0, 7 ou 30 jours, etc.

Une fois la variante optimum identifiée pour un événement d'intérêt, celle-ci a été répliquée dans la population totale, puis calibrée. Les paramètres d'une « distribution empirique sous l'hypothèse nulle » ont été calculés en supposant la distribution des mesures d'association des témoins négatifs, y compris ceux pour lesquels $p < 0.05$, comme gaussienne. Cette distribution modélise le biais résiduel sous l'hypothèse nulle, c'est-à-dire le bruit de fond inhérent à l'application de l'analyse dans le SNDS. Les paramètres calculés (moyenne et variance) ont par la suite été utilisés pour estimer des valeurs « calibrées » de la p-value, tenant compte de l'erreur aléatoire et de l'erreur systématique.

▪ Résultats

- *Mise en place et identification des populations d'intérêt au sein du SNDS*

La définition spécifique d'ALI a conduit à l'identification de 5 152 patients dans le SNDS sur la période 2009-2014. Un total de 355 patients supplémentaires a été identifié à partir de la définition sensible. Sur la même période 304 369 patients ont présenté un MI (définition spécifique). Ce chiffre passait à 717 920 lorsque que l'on tenait compte de l'angor instable (définition sensible). La définition spécifique d'AKI a permis d'identifier 12 317 patients. La définition sensible portait ce nombre à 89 186. Pour UGIB les définitions spécifique et sensibles ont permis d'identifier sur la période d'étude 139 172 et 178 384 patients, respectivement.

Les événements d'intérêt avec les effectifs les plus élevés étaient ceux présentant le plus grand nombre de médicaments témoins potentiellement détectables (*i.e.* avec un risque relatif minimum détectable $< 1,30$). Pour les définitions sensibles de MI et UGIB, 64 témoins positifs et négatifs sur 70, et 59 sur 64 ont été respectivement considérés comme détectables. Seuls 18 témoins positifs sur 58 et 7 témoins

négatifs sur 23 ont été considérés comme détectables pour la définition spécifique de ALI. L'échantillonnage des populations présentant les plus grands effectifs à des fins de diminution des temps de calcul a été accompagné d'une réduction du nombre de témoins détectables.

- *Résultats Généraux*

Globalement, les SCCS semblaient générer de meilleurs résultats que les CC et CP pour l'identification des médicaments associés aux ALI, AKI, MI et UGIB. Que ce soit pour les définitions spécifiques ou sensibles, les SCCS présentaient le plus grand pouvoir discriminant avec des AUC allant de 0,70 à 0,94, et la meilleure précision avec des MSE allant jusqu'à 0,07. Cette différence était moins marquée pour ALI que pour les autres événements d'intérêt. Les AUC des CC et CP étaient relativement proches, mais les CP présentaient toujours des MSE significativement plus élevées.

- UGIB

En ce qui concerne la définition spécifique de UGIB, les AUC s'étendaient de 0,64 à 0,80, 0,44 à 0,61 et 0,50 à 0,67, pour respectivement les SCCS, CC et CP. Les MSE variaient respectivement de 0,07 à 0,39, 0,83 à 1,33 et 1,96 à 4,6. Les régressions univariées ont montré que les AUC élevées étaient obtenues via les SCCS ajustant sur l'utilisation de traitements concomitants et utilisant une fenêtre de risque fixe correspondant au 30 premiers jours suivant la dispensation du médicament d'intérêt plutôt que de la période totale de traitement. Lorsque exécutée dans la population non échantillonnée, la variante la plus performante de SCCS était associée à une AUC = 0,84 et une MSE = 0,14, avec 10 témoins négatifs sur 36 présentant des estimations significatives. Le processus de calibration a mis en valeur une faible erreur systématique potentielle, résultant principalement du biais d'indication et du biais protopathique affectant fortement certains témoins négatifs.

- ALI

En ce qui concerne la définition spécifique de ALI, les AUC s'étendaient de 0,78 à 0,94, 0,64 à 0,92 et 0,48 à 0,85, pour respectivement les SCCS, CC et CP. Les MSE variaient respectivement de 0,12 à 0,40, 0,22 à 0,39 et 1,03 à 5,29. Les variantes ajustant sur l'usage de traitements concomitants présentaient une probabilité de

couverture plus élevée. Les régressions univariées ont montré que les AUC élevées étaient obtenues via les SCCS utilisant une fenêtre de risque correspondant à la période de couverture par le traitement, et non pas une fenêtre fixe de 30 jours. La variante optimale du SCCS avait une AUC = 0,93 et une MSE = 0,22 pour une couverture = 86%, avec 1 témoin négatif sur 7 et 13 témoins positifs sur 18 présentant des estimations significatives. Le processus d'étalonnage tendait à montrer que l'erreur systématique inhérente à l'application du SCCS dans le SNDS pour la mesure de l'association entre des médicaments et ALI était minime, le nombre limité de témoins négatifs inclus dans cette étude étant néanmoins à la source d'une incertitude élevée.

▪ Discussion

Le SNDS est souvent considéré comme la plus grande base de données homogène du monde. L'exposition aux médicaments dans le secteur ambulatoire y est enregistrée de manière précise et quasi exhaustive, même si un doute persiste quant au fait que les médicaments délivrés aient été réellement consommés par le patient. La qualité de l'information relative aux événements est assurée par le travail quotidien des services hospitaliers d'information médicale. Des divergences entre les données saisies et l'état actuel du patient restent possible, mais des études de validation tendent à montrer que les diagnostics que l'on retrouve dans le SNDS présentent une bonne spécificité. La qualité et la richesse de l'information disponible dans cette base de données permettent souvent de surmonter l'absence de certains éléments non collectés tels que certains facteurs de risque (statut tabagique, indice de masse corporel), les événements survenant dans le secteur ambulatoire, ou les résultats d'examen médicaux. La puissance et la représentativité du SNDS en font ainsi un excellent support pour la détection et la validation de signaux de pharmacovigilance.

En comparant différentes approches pharmaco-épidémiologiques, cette thèse visait à identifier comment tirer le meilleur parti de ce potentiel. Bien entendu, les designs identifiés comme optimaux dans ces travaux pourraient encore être améliorés, mais l'évaluation empirique conduite montrent que, en l'état, dans le SNDS, le SCCS est supérieur au CC et au CP pour la détection d'alertes de pharmacovigilance associées à ALI et à UGIB. La présence de faux positifs a révélé que, bien que

faible, de la confusion résiduelle pouvait persister, en particulier à travers les biais protopathique et d'indication. Cependant, des connaissances en pharmacologie clinique et un point de vue médico-pharmaceutique permettent une interprétation correcte des résultats et la différenciation des vrais et des faux positifs. Par ailleurs, des témoins positifs ont donné lieu à des mesures d'associations proche de 1 ou supérieur à 1 mais non significatives (faux négatifs). Ces effets positifs mais non significatif peuvent être attribués à un manque de puissance résultant d'un faible nombre d'événements ou à une faible utilisation du médicament étudié au niveau populationnel. On peut dans ce cas de figure être amené à s'interroger sur l'intérêt porté à un traitement dont l'impact potentiel en termes de santé publique est limité. En outre, comme la pharmaco-épidémiologie capture l'effet du médicament dans ses conditions réelles d'utilisation, une absence d'effet mesuré peut raisonnablement donner à penser que l'événement qui nous intéresse ne constitue pas un problème dans la pratique clinique quotidienne. Cela peut être la conséquence de l'innocuité réelle du médicament ou bien d'un biais de (contre-)indication, traduisant la gestion correcte du risque existant en situation réelle de soins.

Au total, ces travaux ont montré que le SCCS était une approche adaptée à la détection d'alertes de pharmacovigilance associées à ALI et à UGIB dans le SNDS. Un point de vue clinique demeure toutefois nécessaire pour écarter tout risque de faux positif résultant de potentiels biais résiduels, les faux négatifs, quant à eux, ne posant pas de réel problème. Outre leur utilisation pour la validation de signaux de sécurité générés par d'autres sources (ex : la notification spontanée) et la quantification du risque associé, ces outils calibrés pourraient être utilisés en routine au sein du SNDS, de manière automatisée, pour la génération d'alerte de sécurité dans de larges panels de médicaments, en particulier ceux nouvellement commercialisés. Dans un premier temps, une analyse serait exécutée sur l'ensemble des médicaments commercialisés au niveau national. Les risques déjà documentés seraient exclus et les alertes émergentes soigneusement étudiées afin de distinguer celles d'intérêt. Les médicaments nouvellement commercialisés ou suspects feraient par la suite l'objet d'une analyse périodique. L'extension d'une telle approche à des événements d'intérêt autres que ceux étudiés dans cet ouvrage (ex : rhabdomyolyse, Syndrome de Lyell, etc.) constituerait un progrès substantiel en matière de pharmacovigilance en France.

TABLE OF CONTENTS

Acknowledgments	3
Résumé	9
Abstract	13
Résumé substantiel des travaux en langue française	17
List of tables	31
List of figures	31
List of abbreviations	33
1. Introduction	39
1.1. Pharmacoepidemiology	40
1.1.1. Definition	40
1.1.2. Scope	40
1.1.3. Interface with pharmacovigilance	41
1.1.4. Strengths and weaknesses	42
1.1.5. Data sources	43
1.2. Consortia and networks involved in pharmacoepidemiology research	45
1.2.1. EU-ADR	45
1.2.2. OMOP	46
1.2.3. Sentinel	46
1.2.4. AsPEN	47
1.2.5. CNODES	47
1.2.6. OHDSI	48
1.3. Signal generation in longitudinal databases	49
1.3.1. Disproportionality analysis approach	49
1.3.2. Sequence symmetry analyses	50
1.3.3. Sequential statistical testing approach	50
1.3.4. Temporal association rule approach	51
1.3.5. Supervised machine learning approach	51
1.3.6. Tree-based scan statistic method	51
1.3.7. Traditional pharmacoepidemiological approach	52
1.4. Drug safety monitoring in France	55
1.4.1. Actors	55

1.4.2.	National resources	56
1.4.3.	Safety signal management	59
2.	<i>The ALCAPONE project</i>	61
2.1.	Background and thesis context	62
2.2.	Transparency, scientific independency and legality	63
2.3.	Research question and objectives	64
2.4.	Methodology for the assessment of case-based methods in the SNDS	64
2.4.1.	Submitted article: Empirical assessment of case-based methods for drug safety alert identification in the French National Healthcare System database (SNDS): Methodology of the ALCAPONE project	65
2.4.2.	Methodological aspects.....	95
2.5.	Feasibility study	103
2.5.1.	Assessment of the mapping of the SNDS to the OMOP CDM	103
2.5.2.	Adaption of the reference set and power consideration.....	106
2.5.3.	Conclusion of the feasibility study.....	108
2.6.	Preparation of SNDS data	108
2.6.1.	Data extraction.....	108
2.6.2.	Data mapping.....	110
2.7.	Results	111
2.7.1.	General results.....	111
2.7.2.	UGIB example, submitted article: Empirical assessment of case-based methods for identification of drugs associated with upper-gastrointestinal bleeding in the French National Healthcare System database (SNDS)	113
2.7.3.	ALI example, submitted article: Empirical assessment of case-based methods for identification of drugs associated with acute liver injury in the French National Healthcare System database (SNDS).....	157
3.	<i>General discussion, conclusion and perspective</i>	201
3.1.	IT constraints	202
3.2.	Interest of health professional point of view	203
3.3.	Routine implementation and generalizability	204
	References	207
	Appendices	223
	Appendix 1 Communications related to the thesis	225

Appendix 2	Award letter for ENCePP Seal	227
Appendix 3	Décision d’approbation n°187 de l’Institut des Données de Santé du 25 mai 2016	228
Appendix 4	Notification d’autorisation CNIL pour le projet ALCAPONE.....	229
Appendix 5	Poster displayed at the 2018 European OHDSI Symposium, March 23rd, Rotterdam, Netherlands	231
Appendix 6	Poster displayed at the 33rd ICPE, August 2017, Montréal, QC, Canada	232
Appendix 7	Complete Extract, Transform and Load process of SNDS to OMOP CDM v5 for the ALCAPONE project.....	233

LIST OF TABLES

Table 1. Contingency table for CP per-user approach..... 97

Table 2. Contingency table for CP person-time approach 98

Table 3. Contingency table computed for each design variant of each health outcome of interest 100

Table 4. Percentage of loss of positive (CTR+) and negative (CTR-) controls resulting from the restriction of the OMOP reference to the drugs available on the French market..... 107

Table 5. Calculated number of detectable controls in the EGB and estimated number of detectable controls in the SNDS 108

LIST OF FIGURES

Figure 1. Overview of the Extract Transform and Load process from the SNDS to the OMOP Common Data Model (CDM) v5..... 110

Figure 2. Overall performances of self-controlled case series (SCCS), case-population (CP) and case-control (CC) designs in the detection of drugs associated with upper-gastrointestinal bleeding (UGIB), myocardial infarction (MI), acute liver injury (ALI) and acute kidney injury (AKI), specific definitions 111

Figure 3. Overall performances of self-controlled case series (SCCS), case-population (CP) and case-control (CC) designs in the detection of drugs associated with upper-gastrointestinal bleeding (UGIB), myocardial infarction (MI), acute liver injury (ALI) and acute kidney injury (AKI), sensitive definitions 112

LIST OF ABBREVIATIONS

ALCAPONE	Alert generation use the case-population approach in the French nationwide healthcare databases
ANSM	<i>Agence nationale de sécurité du médicament et des produits de santé</i>
AP-HP	<i>Assistance Publique - Hôpitaux de Paris</i>
AsPEN	Asian Pharmacoepidemiology Network
ASSAPS	<i>Agence française de sécurité sanitaire des produits de santé</i>
AUC	Area under the receiver operating characteristic curve
BCMD	<i>Base de Causes Médicales de Décès</i>
BNPV	<i>Base nationale de pharmacovigilance</i>
CC	Case-control
CDM	Common Data Model
CEREES	<i>Comité d'expertise pour les recherches, les études et les évaluations dans le domaine</i>
CIHR	Canadian Institutes of Health Research
CNAM	<i>Caisse Nationale d'Assurance Maladie</i>
CNIL	<i>Commission Nationale de l'Informatique et des Libertés</i>
CNODES	Canadian Network for Observational Drug Effect Studies
CP	Case-population
CPR	Case-population Ratio
CPRD	Clinical Practice Research Datalink
DRUGS-SAFE	Drugs systematized assessment in real-life environment
DSEN	Drug Safety and Effectiveness Network
EGB	<i>Échantillon Généraliste de Bénéficiaires</i>
EHDEN	European Health Data & Evidence Network
EHR	Electronic health record
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EU-ADR	Exploring and Understanding Adverse Drug Reactions by integrative mining of clinical records and biomedical knowledge
FDA	Food and Drug Administration
GP	General practitioner
HOI	Health Outcome of Interest
ICD10	International classification of Diseases, 10 th revision

ICPE	Conference on Pharmacoepidemiology & Therapeutic Risk Management
IMEDS	Innovation Medical Evidence Development and Surveillance
IMI	Innovative Medicines Initiative
INDS	<i>Institut National des Données de Santé</i>
INSEE	<i>Institut National de la Statistique et des Études Économiques</i>
INSERM	<i>Institut National de la Santé et de la Recherche</i>
LTD	<i>Long-term disease registration</i>
MDRR	Minimal detectable relative risk
OHDSI	Observational Health Data Sciences and Informatics
OMOP	Observational Medical Outcomes Partnership
PEPS	<i>Pharmaco-Épidémiologie des Produits de Santé</i>
PMSI	<i>Programme de médicalisation des systèmes d'information</i>
PREPS	<i>Programme de recherche sur la performance du système des soins</i>
pRR	Predicted Relative Risk
RCT	Randomized clinical trial
ROC	Receiver operating characteristic
SCAN	Surveillance of Health Care in Asian Network
SCCS	Self-controlled case series
SNDS	<i>Système National des Données de Santé</i>
SNIIRAM	<i>Système National d'Information Inter Régimes de l'Assurance Maladie</i>
US	United States
WHO	World Health Organization

*All drugs are dangerous,
some may also be useful.*

Moore N. Assessing drug safety. Bmj 2005, 330:539-540.

Competing interests: NM has worked in drug safety for over 20 years and is convinced he is right, most of the time.

1. INTRODUCTION

1.1. Pharmacoepidemiology

1.1.1. Definition

Pharmacoepidemiology can be defined as the study of the use and the effects of drugs on large number of people.¹ This research area borrows from both pharmacology and epidemiology. On one hand clinical pharmacology focus on the effects of drugs on humans, looking to understand the relationship between drug exposure and its response. Epidemiology on the other hand analyses the distribution and determinants of diseases in population through various descriptive and analytical methods.² Thus, pharmacoepidemiology can be seen as the application of epidemiological methods to pharmacological issues, combining the interest of epidemiology for real life and large populations to the concern of pharmacology for drug effects. One might consider it as just another method in experimental pharmacology. However, instead of testing drugs on individual animals or humans, pharmacoepidemiology works on a much larger scale, considering overall populations. Therefore, pharmacoepidemiology studies the determinants and consequences of drug utilization among populations in real life settings.

1.1.2. Scope

At the time of their marketing, the effects of drugs and especially their efficacy have been studied mostly in randomized controlled clinical trials (RCT), comparing them to placebo or to existing drugs. RCTs have the greatest inferential power: they ensure that the treatment “causes” the outcome.³ However, they are by nature limited. They are restrictive in terms of age (children and elderly often being excluded), gender (pregnant women being excluded), co-morbidities (patients with liver or renal diseases, or multiple comorbidities being excluded) and co-prescriptions (patients treated with several drugs being excluded).⁴ Since RCTs are very expensive to set-up, they are usually undersized and too short in time to detect delayed outcome or outcome of low frequency. An outcome occurring 1/1000 will not be detected in an RCT including 500 patients. Similarly, delayed effects such as cancer will not be detected in a 12-month RCT. Hence, proof is needed to ensure that the benefit-risk balance observed during clinical trial phases and justifying the marketing authorization, is still true in real world situations, and that the drug is used correctly in the day-to-day practice.⁵ Where RCTs aim to demonstrate efficacy,

pharmacoepidemiology focuses on effectiveness: efficacy under real conditions of prescription and use.⁶ Furthermore, when several concurrent drugs are launched within a short time frame (e.g. direct acting anticoagulants), in most cases, there is no comparative RCT:⁷ it is very unlikely that any pharmaceutical company will conduct at great cost a directly comparative RCT, taking the risk to conclude on the inferiority of his own product with respect to direct competitors. Because the introduction of new drugs on the market may impact the benefit-risk balance of other treatments, comparative effectiveness research – the assessment of the different therapeutic options for a given medical condition, taking into account the reality of the daily clinical practice – is crucial. Using databases, exploiting registries, enrolling patients, pharmacoepidemiology tries to take up these challenges, avoiding as much as possible any modification to the real life that could be caused by the study itself.⁶ The rising need of real world evidence in the recent years has considerably increased the public health authorities' interest in pharmacoepidemiology. This bridge science provides information about the impact of a drug in the population and therefore can also be useful to assess practical consequences and cost-effectiveness of public health decision.

1.1.3. Interface with pharmacovigilance

Pharmacovigilance is defined by the World Health Organization (WHO) as “the science and activities relating to the detection, assessment, understanding and prevention of adverse effects or any other drug-related problem”.⁸ In practice, pharmacovigilance refers almost exclusively to the spontaneous reporting systems, which allow health care professionals and others to report adverse drug reactions to a central agency.⁹ However, other activities such as case-control (CC) or cohort studies can be used to monitor drug safety. Although pharmacovigilance mainly considers safety outcome, one of its ultimate goals is still the assessment and communication of the risks and benefits of drugs on the market.¹⁰ In this scope, pharmacoepidemiology meets the pharmacovigilance. Benefit-risk assessment cannot be conducted without both safety and effectiveness: safety data are not self-sufficient to rule on drug effectiveness, and pharmacoepidemiology makes no sense if safety outcomes are not taken into consideration. In that respect, we can say that pharmacovigilance and pharmacoepidemiology are two aspects of a same activity: assessing drugs in real world settings for a rational, efficient and safe use.

1.1.4. Strengths and weaknesses

The strengths of pharmacoepidemiology respect of RCTs have already been partially discussed. The rise of healthcare data (e.g. national claims database, electronic medical records) allows pharmacoepidemiology to assess in reduced timelines and for tiny cost drugs in large population, even sometimes in overall national population, over years of follow-up. Pharmacoepidemiology also borrows two classical epidemiology advantages: the presence of a well-defined “denominator” corresponding to the study population, and the ability of implementing a comparison group. Those two elements enable risks and benefit quantification.¹¹ As a global population is considered during a study, the proportion of patients exposed and/or the proportion of cases observed can be used to estimate relative risks, which is impossible with spontaneous reporting.

The main strength of pharmacoepidemiology is also its weakness. Because by definition, exposure and outcomes are assessed in real life settings, surrounding factors are not controlled and randomization is not present to balance them.² Some methodological and mathematical solutions exist to address such issues (e.g. active comparator new user study design,^{12, 13} high-dimensional propensity scores¹⁴), however we do not know to what extent residual confounding could persist. Thus, pharmacoepidemiology studies are often criticized since the strength of causal inferences that can be drawn remains uncertain.⁴

The limits of pharmacoepidemiology directly depend on the source of data available. Just any database cannot be used to answer any questions. The consistency between the topic of interest and the data source considered is crucial. For example: minor health issues usually addressed by general practitioners (GP) will not be found in hospital discharge summaries. Similarly, it is very unlikely that serious events such as myocardial infarction are systematically recorded in GP electronic charts. Furthermore, intrinsic weaknesses and biases relative to the data source chosen will also impact the study when interpreting the results (e.g. missing data, discrepancies between prescribed, dispensed and administered drug, etc.).

1.1.5. Data sources

1.1.5.1. Primary data collection

Ad-hoc studies can be conducted through primary data collection. In that case, specific data such as quality of life, lifestyle, blood or DNA samples, *etc.* can be collected for a well-defined research problem.¹⁵ This will consist mainly in field studies. Exposure and outcome are directly collected from physician and/or patients by a study center. Patients can be recruited using pre-existing disease or exposure registries, or through hospitals, specialists or surveillance centers, most of the time to answer etiological questions.^{16, 17} Cohort of patients can be set up and followed during years for etiological but also more descriptive purposes.^{18, 19} Primary data collection is important in pharmacoepidemiology: allowing in-depth case assessment by clinical experts for a large source population, it enables the evaluation of drug-disease associations for rare complex conditions.²⁰ However, such studies are often affected by selection bias during the recruitment process: patients participating to the study may not be representative of the whole patient population.^{21, 22} They may also be affected by what patients or physicians remember (recall bias),²³ or subsequent changes in patients or physicians behavior caused by research participation (Hawthorne effect).^{24, 25} Moreover studies relying on primary data collection remain expensive and complex to set up. As with RCTs costs are proportional to the size of the population and the duration of the follow-up.²⁶

1.1.5.2. Secondary use of data

Most recent pharmacoepidemiology studies rely on secondary use of data, most of the time already available electronic patient healthcare data. Healthcare databases are longitudinal databases. Patients are followed over time. As data are collected prospectively and systematically, they are affected neither by recall bias nor by the Hawthorne effect. In this scope, they are considered as more exhaustive than primary data collection, especially over the long-term. Healthcare databases can be divided in two types of records: electronic health records (EHRs) and claims databases.

- ***Electronic health records (EHRs)***

EHRs are continuously fed after anonymization by the information entered by a health professional as part of patient management (*e.g.* diagnostics, prescribed

drugs, laboratory tests or imaging procedure results, *etc.*). The nature of the information available is directly related to the nature of the EHR itself. EHRs consisting of electronic medical charts from a GP network will group events regularly observed in primary care (*e.g.* fevers, high blood pressure, diabetes, *etc.*). This is the case of the Clinical Practice Research Datalink (CPRD) database in the United Kingdom.²⁷ Conversely, those related to hospital settings will mainly focus on serious events (*e.g.* myocardial infarction, acute renal failure, cancer, *etc.*). Lifestyle details such as smoking status or body mass index are often present. Data completeness depends on the propension of the physician to complete the patient's chart carefully, and on patient loyalty to their physician. In so far as EHRs are not universal, switching physicians may induce interruption of the data. Moreover, the exploitation of collected data may present some challenges since all variables may not be uniformly coded, or sometimes not coded at all. Complex approach such as natural language processing may be necessary to extract the useful information from free text.²⁸

- ***Claims databases***

Claims databases were initially developed for billing purposes. They contain reimbursed healthcare expenditures coded using national or international classification and continuously collected from hospital, laboratory, pharmacy or healthcare professional visits. The data holder, usually national or private health insurances, then uses the information to reimburse patients and health care providers (hospitals, pharmacists, *etc.*). Claims databases cover very large populations. For example, the Market Scan Research Databases²⁹ in the US counts more than 250 millions patients. In countries where national healthcare insurance is provided to each citizen, the overall national population is included (*e.g.* France, Finland, Denmark, Finland, Iceland, Norway, Sweden). As long as a health care is reimbursed, the information is recorded for affiliated patients. The exposure is exceptionally well captured, which is a huge advantage when conducting pharmacoepidemiology studies. The exact form, dosage and brand of the drug dispensed as well as the quantity and the date are fully recorded. Date and type of procedures, lab tests and medical visits are usually present, as well as hospital discharge diagnostic codes and basic sociodemographic data. However, medical details are missing: no lab test results or medical examination reports are available.

Risk factors such as weight, smoking status, alcohol consumption or allergies are not included either.

According to national legislations, healthcare databases (claims databases, EHRs, death and pathologies registries, *etc.*) can be linked together offering additional opportunities to the investigators. In the last years we have witnessed the development of the possibility of linkage: in France, merging primary collected data to healthcare database became possible. Such practices allow to address respective limits of each database, for example gathering specific *ad-hoc* medical data of the field to large and exhaustive healthcare histories available in claims data.³⁰ CONSTANCES cohort is an example of this linkage between primary collected and claims data.³¹

1.2. Consortia and networks involved in pharmacoepidemiology research

In 2004, following the unexpected worldwide withdrawal of rofecoxib (Vioxx) an international reflection was introduced on how to better assess benefit-risk balanced throughout the life of a drug.^{32, 33} The permanent desire to study ever-rarer events and exposures, led to try to include ever more patients. To do so, research networks were set-up at a national or international level. Additional knowledge on the consistency of the information collected and on the generalizability of the results generated were also provided by this approach: is the effect observed in several countries and if not, why?²⁰ This search for power came with the necessity to develop new tools to improve accuracy and consistency of the results. Hence, different networks have been set-up in the last years, sometimes prompted by politics or health authorities and sometimes as part of public-private partnership.

1.2.1. EU-ADR

Exploring and Understanding Adverse Drug Reactions by integrative mining of clinical records and biomedical knowledge, the EU-ADR project, was launched by the European Commission in 2007 to develop new methodologies for drug safety monitoring on large databases following the withdrawal of Vioxx.³⁴ Eight European databases located in four countries (Italy, Netherland, United Kingdom and Denmark) were involved in a common data framework through a “distributed network approach” for a total study population of almost 20 million people. Using standardized input files, the *ad-hoc* developed software Jerboa[®] queried and aggregated patient-level data in

each participating database. Resulting data were then pooled and analyzed. EU-ADR project worked a lot on signal detection approaches.³⁵⁻³⁷ The framework developed was also used in multidatabase European studies such as the SOS project, on the safety of non-steroidal anti-inflammatory drug.³⁸ On the results of EU-ADR project, the European Commission funded the EU-ADR alliance, a stable collaboration framework for running drug safety studies on 8 databases from five countries: the previous four and Germany.³⁹

1.2.2. OMOP

One of the consequences of Vioxx history in United States (US) was the Food and Drug Administration (FDA) Amendments Act of 2007 calling for establishing at the national level a Risk identification and analysis System: a nationwide network of databases targeting to capture by 2012 data on more 100 millions subjects, which will take the form in 2008 of the Sentinel initiative.^{40, 41} In the same time, a public-private partnership was established, the Observational Medical Outcomes Partnership (OMOP), with the objective of informing the appropriate use of observational healthcare database. Between 2009 and 2013, OMOP conducted methodological research to empirically evaluate performance of various analytical methods in different type of databases.⁴²⁻⁵⁰ One of the side effect of these OMOP experiments was the creation of a Common Data Model (CDM).⁵¹ The latter was initially designed to be able to apply same methodology across all the involved databases for comparison purposes.⁵² In 2013, at the end of the 5 years, the research laboratory moved to the Reagan-Udall Foundation for the FDA under the Innovation Medical Evidence Development and Surveillance (IMEDS) Program⁵³ and OMOP research investigators initiated the Observational Health Data Sciences and Informatics (OHDSI).⁵⁴

1.2.3. Sentinel

As previously mentioned, the Sentinel initiative started in 2008 with the aim of obtaining information from multiple existing databases to assess the safety of approved medical products.⁴⁰ As part of this initiative, the Mini-Sentinel program was launched in 2009. This pilot program included 31 academic and private organizations.⁵⁵ Mini-sentinel was based on a distributed data-system: each data partner maintained physical and operational control over their own data, but also

transformed them into a CDM according to well defined specifications.⁵⁶ Then, *ad-hoc* or standardized program data were distributed and locally executed. Generated outputs were reviewed and transferred to the Operation Center. With Mini-Sentinel semi-automated method to address confounding issue were developed or assessed, such high-dimensional propensity scores⁵⁷ or instrumental variables.⁵⁸ After 2011, involved databases were usable for distributed queries, enabling the conduct of original investigations⁵⁹ In 2016, Mini-Sentinel transitioned from its pilot stage to a full-fledged program: the Sentinel System, an active surveillance system to monitor the safety of regulated medical product.⁴⁰

1.2.4. AsPEN

The Asian Pharmacoepidemiology Network⁶⁰ (AsPEN) started during the 4th Asian Conference on Pharmacoepidemiology in 2009 in Tainan, Taiwan. It is a multinational research network formed to provide mechanism to support the conduct of pharmacoepidemiological studies and the prompt identification and validation of emerging safety issues among the Asian Countries.⁶¹ As Sentinel, AsPEN uses a distributed network approach. Data holders maintain raw data in their site. When a study is set-up, each participating center is in charge of managing their data under the instructions of the coordinating center to fit a common structure. Unique analytic program is then applied at the local level. Results from multiple sites are then pooled and analyzed. In the early phase of AsPEN, this common structure was study-specific.⁶²⁻⁶⁴ In 2014, at the 30th International Conference on Pharmacoepidemiology & Therapeutic Risk Management (ICPE), AsPEN started a reflection on its infrastructure development through the Surveillance of Health Care in Asian Network (SCAN) project, with the view of implementing a formal global CDM and chose the OMOP one.^{62, 65} AsPEN members count with Taiwan, Japan, Korea, Australia, China Thailand, Singapore, Hong-Kong, Canada, US and Sweden. The most often involved databases are those from Australia, Hong-Kong, Japan Korea and Taiwan.^{64, 66, 67}

1.2.5. CNODES

The Canadian Network for Observational Drug Effect Studies⁶⁸ (CNODES) was created in 2011. It is part of the Drug Safety and Effectiveness Network (DSEN), a joint initiative of Health Canada⁶⁹ and the Canadian Institutes of Health Research (CIHR).⁷⁰ CNODES was set up to answer Canadian regulators queries about drug

safety and effectiveness. Methodological research is also conducted.⁷¹⁻⁷³ This network consists of seven Canadian provincial databases and two international ones: the CPRD from United Kingdom and the US MarketScan database, gathering over 100 millions patients in total. In order to take advantage of local expertise and analytical capacities, CNODES also chose a distributed network model but applies it in a different way than Sentinel or AsPEN. When Health Canada submit a question to CNODES, a project team including people from each province is set up to develop common scientific and analytical protocols. After getting local approval for the research, analyses are conducted separately in each database.⁷⁴ Algorithm or statistical analysis software can be share but do not involved the use of any CDM. Results from the different databases are then pooled together through a meta-analysis.⁷⁵⁻⁷⁷

1.2.6. OHDSI

OHDSI ⁷⁸ is a multi-stakeholder, interdisciplinary collaborative initiated by the former OMOP investigators. The initiative is coordinated in the US by the Columbia University (New York), and in Europe by the Erasmus Medical Center (Rotterdam, Netherlands). OHDSI comprises academics, industry scientists, health care providers and regulators. Funds are provided by a variety of sources, including grants from government agencies, foundations and industry.⁷⁹ The works conducted rely on the OMOP CDM, which the group maintained and developed with its associated standardized vocabulary. Whereas OMOP was restricted to methodological research, OHDSI develops and applies methods to observational data to answer real-world clinical questions. OHDSI's overall approach is to create an open network of data holders, and require that they translate their data to the OMOP CDM and its standardized vocabularies.⁵⁴ Tools are provided to help this conversion. Besides the community developed a large library of visualization tools and analytical methods in free access that can be implement in whatever database meeting the OMOP CDM requirement.⁸⁰⁻⁸⁴ This enables to conduct multicenter study on a distributed network model, carrying out analyses locally and transmitting results to the coordinating center. As OHDSI is fully collaborative, any collaborator can propose to the community a study project. Once the study synopsis posted on the OHDSI research forum,⁸⁵ the project will be reviewed by other collaborators. If multiple sites show

interest for it, a common protocol will be produced. Interested collaborators will then be invited to join the study, run the code and share the results.^{86, 87}

Large databases in US are available to members of the OHDSI community and have adopted the OMOP CDM, for example MarketScan Commercial Claims and Encounters (250 million patients), Optum ClinFormatics (40 million patients), GE centricity (33 million patients), *etc.* In the recent years several Asian databases in Hong-Kong, Japan Taiwan and South Korea joined. In Europe, so far, very few databases are mapped to the OMOP CDM. To support OMOP CDM adoption in Europe, in 2018, the European Health Data & Evidence Network (EHDEN) was launched.⁸⁸ EHDEN is a consortium depending of the Innovative Medicines Initiative⁸⁹ (IMI) a public-private partnership between the European Commission and pharmaceutical companies, aiming to build a large-scale federated network of data sources to optimize the generation of real world evidence, taking advantage of the OHDSI achievements.

1.3. Signal generation in longitudinal databases

Since their creation, pharmacoepidemiology networks and consortium devoted considerable effort to methodology research, especially safety signal detection, the availability of these large healthcare databases opening new opportunities to overcome traditional spontaneous reporting limits.⁹⁰⁻⁹³ Numerous data-mining approaches for drug safety signal detection resulted from this need of better assessing drug at the post-marketing stage. This section aims to describe the global concepts of those methods, their related strengths and limitations, relying, amongst other sources, on the article cited below:

Mickael Arnaud, Bernard Bégaud, Nicolas Thurin, Nicholas Moore, Antoine Pariente & Francesco Salvo (2017): Methods for safety signal detection in healthcare databases: a literature review, Expert Opinion on Drug Safety, DOI: 10.1080/14740338.2017.1325463

1.3.1. Disproportionality analysis approach

Disproportionality analysis methods were originally developed on spontaneous reporting databases, which are transversal databases, with neither patient history nor follow-up. In this approach, all spontaneous reports are displayed in the form of a large contingency table with dimensions corresponding to all the drugs and events encountered. For each drug-pair a 2x2 contingency table is generated. Ratio of the

observed-to-expected count of each drug-outcome pair is then computed. Disproportionality methods differ by the way they “field” the contingency table (how they generate drug safety reports from healthcare longitudinal data), and how the ratio observed-to-expected is calculated.^{36, 94} Such methods are easy to implement but do not provide risk estimates. Moreover, despite methods enhancement, they are still very sensitive to numerous confounders including protopathic bias and channeling by indication, resulting in bad performances.⁵⁰

1.3.2. Sequence symmetry analyses

The idea is to compare the sequence of initiation of two drugs, A and B, within a given time window where drug A is the exposure of interest and drug B a surrogate for the potential adverse event.⁹⁵ If drug A actually induces an adverse event that needs to be treated with drug B, the number of patients initiating first drug A and then drug B will be higher than the number of patient initiating drug B before drug A. Even if some adjustments are proposed, such approach can be affected by time-varying confounding as changes in prescription trends with an uncorrelated increase or decrease of drug A or B prescription during the study period.⁹⁶ Protopathic bias and confounding by indication can also impact the results. Even if the method is not applicable for death, sequence symmetry analysis remains very interesting since it is easily understandable and shows good performances.⁹⁷

1.3.3. Sequential statistical testing approach

Those methods aim to test sequentially (e.g. on a monthly basis), on prospective cohort data, the null hypothesis: “Is the event rate higher among exposed patients compared to unexposed?”. Each new analysis takes into account the number of new exposed and unexposed patients since the last one, and the increment in exposure time for patients previously included. A signal is raised when the generated test statistic exceeds a predefined value. Approaches differ by the way in which non-cases are handled and the way they managed exposed and non-exposed subjects.⁹⁸⁻¹⁰⁰ Maintaining the type I error at 0.05 across multiple testing is sometimes challenging but the main limitations of such methods are the absence of risk estimate and its sensitivity to protopathic bias and confounding by indication. Moreover they seem to perform worse than random signal detection.¹⁰¹

1.3.4. Temporal association rule approach

Such approaches rely on two rules: (1) the event must follow drug exposure, and (2) the event must occur during a prespecified time window (*i.e.* the period considered at risk). For a given drug, all potential events are mined sequentially, and a correlation score is computed using measure of interestingness. Control periods are set-up and considered differently according to the method.^{102, 103} Additional metrics may be computed and integrated or not to the final score to make them more robust.^{104, 105} However, most of those methods neither provide any risk estimate nor a natural threshold for discriminating positive and negative signals. As the disproportionality analyses, this approach is very susceptible to confounding and show poor performances.¹⁰⁶

1.3.5. Supervised machine learning approach

Supervised machine learning approach consists of two stages. The first one is to train a classifier (a random forest model) by using a sample that includes drug-outcome pairs with previously known related or not-related association. For each pair, a vector of predetermined parameters corresponding to proxies for the association is generated. Using resampling methods and an impurity criterion, the classifier selects the best parameters: the ones enabling the detection of the true associations in the training sample. The second stage consists in the extraction of the selected parameters and to apply the trained classifier in the data of interest to predict associations that could be new drug safety issues. Supervised machine learning does not generate risk estimates. Running such an approach calls for a significant computational load: one random forest model needs to be defined by drug screened, and to be defined efficiently, a large training sample with numerous drug-outcome pairs is required. Performances seem to be excellent,^{107, 108} but it is important to note that the condition of evaluation directly impacts method's ability for prediction.

1.3.6. Tree-based scan statistic method

The fundamental principle of the tree-based scan statistic method proposed is to map a tree according to the basis of the hierarchical structure of classifications used for coding events: the root corresponds to the broadest definition of a given event, the nodes correspond to the different sublevel definitions, the leaves correspond to the

codes with the finest definitions, and the branches link the three elements together.¹⁰⁹ For each leaf, the observed and the age and sex-adjusted expected number of exposed persons who experienced the selected event is computed. Then, all the possible samples of a given root–node–leaf event pathway are tested simultaneously using the log-likelihood ratio test statistic. The method uses Monte-Carlo-based *p*-values to formally adjust them for multiple testing due to the many overlapping definitions of the events to maintain the overall type I error at $\alpha = 0.05$.¹¹⁰ Like most of the methods presented, this approach does not generate risk estimate and is unable to handle confounders including protopathic bias. No formal assessment using a reference set was conducted to assess its performances, however a safety signal justifying further investigation was detected through this approach.¹⁰⁹

1.3.7. Traditional pharmacoepidemiological approach

Several of those pharmacoepidemiological designs have been tested in the US during the OMOP experiments, and in Europe by EU-ADR.^{36, 37, 42, 49, 101} Most of those designs, which have been extensively used in *ad-hoc* studies in epidemiology and pharmacoepidemiology, consist in a two-step process: (1) to constitute from the database two groups of patients based on exposures (cohort approach) or events (case-based approach), and (2) to compare the rate of the drug-outcome association in these groups. Statistical tools are usually available to control for putative confounders (*e.g.* co-prescription). The strength of such methods is that risk estimates are provided, allowing a first approximation of the impact of a potential signal.

1.3.7.1. New user cohort design

Cohorts of patients are followed prospectively from the first start of a drug exposure: one cohort includes patients initiating the drug of interest, while another cohort includes patients newly exposed to another drug.⁴⁶ The drug used for comparison generally shared the same indication with the drug of interest to address confounding by indication.¹² The rate of occurrence of the event(s) of interest is then compared in those two populations. The cohort design provides many solutions for addressing confounder such using age and sex, propensity scores or high dimensional propensity scores, to match patients in both cohorts or weight incidence rate ratio or

adjust model.¹⁴ In the assessment conducted, new user cohort designs seem to perform better when high dimensional propensity scores are used.^{36, 37, 101} To be able to detect signals relative to rare events with a cohort design, a dataset including a large number of persons is required. Moreover, since a (high dimensional) propensity score represents the probability for a patient to be treated with a defined drug, one score has to be computed for each drug of interest (or drug pair in case of a comparison), which can be problematic when screening a large set of drugs. In the context of a cohort study considering only few outcomes of interest but multiple exposures, disease risk score adjustment appears to be a valuable alternative to adjust for confounding.¹¹¹ Like the propensity score, the disease risk score is a summary measure derived from the observed values of the covariates.^{112, 113} However, disease risk score derives the probability of disease occurrence, characterizing the relationship of risk factors with a given outcome of interest.

1.3.7.2. Matched case-control designs

The basic principle of matched case-control design is to analyze retrospectively, starting from a given date, prior drug exposure(s) among two groups of subjects matched on confounders (*e.g.* age, gender).⁴⁴ The first group includes patients (*i.e.* the cases), and the second, patients free of the event (*i.e.* the controls). Odds of exposure to the drug(s) of interest are then compared in the two groups. The availability of large health-care databases popularized the use of CC designs nested in a cohort of patients, which improves the comparability across groups.¹¹⁴ Again, a large number of subjects is required to handle rare event and/or exposure. Classical matched CC design showed comparable performances to new-user cohort designs, but approaches using more advanced settings to control for confounders such as propensity scores, or disease risk score, have not been tested for signal detection so far.^{36, 37, 42, 101} Particular care must be taken in control sampling and in the definition of covariate assessment period to avoid as much as possible potential confounding.^{115, 116}

1.3.7.3. Case-population design

This approach was not presented in the original literature review. In the case-population (CP) design, the rate of exposure to the drug of interest in cases is compared to the rate of exposure in the entire population from which cases were extracted.¹¹⁷⁻¹¹⁹ This approach is also called population-based case-cohort study, or

case-only study, since this design only differed from the CC design by the way the control group is selected.¹²⁰ Here, the global population is used as controls. No individual data are required for them. Exposure can be approximate by drug sales or from aggregated data, as long as data are representative of the considered population.^{119, 121} In circumstances where exposure in the population and cases are either rare or very rare, cohort and CC approach may be ineffective because of power issues. In this context CP can be helpful as long as the exhaustivity of the cases and a measure of the exposure for a territory are available.¹¹⁷ Studies have compared results provided by the CP to other study designs: generated estimates were of the same order of magnitude as the corresponding relative risks estimated with a standard control group.^{118, 122-124} Only one study exploring the performances of the CP design as tool for signal detection has been identified.¹²⁵ Results showed that this population-based approach was able to detect known teratogen risks for several widely used nervous system drugs and to not detect association for drugs considered to be safe. By definition case-population is very sensitive to bias, and performances of the design seem to be directly correlated to the incidence of the exposure and the event in the population.

1.3.7.4. Self-controlled designs

Self-controlled designs differ from the previous approach in that only one cohort of patients is considered and each patient is his/her own control. The effect of a drug on the occurrence of an event is measured for each patient by comparing the event rate in risk periods (usually the exposed periods) to the control periods (usually the unexposed periods). One strength of this kind of design is that all time-invariant confounders are implicitly controlled (e.g. chronic comorbidities, gender, genetic factors, area of residence, *etc.*). The self-controlled case series (SCCS) design considers only patients who have been both exposed to the drug of interest and experienced the event at least once. Statistical tools through penalized regression models were developed to apply high-dimensional multivariate adjustment to control for time-varying confounders (seasonality, comedications, age, *etc.*).^{48, 126}

Self-control cohort differs from SCCS by considering all the exposed patients whether or not they have experienced the event. However, contrary to the SCCS, no specific statistical tool has been yet developed to control for time-varying confounder.¹²⁷ The case crossover design is similar to SCCS in the sense that only cases exposed at

least once to the drug of interest are considered. However, we are not looking here to the event rate during exposure period. Cases are considered retrospectively from their event onset. Risk and control periods are defined arbitrarily with respect to the event. Exposed/unexposed status is then determined based on whether exposure to the drug of interest overlaps with one of the risk periods. The probability of exposure during risk and control periods are then compared.¹⁰¹ An interesting variant of the case-crossover design is the case-time-control design where adjustment for time-trends in exposure is obtained using a set of control subjects.¹²⁸ OMOP experiments showed very good performances for SCCS and self-controlled cohort. Case-crossover has been less studied and no assessment of case-time-control design has been conducted.^{46, 48, 101}

1.4. Drug safety monitoring in France

1.4.1. Actors

In France, the agency in charge of the safety of health products (drugs, biologics, medical devices, cosmetics and biocides) is the *Agence nationale de sécurité du médicament et des produits de santé*¹²⁹ (ANSM), formerly called *Agence française de sécurité sanitaire des produits de santé* (AFSSAPS) from 1999 to 2012 and *Agence du médicament* from 1993 to 1999. ANSM is involved in the safety surveillance of drugs throughout their lifecycle. It issues RCT and marketing authorization. Once the product launched on the market, the agency is in charge of the regular assessment of its benefits-risks balance. To this end, the ANSM coordinates the national pharmacovigilance system with its network of 31 regional pharmacovigilance centers in charge of the adverse event spontaneous reporting system.^{130, 131} In the last years, as a result of the Mediator[®] crisis, the agency has demonstrated an increasing interest on pharmacoepidemiology. In addition to studies it may require from the drug marketing authorization holders, ANSM decided to build capacity for independent pharmacoepidemiology investigation. In 2012, it set up a department of health product epidemiology and initiated a collaboration with the *Caisse Nationale d'Assurance Maladie* (CNAM), the data holder of the *Système National des Données de Santé* (SNDS) database (the French National Healthcare System database), which already had experience in conducting such studies.¹³² In 2014, a call for application was launched to fund for 4 years two platforms with

pharmacoepidemiology abilities and no links with pharmaceutical industries, whereupon in 2015, agreements were made with the “Drugs systematized assessment in real-life environment” (DRUGS-SAFE) platform¹³³ in Bordeaux led by Pr. Antoine Pariente, the *Pharmaco-Épidémiologie des Produits de Santé* (PEPS) consortium in Rennes led by Pr. Emmanuel Oger and the ANSM.¹³⁴ In December 2018 the teams involved in health product epidemiology from both the ANSM and the CNAM merged into a single structure, EPI-PHARE, with the aim of efficiently conduct and coordinate health product epidemiology studies using SNDS data.^{135, 136} In September 2009, the DRUG-SAFE platform obtained a renewal of its 4-year funding to support EPI-PHARE in its pharmacoepidemiology missions.¹³⁷

1.4.2. National resources

1.4.2.1. BNPV

The *Base Nationale de Pharmacovigilance* (BNPV) holds all the anonymized suspected adverse event cases spontaneously reported by health professionals and patients to the regional pharmacovigilance centers.¹³⁸ Before their inclusion in the database, a causality assessment is conducted for each report where experts estimate the putative causal relationship between the drug(s) involved and adverse event reported.¹³⁹ Data from the BNPV are regularly uploaded to EudraVigilance which centralized all the adverse events reported in the countries where drugs from the European economic space are available. Cases are also uploaded to Vigibase, which is the WHO global database of individual case safety reports managed by the Uppsala Monitoring Centre. Researchers from regional pharmacovigilance centers can access BNPV, EudraVigilance and Vigibase.

1.4.2.2. SNDS

The SNDS covers about 99% of the French population (about 66.6 million persons) from birth (or immigration) to death (or emigration).¹⁴⁰ SNDS is a database that using a unique pseudonymized identifier merges information from the *Système National d'Information Inter Régimes de l'Assurance Maladie* (SNIIRAM) containing all reimbursed claims from all French health care insurance schemes, the *Programme de médicalisation des Systèmes d'information* (PMSI) gathering hospital-discharge summaries from the French public and private hospitals, and the *Base de Causes Médicales de Décès* (BCMD) which is the national death registry managed by the

CépiDc.¹⁴¹ A 1/97th representative and permanent sample of this database, counting now around 800 000 subjects, is also available: the *Échantillon Généraliste de Bénéficiaires* (EGB). Both EGB and SNDS contains individual pseudonymized information on:^{140, 142}

- General characteristics: gender, year of birth, affiliation scheme, area of residence;
- Date and cause of death for those concerned
- Long-term disease registration (LTD – ALD: *Affection longue durée* in French) with the corresponding diagnostic code according to the International classification of Diseases, 10th revision (ICD10). LTD Registration is obtained at the request of a patient's GP to obtain the full reimbursement of expenditures related to a chronic diseases or expensive treatments. Registration must be validated by a health insurance's physician.
- Outpatient reimbursed healthcare expenditures: visits to GP and specialists, medical and diagnostic procedures, nursing acts, physiotherapy, laboratory tests, drugs with dosage and number of boxes dispensed, medical devices, transports, sick leaves, *etc.* For each expenditure, associated costs, prescriber and caregiver information (specialty, private/public practice) and the corresponding dates are provided. However, neither medical indication nor result are recorded;
- Inpatients details: ICD10 diagnostic codes resulting from hospital discharge summaries with the date and duration of the hospital stay, the medical procedures performed, and the related costs. Three different kind of diagnostic codes can be found:
 - Primary diagnosis is the health problem that motivated the admission in the hospital. It is determined at hospital discharge.
 - Related diagnosis can exist only if the primary diagnosis is a medical procedure with a code Z of the ICD10 classification (e.g. chemotherapy session), and indicates the pathology at the origin of procedure.
 - Associated diagnoses correspond mainly to underlying chronic diseases increasing the cost of the patient management.

In case a patient visits different medical units, units' summaries are conserved. Drugs included in the diagnosis related group cost are not captured. However, expansive drugs (*i.e.* the one charged in addition to the group cost) are.

Because data stored in the SNDS are pseudonymized and not anonymized, a re-identification risk remains. Thus, access to SNDS is carefully regulated. The *Institut National des Données de Santé* (INDS) is in charge of it. All the data requests are reviewed by the *Comité d'expertise pour les recherches, les études et les évaluations dans le domaine de la santé* and the French data protection commission (CERREES), an expert committee in charge of assessing the global scientific quality of the project, including the consistency with the data requested. In a second stage the demand is forwarded to the French data protection commission (*Commission Nationale de l'Informatique et des Libertés* - CNIL), which is the only one legally allowed to issue the data processing authorization. Works on data can be then conducted through a remote secured access on CNAM platform "*le portail*", or data extractions can be provided to research teams that have an homologated secured IT framework.¹⁴³

Data from SNIIRAM are uploaded to the SNDS throughout the year. It is admitted that a lag of around 6 months is required to catch 90% of the dispensings. PMSI is uploaded in one time, at the end of the following year. Hence, we consider complete data of year Y are available in January of the year Y+2. This lag of around 13 months allows to follow drug consumption and potential related issues in quite a small delay, some say "almost in live".

1.4.2.3. EHR

Some private EHR suitable for pharmacoepidemiology research (and so, potential signal detection) are available in France: the Longitudinal Patient Database and the Disease Analyzer.^{144, 145} Both of them were owned by IMS Health™, now forming part of IQVIA™. They are fed by data proceeding from patients charts of GP samples (around 1 200 GP each). Elements such as symptoms, diagnostics, prescriptions, demographics, risk factors, laboratory test results as well as GP profile are captured.¹⁴⁶

On the public sphere, in 2008, the George Pompidou European Hospital, an 890-beds university hospital in Paris, was one of the first hospital to set-up a clinical data warehouse gathering 1.2 million patient records to research purposes. After anonymization, laboratory results, drug prescriptions, clinical observations, medical procedures codes and some free text reports were charged into a I2B2 data scheme.

^{147, 148}

In 2017, CNIL allowed the construction of the data warehouse of the *Assistance Publique – Hopitaux de Paris*¹⁴⁹ (AP-HP), which involves 39 hospitals in Paris area, including the former one, for a total of 20 700 beds and around 28 000 physicians and residents.¹⁵⁰ At the end of 2018, this warehouse contained data from 8.8 millions patients. So far, on the 33 projects having received an approval for using EHRs from AP-HP (also called “Greater Paris University Hospitals Clinical Data Warehouse”), only one focused on drug misuse. AP-HP data framework example has recently led other hospitals to set-up/enhance their own data warehouse (e.g. Montpellier, Rennes, Nantes, etc.).¹⁵¹

1.4.3. Safety signal management

In recent years, automated safety signal detection has been implemented in BNPV using a variant of the Gamma Poisson shrinker method, a Bayesian approach of disproportionality analyses.^{152, 153} Although this approach may contribute to signal generation, the impossibility to quantify the potential risk and the high number of false positives generated remain major issues.^{5, 154, 155}

In the frame of DRUG-SAFE platform and as part of his thesis works, Arnaud M. implemented sequence symmetry analysis design to EGB data, developed a prioritization algorithm and assess the full system through one drug class. This approach allowed to detect already known signals as proof as good performances of the methods, and also one previously unknown signal, showing its potential interest in routine detection. However, further research is needed to fully validate this system across other drug classes.^{156, 157}

Since 2014, the CNAM and the *École Polytechnique* have been collaborating together to develop data-mining approach to enjoy the wealth of SNDS data. Their works consist in designing a new data model, “flattening” data to made them easily accessible, and to develop an algorithm based on SCCS for adverse event screening.^{158, 159} Pilot stages have been completed, tools are currently being implemented in a raw extraction of the SNDS but no results have been published so far.

Whatever the source, once a signal is generated, it needs to be confirmed, usually through an expert board or/and a literature review, and assessed. Risk evaluation of suspected drugs is conducted on the behalf of the ANSM by platforms with

pharmacoepidemiology skills such as DRUGS-SAFE and EPI-PHARE through *ad-hoc* studies.¹⁶⁰⁻¹⁶²

2. THE ALCAPONE PROJECT

2.1. Background and thesis context

Frequent adverse events or events resulting from a widely used drug are most of the time already known. Hence, signal generation takes place in a context where exposure and outcomes are either rare or very rare, for instance in the first months of a drug launch. This is particularly true since health authorities ask for real-world safety proof even earlier in product lifecycles, as reflected by the implementation of the risk management plans.¹⁶³ As discussed earlier, one of the answers provided to this need of power at the international level was the development of research networks using distributed network approaches to set up source population of significant size. However, such approaches are not without problems. Heterogeneity resulting from pooling data sources covering different populations remains a potential problem.¹⁶⁴ Although some consortia try to increase homogeneity using a CDM, the consecutive loss in data details (e.g. vocabulary translation) do not help much to overcome the discrepancies resulting from the measured and unmeasured confounders present in the original source populations.

SNDS is an exception in the international database landscape. It combines the large size of some US claims databases to the completeness and the quality of the data available in the European Nordic countries. Hence, it is often seen as the largest homogeneous health database worldwide, enabling studies on a very large population without heterogeneity issues. However, SNDS data were not as available as now at the time of EU-ADR and OMOP experiments, and no empirical assessment of signal generation methodology were conducted on it.

In 2014, a letter of intent was submitted in the frame of a national call for projects about healthcare performances organize by the French Ministry of Health (*Programme de recherche sur la performance du système des soins – PREPS*), the ALCAPONE project: Alert generation using the case-population approach in the French National healthcare databases. ALCAPONE was approved (PREPS 2014, 0635) and was initiated in 2016. The main idea was to empirically compare signal detection methods in the SNDS to find the most suitable for each one of the four health outcomes of interest (HOI) included: acute liver injury (ALI), acute kidney injury (AKI), myocardial infarction (MI), and upper-gastrointestinal bleeding (UGIB). To do so, pharmacoepidemiological approaches were considered. Because they provide risk estimates, they enable to quantify the risk at the same time they detect it,

switching from signal detection to “alert generation”. By law, only extractions of the SNDS are available for research purposes. The easiest way to fulfill our drug screening perspective in light of this technical limitation was to deal with case-based extractions (one by HOI), *de facto* excluding the possibility to run cohort-based approaches. SCCS, CC and CP approaches were chosen. The first two because OMOP experiment provided a sort of benchmark and they were available through R package in the OHDSI library,^{126, 165} and the case-population because precisely, it has never been properly tested. Case-crossover was not retained because of the similarity with SCCS.

The present thesis relies on the ALCAPONE project. Works achieved range from protocol and statistical analyses plan redaction, to results interpretation, including data management and statistical analyses stages. For the last two, technical work was supported by the statisticians and the data management team of the Bordeaux PharmacoEpi platform. Results generated throughout ALCAPONE progress were regularly presented to national and international conferences (Appendix 1). Some of these presentations are introduced here. A total of five publications are planned to cover the overall project, one focusing on the methodology and one for each HOI. The methodology manuscript depicting the overall process of the project is presented in the next pages. UGIB and ALI manuscripts are then developed to enrich general results and to give practical applications.

2.2. Transparency, scientific independency and legality

The ALCAPONE project was fully funded by the French Ministry of Health (PREPS 2014, 0635) and implemented by Bordeaux PharmacoEpi platform. Scientific aspects were supervised by an independent scientific committee. ALCAPONE project is present in the European Union electronic Register of Post-Authorization Studies¹⁶⁶ (EU PAS register) under number 13031 and fulfills European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Seal requirements¹⁶⁷, meaning that researchers follow the provision of the ENCePP Code of Conduct in its entirety.¹⁶⁸ Award letter for ENCePP seal is available in Appendix 2.

As a research team affiliated to the *Institut National de la Santé et de la Recherche* (INSERM), the Bordeaux PharmacoEpi platform is allowed to directly access the EGB for academic purposes upon declaration. Moreover, the approval for SNDS extraction was received from the *Institut des données de santé* (approval #187,

Appendix 3), which was in charge of the scientific evaluation of the submitted protocol at the time. CNIL approved SNDS data processing 10 months later, in March 2017 (decision DE-2017-040, Appendix 4).

2.3. Research question and objectives

Considering the previous works conducted by OMOP and OHDSI, ALCAPONE was set up to determine to what extent, the French Healthcare databases – the SNDS and the EGB – were suitable to perform drug safety signal detection, what the preferred methodology was to identify ALI, AKI, MI and UGIB related risk, and how accurate were the estimates generated.

To do so, the following main objectives were defined:

- To develop the CP approach in the SNDS
- To assess the performances of SCCS, CC and CP in the SNDS for drug safety signal detection based on the OMOP reference set

Derived specific objectives were:

- To describe qualitative changes required for the adaptation of the OMOP reference set to SNDS
- To assess the feasibility of the project using the EGB
- To apply SCCS, CC and CP to SNDS and compare their performances

2.4. Methodology for the assessment of case-based methods in the SNDS

ALCAPONE involves 3 different pharmacoepidemiological approaches, 4 distinct HOIs, and 136 drug controls for a total of 546 drug-outcome pairs. To have the opportunity to present the overall project design and to discuss in depths the methodological choices, we chose to publish the research protocol separately. Herewith, the manuscript below deals with the methods developed and applied to conduct ALCAPONE assessment and validation process, the rationale for important aspects of the considered designs, the difficulties encountered, and the preliminary results. It prepares future papers exploring systematically the different safety domains.

2.4.1. Submitted article: Empirical assessment of case-based methods for drug safety alert identification in the French National Healthcare System database (SNDS): Methodology of the ALCAPONE project

Nicolas H. Thurin, Régis Lassalle, Martijn Schuemie, Marine Pénichon, Joshua J. Gagne, Jeremy A. Rassen, Jacques Benichou, Alain Weill, Patrick Blin, Nicholas Moore, Cécile Droz-Perroteau, Pharmacoepidemiology and Drug safety (1st submission)

**Empirical assessment of case-based methods for drug safety alert
identification in the French National Healthcare System database (SNDS):
Methodology of the ALCAPONE project**

Short title: Methodology for the assessment of case-based methods in the SNDS

Nicolas H Thurin^{1,2,*}, Régis Lassalle¹, Martijn Schuemie^{3,4}, Marine Pénichon¹,
Joshua J Gagne⁵, Jeremy A Rassen⁶, Jacques Benichou^{7,8}, Alain Weill⁹, Patrick
Blin¹, Nicholas Moore^{1,2}, Cécile Droz-Perroteau¹

¹Bordeaux PharmacoEpi, INSERM CIC1401, Université de Bordeaux, Bordeaux, France.

²INSERM U1219, Université de Bordeaux, Bordeaux, France.

³Epidemiology Analytics, Janssen Research and Development, Titusville, NJ, USA.

⁴Observational Health Data Sciences and Informatics (OHDSI), New York, NY, USA.

⁵Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA USA.

⁶Aetion, Inc., New York, NY, USA.

⁷Department of Biostatistics and Clinical Research, Rouen University Hospital Rouen, France.

⁸INSERM U1181

⁹Caisse Nationale de l'Assurance Maladie, Paris, France.

* Corresponding authors

nicolas.thurin@u-bordeaux.fr
Bordeaux PharmacoEpi
Université de Bordeaux
case 41 - 146 rue Léo Saignat
33076 Bordeaux cedex - France

Key words: calibration, case population, claims database, SNDS, method assessment, pharmacoepidemiology, pharmacovigilance

Key points:

- French National Healthcare System database (SNDS) is a powerful source for drug safety alert generation
- ALCAPONE aims to establish in the SNDS gold standard calibrated methods to measure associations between drugs and specific health outcomes of interest (HOI): acute liver injury, acute kidney injury, myocardial infarction, and upper gastrointestinal bleeding
- The performances of case-control, case-population and self-controlled case series designs will be compared using positive and negative drug controls
- The best performing method (gold standard) can be used to validate drug safety signals obtained elsewhere or to directly generate new HOI-related alerts.

Word count excluding abstract, tables, figures and references: 4211

Statement about prior postings and presentations, name(s) of any sponsor(s) of the research contained in the paper, along with grant number(s):

ALCAPONE (Alert generation using the case population approach) methodology and related results were presented during the 33rd and 35th International Conference on Pharmacoepidemiology & Therapeutic Risk Management and at the 1st and 2nd OHDSI European Symposium.

The ALCAPONE project is funded by the French Ministry of Health (PREPS 2014, 0635).

ABSTRACT

227/250 words

Objectives: The ALCAPONE project aims to assess empirically in the National healthcare data system (SNDS) case-based designs for alert generation related to 4 health outcome of interest (HOI).

Background: SNDS, the French nationwide healthcare system database covering 99% of the French population, provides a potentially valuable opportunity for drug safety alert generation.

Methods: ALCAPONE uses a reference set adapted from OMOP and EU-ADR, with 4 HOI - Acute Liver Injury (ALI), Myocardial Infarction (MI), Acute Kidney Injury (AKI), and Upper Gastrointestinal Bleeding (UGIB) – and positive and negative drug controls. ALCAPONE consists of 4 main stages: data preparation to fit the OMOP Common Data Model and select the drug controls; detection of the selected controls via 3 case-based designs: case-population, case-control, and self-controlled case series, including design variants (risk window, adjustment strategy, etc.); assessment and comparison of design performance (area under the ROC curve, mean square error, etc.); selection of the best design variants and their calibration for each HOI.

Results: Over 6 years, considering specific definitions, 5 152 cases of ALI, 12 317 KI, 304 369 MI and 139 172 UGIB were identified. The number of detectable drugs ranged from 61 for MI to 25 for ALI providing enough power for a valid assessment of methods.

Conclusions: If successful, ALCAPONE will provide a better understanding of the performance of different case-based designs for signal identification in SNDS.

1 INTRODUCTION

Spontaneous reporting has historically been the mainstay approach for drug safety alert generation.¹ The increasing accessibility of population databases has brought new opportunities to identify drug-related alerts, using very different methods from those used to analyze spontaneous reporting data.

Unlike spontaneous report data, longitudinal healthcare databases, such as claims databases, typically accrue automatically and prospectively. Information about events and exposures are collected independent of any research project and are therefore largely unaffected by recall biases.² Electronic health records (EHRs) may also contain a large numbers of time-stamped medical records from routine clinical practice.³ As a result, such data represent a valuable source of information for safety signal strengthening and validation. Furthermore, because they capture the very first prescriptions of new drugs prospectively, these databases have potential for early detection of drug safety signals. Methods used to explore longitudinal observational data can be divided into a few main categories based on entry in the study through exposure (cohort-based designs) or events (case-based designs). Other general design options can be considered, especially concerning control groups that may range from self-controlled methods to population-wide controls.⁴

A number of initiatives have been undertaken to develop methods and systems for safety signal identification and evaluation in longitudinal healthcare databases. In Europe, the Exploring and Understanding Adverse Drug Reactions project (EU-ADR) has combined data from several databases from different countries and settings to demonstrate the capacity to assess drug safety signals.⁵⁻⁷ In the United States, the Food and Drug Administration (FDA) has developed the Sentinel System to monitor the safety of regulated medical products with a network of healthcare databases.⁸ The Observational Medical Outcomes Partnership (OMOP) performed an empirical assessment of analytical methods for signal identification in healthcare data standardized to a Common Data Model (CDM).⁹⁻¹⁶ The performance of various methods was evaluated in five large 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 drug-event pairs across four health outcomes of interest (HOI): acute liver

injury (ALI), myocardial infarction (MI), acute kidney injury (AKI), and upper gastrointestinal bleeding (UGIB). These four events have led to drug withdrawals from the market.¹⁷ Positive pairs represented known drug-event associations, and negative pairs represent drug event pairs with no known association. The same methods were also applied to six European EHRs or claims databases from EU-ADR covering nine million persons, using the same test cases.^{5, 6} The OMOP CDM was then improved and used for the development of updated analysis packages maintained by the Observational Health Data Sciences and Informatics (OHDSI) consortium.¹⁸

France has a large nationwide longitudinal claims and hospital database -- the National healthcare data system (*Système National des données de Santé*, SNDS) -- which currently includes about 99,9% of the French population (66.6 million persons), from birth or immigration to death or emigration. It includes all reimbursed medical and paramedical encounters, including among others all dispensed drugs, hospital-discharge summaries and dates of death.¹⁹ Because SNDS was not available at the time, it was not included in the European replication of the OMOP experiment.

The case-population design (CP), a case-based approach where drug exposure in cases is compared to aggregate data from the entire population^{4, 20-22} was not included in OMOP experiment. This approach, made possible by the complete national coverage by SNDS, seems well suited to an alert generation environment. Case-population provides absolute event rates in addition to the relative risks or odds ratios also found from other case-based approaches⁴, providing another measure of risk to help decision-making. As CP neither involves control selection nor complex calculation, it is easy to implement and the results generated are comprehensible to non-specialists.

This paper presents the ALCAPONE project (Alert generation using the case population approach), which is funded by the French Health Ministry (PREPS 2014, 0635) and aims to empirically compare and calibrate within SNDS case-based methods (including the CP design) using the OMOP methodology and the OHDSI environment, *i.e.* a collection of positive control and negative control drug-outcome pairs across all four HOI (ALI, AKI, UGIB and MI).

2 DESIGN AND RESEARCH PLAN

ALCAPONE tests case-based approaches using case data from the SNDS associated with a drug control reference set.

The ALCAPONE process consists of 4 main stages (Figure 1): (I) the preparation of SNDS data to fit the OMOP CDM and the selection of the detectable positive and negative drug controls; (II) the application of 3 case-based designs: CP, case control (CC), and self-controlled case series (SCCS), including design variants for each method; (III) the assessment and comparison of design performance; and (IV) the identification of the best design variants and their calibration.

2.1 DATA SOURCE AND DATA MAPPING PROCESS

France has a universal single-payer health care system covering most outpatient medical expenses including drugs, medical devices, lab tests, medical visits and paramedical activities, as well as hospital stays, and medical leaves. Claims are collected in a single database called SNIIRAM (*Système National d'Informations Inter-Régimes de l'Assurance Maladie*). Hospital discharge summaries from public and private hospitals, including International Classification of Disease, 10th Edition (ICD10) primary and secondary codes as well as procedures and other information, are available through the PMSI (*Programme de médicalisation des systèmes d'information*). SNIIRAM, PMSI and the National Death registry (CépiDC) are linked in a global system called *Système National des Données de Santé* (SNDS). SNDS covers over 99% of the 66.6-million-person French population, resulting in one of the largest nationwide claims and hospital databases in the world. A 1/97th representative and permanent sample of this database is also available: the EGB (*Echantillon Généraliste de Bénéficiaires*).¹⁹

All recorded hospital admissions for ALI, AKI, MI and UGIB between 01/01/2009 and 12/31/2014 were extracted from SNDS. The resulting datasets were managed according to the Bordeaux PharmacoEpi standardized operating procedures to generate 14 datasets including socio-demographics, medical visits, chronic medical conditions, hospitalization diagnoses, drug dispensing, lab tests, and outpatient and inpatient procedures. This dataset was then mapped to the OMOP CDM v5.0.1.

Durations of drug exposures were estimated at the ATC level according to drug dispensings and medians of treatment durations.

2.2 REFERENCE SET CONSTRUCTION

The reference set used to assess the different methodological approaches includes the four HOI (ALI, AKI, UGIB and MI) with a collection of positive and negative drug controls for each HOI. Positive controls are drugs with a known association with the outcome. Negative controls are drugs with no evidence to support causal association with the outcomes. This set results from a fusion of the OMOP and EU-ADR reference sets, adapted to drug availability on the French market.^{23, 24}

2.2.1 Health outcome of interest

HOI were identified from hospital discharge summaries ICD10 codes. Specific and sensitive definitions were developed for each HOI.

AKI and UGIB were selected using the ICD10 codes corresponding to the ICD9 codes used in the original OMOP exercise. Patients with previous renal transplantation, metal intoxication or specific kidney diseases were excluded from the AKI population.

Relevant codes resulting from *ad-hoc* SNDS studies were chosen to enrich OMOP codes for ALI and MI.²⁵ The specific definition applied to MI covered acute transmural MI and acute subendocardial MI (STEMI, NSTEMI). The sensitive definition also included unstable angina. The corresponding codes have been used in several studies in the SNDS.²⁶⁻²⁹

For AKI, UGIB and MI, patients with more than 15 cumulative hospital days in the month preceding the outcome onset were excluded, because in-hospital exposure to common drugs is not ascertainable.

The patient selection process for ALI and the corresponding inclusion and exclusion criteria were replicated from the EPIHAM study.^{30, 31} They were defined as “toxic liver disease” and “acute and subacute hepatic failure”. Patients with codes related to liver injury resulting from other causes than potential drug toxicity were excluded (e.g. chronic viral hepatitis, alcoholic liver disease, *etc.*). The full list of inclusion and exclusion criteria is available in Appendix 1.

2.2.2 Positive and negative controls

The original OMOP and EU-ADR reference sets were merged and screened at the ATC level to rule out drugs that are not marketed or not reimbursed in France. Only systemic forms of the drugs were conserved as positive controls. The control selection process is presented in Figure 1. For ALI, the original OMOP reference set included 118 controls, of which 75 were available and reimbursed on the French market. Adding 2 positive controls and the 4 negative ones from the EU-ADR reference set, a total of 58 positive and 23 negative ALI controls were retained. Twenty-two positive and 36 negative controls were identified for AKI. The MI reference set was formed from 28 positive and 42 negative controls, including 7 from EU-ADR. Twenty-two positive and 42 negative controls were retained for UGIB. The ALCAPONE reference set thus included 139 distinct drugs and 273 drug-outcome pairs. Based on the number of exposed patients in the relevant HOI population, we restricted the drug-outcome pairs to those with a sufficient power to detect a minimum relative risk of 1.30.³²

2.3 METHODOLOGICAL APPROACHES

Three case-based approaches – SCCS, CC and CP – were applied with different settings to identify the variants with the best performances. To optimize machine time processing the MI, UGIB and AKI population were sampled at respectively 1/20th, 1/10th, and 1/3rd proportions before screening. The best variant of each approach was then replicated in the whole population.

The characteristics of the servers and software and package versions used in this experiment are available in Appendix 2.

2.3.1 Self-controlled case series

The self-controlled case series was tested using the OHDSI *SelfControlledCaseSeries* R package. This consists of comparing each case to itself: the event rate during periods exposed to the drug of interest is compared to the event rate during unexposed periods. The self-controlled case series assumes that outcomes associated with the target condition arise from a non-homogeneous Poisson process.¹⁵ This design requires two additional strong assumptions: (i) outcomes must not influence the occurrence or timing of subsequent exposure; and

(ii) outcomes must not influence the duration of the observation period. Violation of these assumptions will induce bias.³³ To address these and other potential biases, further settings of the OHDSI R package were tested resulting in 96 SCCS variants. The minimum duration of a subject for inclusion in the analysis was set to 182 days. Two different risk windows were considered. In order to address potential indication bias, three different pre-exposure windows were tested: 0, 7 or 30 days. The model was applied to all occurrences of the HOI or only to the first one. Some variants also adjusted for age, seasonality and for multiple drugs.

2.3.2 Case-control

Case-control methods compare the distribution of exposure prior to outcomes in cases with the distribution in patients at risk for the outcome.¹⁰ We used the OHDSI *CaseControl* R package. As explained above, controls were selected from the EGB subjects (SNDS 1/97th sample) that did not present with the HOI. The age of each potential control was calculated for each year of the data extraction sample (6 ages in total). Controls were matched with cases according to their gender and their age at index-date. Each selected control was given the same index date as their corresponding case. The number of controls per case was set to 2 or 10 according to the variant. Unmatched cases were removed. To be included, cases and controls must have had at least 182 days of observation prior to their “index date”. When only the first occurrence of the HOI was considered, the patient was excluded if it occurred within the 182 days of the washout period. To address protopathic and confounding by indication a lag period of 7 or 15 days was introduced before the event onset in some variants. The risk windows applied was of 7, 30 or 60 days.

2.3.3 Case-population

As in the CC design, the exposure distribution among cases is compared to the distribution among controls. The novelty of the method is that the control group consists of the complete population, which increases statistical precision. In our application, exposure distribution for the complete population was extrapolated from the EGB over the study period using 1) an age and sex stratified extrapolation, and 2) a raw extrapolation (*i.e.* no stratification on age or sex). To be included in the case group or the aggregated control data, a patient had to be enrolled in the database for

at least 182 days. Risk windows, exclusion periods and outcome selection were defined in same way as for CC. Two approaches were tested: (1) a count data approach, considering the number of patients exposed or not in the control population; and (2) a person-time approach, considering the person-time units of exposure in the reference population. (e.g. person-months). Two measures of associations were calculated: the case population ratio (CPR)⁴, and the predicted relative risk (pRR).²² In the CPR calculation we assumed that the number of cases and the exposure rate are so small that the overall number of cases and the overall population can respectively approximate the number of unexposed cases and the unexposed population. In the pRR calculation the proportion of unexposed persons in the case group and in the population are not disregarded. In addition, CP allows the measure of relative risks based on per-patient exposure or per patient-time exposure. The 80 CP variants were executed using an in-house program developed in R. A part of the analyses was replicated in SAS[®] to ensure internal validity.

2.4 METRICS

To assess the ability of the methods and their variants to distinguish between positive and negative controls, the area under the receiver operating characteristic curve (AUC) was estimated. Sensitivity and specificity were also estimated. These elements give information about the discriminating power of the approaches, but we were also interested in the accuracy of the magnitude of the effect estimated. Even if the true effect size of positive controls remains unknown, we can assume that the log relative risk of the negative controls is zero. Using this assumption, we computed for each method and its specific settings the mean square error (MSE), which is the average squared difference between the log relative risk and zero. The smaller the MSE is, the better the estimation is. The “coverage probability” is the frequency over many replications (one for each negative control) that the 95% confidence interval contains the true RR: 1. In the case of an unbiased estimator we would expect the coverage probability to be 95%.

2.5 CALIBRATION

Calculation of p-value in traditional significance testing relies on the use of the theoretical null distribution. In ALCAPONE, once the best performing design variant

for an HOI was identified and replicated in the non-sampled population, the *EmpiricalCalibration* R package was used to derive an empirical null distribution from the observed effect estimates for the negative controls, and to generate a “calibrated” p -value.^{34 35} This calibrated p -value takes into account the random error (as the traditional p -value does), but also the systematic error distribution inherent to the application the SNDS.

3 RESULTS

3.1 HEALTH OUTCOMES OF INTEREST

5 152 patients were identified in SNDS based on the specific ALI definition. The sensitive definition identified 355 more cases.

During the study period, 304 369 patients experienced a myocardial infarction (specific definition). When adding unstable angina, the number of events identified doubled, from 354 109 to 717 920.

The specific AKI definition identified 12 317 patients, the sensitive definition brought the number of outcomes to 89 186.

For UGIB, 156 057 and 204 442 outcomes (specific and sensitive definitions, respectively) were identified during the study period, corresponding to 139 172 and 178 384 patients, respectively.

Numbers of included patients are presented in Table I.

3.2 POSITIVE AND NEGATIVE CONTROLS

Table II shows the number of positive and negative controls with a minimum detectable relative risk < 1.30 , *i.e.* those for which the SNDS is powerful enough to detect such an association, based on exposure. HOIs with the largest numbers are the one with the most detectable controls: for MI sensitive definition, 64 detectable controls out of 70 total potential positive and negative controls, and for UGIB 59 out of 64. The sampling of these large populations for computational purposes was accompanied by a decrease in power and in number of detectable drugs. The ALI specific definition enabled 18 detectable positive controls out of 58 and 7 negatives out of 23.

4 DISCUSSION

In this large-scale analysis, we empirically compared three case-based approaches for alert generation for each of four HOIs. Results corresponding to the methods comparison and calibration for ALI, AKI, MI and UGIB will be presented in forthcoming papers. Because the central idea of the ALCAPONE study is to develop alert generation based on pre-specified HOIs across a large number of drugs, it was easier to perform an extraction for the four HOI rather than one for each drug examined, which is why we focused only on case-based methods. The exposure of these cases to different drugs, here the reference sets, is compared to the exposure in selected controls populations, from the same patient in SCCS, to matched patients in CC, and to unmatched and unselected controls in the CP approach. We need to identify cases and controls from the same population, and to this end we can access either the main SNDS database (66 million persons) or its 1/97th sample, EGB, over one or several years. Though in some cases it might be sufficient to identify cases in the EGB (e.g., myocardial infarction, with about 500 cases per year), for rarer events, such as ALI, the only way to accrue sufficient numbers of cases for meaningful analyses is to access the full nationwide database over several years. However, controls do not need to be taken from the whole national database. EGB, with 700 000 persons, should be enough, providing two-step sampling: first, a first random sampling of the nationwide SNDS database to its 1/97th representative sample, EGB; then either using that sample entirely in the CP approach, or selecting specific controls in the CC.

The choice to include CP to the methods screened arose from the observation that in the alert generation environment, exposure in the population and cases may be both very rare, especially in the early stages of marketing a drug. In this context, the low level of exposure could make matching or adjustment impossible in a classical case-control design. The case population is a potential solution as the control group is the whole population. The case population approach has been compared to several other study designs and results show that the CPR computed were of the same order of magnitude as the corresponding relative risks estimated with a standard control group.^{22, 36-39} A theoretical analysis showed that the rarer the exposure and the outcome rate, the better the CPR approximates the actual RR.⁴ In addition, because

all cases are identified and there is an analysis of the whole-country exposure, actual event rates per patient and per patient-year can be quantified.

Originally, we planned to conduct this project using only the controls of the OMOP reference set, with a threshold to consider a control as detectable set to 1.25 as in the original OMOP experiment. The OMOP set resulted from work conducted in the USA and counted 165 positives and 233 negative controls. Of these, only 120 positive and 126 negative control drugs are on the French market. In small populations, as with the specific ALI definition, only two drug controls were considered detectable with these parameters. In addition to the population size, this could be explained by differences between the US and French prescription patterns. We decided to add control drugs from EU-ADR, which includes drugs more often prescribed in the French market. We also increased the minimum detectable relative risk threshold from 1.25 to 1.30. These adjustments allowed us to reach 7 detectable negative controls for the 5 152 patients meeting the ALI specific definition. Almost all the control drugs tested were identified as detectable in the large HOI populations: in the UGIB sensitive definition population, 59 controls out of 66 were considered detectable.

Even though the selection process of the control drugs was carefully described^{23, 24}, one of the limitations of our project is the possibility that negative controls actually form drug-outcome pairs with a causal relationship. A prior investigation found that 17% of the OMOP negative controls are misclassified or potentially misclassified but without any assertion of the causal relationship with the HOI.⁴⁰ Confounding by indication (or contraindication) and protopathic bias could also affect the findings. Although there are many technical options to reduce such biases, a clinical assessment of the negative controls wrongly detected as positive will likely be necessary.

As detailed in the Methods section, we have many drug-outcome pairs and many variants of the method designs: 20 CC, 96 SCCS and 80 CP. If we consider the sensitive MI definition, before sampling, this represents a total of 12 544 analyses to run across 558 538 patients. When we consider all the drug-outcome pairs across all HOI definitions, 73 696 analyses were forecast in ALCAPONE. The large number of cases is not an issue for CP, but requires substantial computing time for the complex calculations involved in some designs. For data protection purpose, French law

assigns SNDS data extraction to a physical research structure and the data are not allowed outside of this structure. As a result, use of external servers such as the ones maintained by Amazon Web Services is not possible. Running all the analysis with our IT capacity would have called for years of machine processing time. So, we chose to sample our biggest data extractions in order to obtain more tractable populations. This process also simplified the case-control selection of controls: even for only one control per case, working with the raw population would have required reselecting the same controls several times in EGB, or going to SNDS to identify them. Extracting data from SNDS involves a complex administrative approach involving third-party access to the full data.

Data extractions were transformed and loaded into the OMOP CDM. The OHDSI environment was chosen to take advantage of a scalable framework with an open source and transparent toolbox. In the first version of the protocols we had planned to assess more design variants. However, during the feasibility and testing phases we experienced some failures with some of the packages and noticed that some package options gave different results according to the machine they were run on. We thus reduced the number of variants to only the stable ones. Even though open-source programs are often considered very reliable because they are peer-reviewable, this maturity takes time and internal validity tests should be systematically conducted when a new package or update is made available and executed in a new IT framework.

ALCAPONE project aims to give rise to calibrated gold-standard method, but so far, we have not addressed calibration. As detailed in the Methods section, metrics were calculated based on the assumption that negative control true effect sizes are 1. However, in real world settings this is not always the case. The distribution of the negative controls point estimates reflects the systematic error distribution inherent to the database and the considered design, *i.e.* the background noise of the database, assuming a similar confounding structure for a drug of interest and all of the negative control drugs, on average. In the environment of alert generation, knowing this distribution could help to reduce the number of false positives when screening drugs.^{34, 35}

While analyses of spontaneous reports often consider associations among all drugs and all HOIs, signal detection approaches in longitudinal healthcare databases may be better

tailored to specific drugs and/or HOIs. In ALCAPONE, we focus on four HOIs that are among the most common safety reasons for withdrawing drugs from the market. We aim to understand the best performing methods for each outcome in SNDS to inform future drug safety surveillance. Signal detection in such large databases could be used to validate drug safety signals obtained through other mechanisms (*e.g.* spontaneous reporting, social media) or to directly generate novel HOI-related alerts. The resulting point estimates can be assessed in view of the SNDS background noise and will provide a quantification of the potential risk. The same methodology could be applied to other HOIs such as the “critical terms” list identified by the Uppsala Monitoring Centre, or reported from the analysis of event-related drug withdrawals.¹⁷ From there a systematic specific alert monitoring system could be set up, for the systematic detection of new drug-related alerts concerning these preselected HOI, over time.

ACKNOWLEDGMENTS

ALCAPONE project is funded by the French Ministry of Health (PREPS 2014, 0635).

REFERENCES

1. Moore N. The past, present and perhaps future of pharmacovigilance: homage to Folke Sjoqvist. *Eur J Clin Pharmacol* 2013, 69 Suppl 1:33-41.
2. Martin-Latry K, Begaud B. Pharmacoepidemiological research using French reimbursement databases: yes we can! *Pharmacoepidemiology and drug safety* 2010, 19:256-265.
3. Schuemie MJ, Coloma PM, Straatman H, Herings RM, Trifiro G, Matthews JN, Prieto-Merino D, Molokhia M, Pedersen L, Gini R, et al. Using electronic health care records for drug safety signal detection: a comparative evaluation of statistical methods. *Med Care* 2012, 50:890-897.
4. Moore N, Gulmez SE, Blin P, Lassalle R, Jove J, Theophile H, Begaud B, Larrey D, Benichou J. Relative risks from case-population data. *Epidemiology* 2013, 24:935-936.
5. Trifiro G, Pariente A, Coloma PM, Kors JA, Polimeni G, Miremont-Salame G, Catania MA, Salvo F, David A, Moore N, et al. Data mining on electronic health record databases for signal detection in pharmacovigilance: which events to monitor? *Pharmacoepidemiol Drug Saf* 2009, 18:1176-1184.
6. 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. *Stud Health Technol Inform* 2009, 148:43-49.
7. Trifiro G, Patadia V, Schuemie MJ, Coloma PM, Gini R, Herings R, Hippisley-Cox J, Mazzaglia G, Giaquinto C, Scotti L, et al. EU-ADR healthcare database network vs. spontaneous reporting system database: preliminary comparison of signal detection. *Stud Health Technol Inform* 2011, 166:25-30.
8. FDA. FDA's Sentinel Initiative. Available at: <https://www.fda.gov/Safety/FDAsSentinelInitiative/ucm2007250.htm>. (Accessed 30 Oct)
9. DuMouchel W, Ryan PB, Schuemie MJ, Madigan D. Evaluation of disproportionality safety signaling applied to healthcare databases. *Drug safety : an international journal of medical toxicology and drug experience* 2013, 36 Suppl 1:S123-132.
10. Madigan D, Schuemie MJ, Ryan PB. Empirical performance of the case-control method: lessons for developing a risk identification and analysis system. *Drug safety : an international journal of medical toxicology and drug experience* 2013, 36 Suppl 1:S73-82.
11. 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 safety : an international journal of medical toxicology and drug experience* 2013, 36 Suppl 1:S107-121.
12. 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 safety : an international journal of medical toxicology and drug experience* 2013, 36 Suppl 1:S59-72.

13. Ryan PB, Schuemie MJ, Madigan D. Empirical performance of a self-controlled cohort method: lessons for developing a risk identification and analysis system. *Drug safety : an international journal of medical toxicology and drug experience* 2013, 36 Suppl 1:S95-106.
14. Schuemie MJ, Madigan D, Ryan PB. Empirical performance of LGPS and LEOPARD: lessons for developing a risk identification and analysis system. *Drug safety : an international journal of medical toxicology and drug experience* 2013, 36 Suppl 1:S133-142.
15. 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 safety : an international journal of medical toxicology and drug experience* 2013, 36 Suppl 1:S83-93.
16. Ryan P, Stang P, Overhage J, Suchard M, Hartzema A, DuMouchel W, Reich C, Schuemie M, Madigan D. A comparison of the empirical performance of methods for a risk identification system. *Drug safety : an international journal of medical toxicology and drug experience* 2013, 36 Suppl 1:S143-158.
17. Onakpoya IJ, Heneghan CJ, Aronson JK. Post-marketing withdrawal of 462 medicinal products because of adverse drug reactions: a systematic review of the world literature. *BMC Med* 2016, 14:10.
18. Hripcsak G, Duke JD, Shah NH, Reich CG, Huser V, Schuemie MJ, Suchard MA, Park RW, Wong IC, Rijnbeek PR, et al. Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers. *Stud Health Technol Inform* 2015, 216:574-578.
19. Bezin J, Duong M, Lassalle R, Droz C, Pariente A, Blin P, Moore N. The national healthcare system claims databases in France, SNIIRAM and EGB: Powerful tools for pharmacoepidemiology. *Pharmacoepidemiol Drug Saf* 2017, 26:954-962.
20. Gulmez SE, Larrey D, Pageaux GP, Lignot S, Lassalle R, Jove J, Gatta A, McCormick PA, Metselaar HJ, Monteiro E, et al. Transplantation for acute liver failure in patients exposed to NSAIDs or paracetamol (acetaminophen): the multinational case-population SALT study. *Drug safety : an international journal of medical toxicology and drug experience* 2013, 36:135-144.
21. Moore N, Gulmez SE, Larrey D, Pageaux GP, Lignot S, Lassalle R, Jove J, Pariente A, Blin P, Benichou J, 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. *Pharmacoepidemiology and drug safety* 2013, 22:160-167.
22. Theophile H, Laporte JR, Moore N, Martin KL, Begaud B. The case-population study design: an analysis of its application in pharmacovigilance. *Drug safety : an international journal of medical toxicology and drug experience* 2011, 34:861-868.
23. Coloma PM, Avillach P, Salvo F, Schuemie MJ, Ferrajolo C, Pariente A, Fourier-Réglat A, Molokhia M, Patadia V, van der Lei J, et al. A Reference Standard for Evaluation of Methods for Drug Safety Signal Detection Using Electronic Healthcare Record Databases. *Drug Safety* 2013, 36:13-23.

24. Ryan PB, Schuemie MJ, Welebob E, Duke J, Valentine S, Hartzema AG. Defining a reference set to support methodological research in drug safety. *Drug Saf* 2013, 36 Suppl 1:S33-47.
25. Reich C, Ryan P, Schuemie MJ. Alternative outcome definitions and their effect on the performance of methods for observational outcome studies. *Drug safety : an international journal of medical toxicology and drug experience* 2013, 36 Suppl 1:S181-193.
26. Duong M, Salvo F, Pariente A, Abouelfath A, Lassalle R, Droz C, Blin P, Moore N. Usage patterns of 'over-the-counter' vs. prescription-strength nonsteroidal anti-inflammatory drugs in France. *Br J Clin Pharmacol* 2014, 77:887-895.
27. Bezin J, Pariente A, Lassalle R, Dureau-Pournin C, Abouelfath A, Robinson P, Moore N, Droz-Perroteau C, Fourier-Reglat A. Use of the recommended drug combination for secondary prevention after a first occurrence of acute coronary syndrome in France. *Eur J Clin Pharmacol* 2013.
28. Blin P, Dureau-Pournin C, Lassalle R, Jove J, Thomas-Delecourt F, Droz-Perroteau C, Danchin N, Moore N. 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* 2015.
29. 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. *British journal of clinical pharmacology* 2016, 81:569-578.
30. Gulmez SE, Unal US, Lassalle R, Chartier A, Grolleau A, Moore N. Risk of hospital admission for liver injury in users of NSAIDs and nonoverdose paracetamol: Preliminary results from the EPIHAM study. *Pharmacoepidemiol Drug Saf* 2018, 27:1174-1181.
31. Moore N, Duret S, Grolleau A, Lassalle R, Barbet V, Duong M, Thurin N, Droz-Perroteau C, Gulmez SE. Previous Drug Exposure in Patients Hospitalised for Acute Liver Injury: A Case-Population Study in the French National Healthcare Data System. *Drug Saf* 2018.
32. Armstrong B. A simple estimator of minimum detectable relative risk, sample size, or power in cohort studies. *Am J Epidemiol* 1987, 126:356-358.
33. Whitaker H, Farrington C, Spiessens B, Musonda P. Tutorail in biostatistics: The self-controlled case series method. *Statistics in medicine* 2006, 25:1768-1797.
34. Schuemie MJ, Ryan PB, DuMouchel W, Suchard MA, Madigan D. Interpreting observational studies: why empirical calibration is needed to correct p-values. *Stat Med* 2014, 33:209-218.
35. Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. Robust empirical calibration of p-values using observational data. *Stat Med* 2016, 35:3883-3888.
36. 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, 56:763-767.

37. 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, 43:211-214.
38. Laporte JR, Ibanez L, Ballarin E, Perez E, Vidal X. Fatal aplastic anaemia associated with nifedipine. *Lancet* 1998, 352:619-620.
39. Laporte JR, Vidal X, Ballarin E, Ibanez L. Possible association between ocular chloramphenicol and aplastic anaemia--the absolute risk is very low. *British journal of clinical pharmacology* 1998, 46:181-184.
40. Hauben M, Aronson JK, Ferner RE. Evidence of Misclassification of Drug-Event Associations Classified as Gold Standard 'Negative Controls' by the Observational Medical Outcomes Partnership (OMOP). *Drug Saf* 2016, 39:421-432.

TABLES

Table I. Outcomes and patients included in ALCAPONE, by Health Outcome of Interest (HOI)

	ALI		MI		AKI		UGIB	
	Narrow Def.	Broad Def.						
n (outcomes)	5225	5580	354109	717920	12633	89186	156057	204442
n (patients)	5152	5490	304369	558538	12317	82610	139172	178384

ALI: acute liver injury; MI: myocardial infarction; AKI: acute kidney injury; UGIB: upper gastro-intestinal bleeding.

Table II. Number of detectable controls ($MDRR \leq 1.30$) according to HOI and population sampling

Drug controls	French market Reference set	SNDS							
		raw		1/3 rd sample		1/10 th sample		1/20 th sample	
		Number of detectable controls ¹		Number of detectable controls ¹		Number of detectable controls ¹		Number of detectable controls ¹	
		Narrow definition	Broad definition						
ALI	+	58	18	20					
	-	23	7	7					
MI	+	28	25	26				26	21
	-	42	36	38				20	30
AKI	+	22	17	21	11	20			
	-	36	13	31	10	17			
UGIB	+	22	22	22			19	19	
	-	42	36	37			22	22	

¹ Drug controls with $MDRR \leq 1.30$

ALI: acute liver injury; MI: myocardial infarction; AKI: acute kidney injury; UGIB: upper gastro-intestinal bleeding.

FIGURE

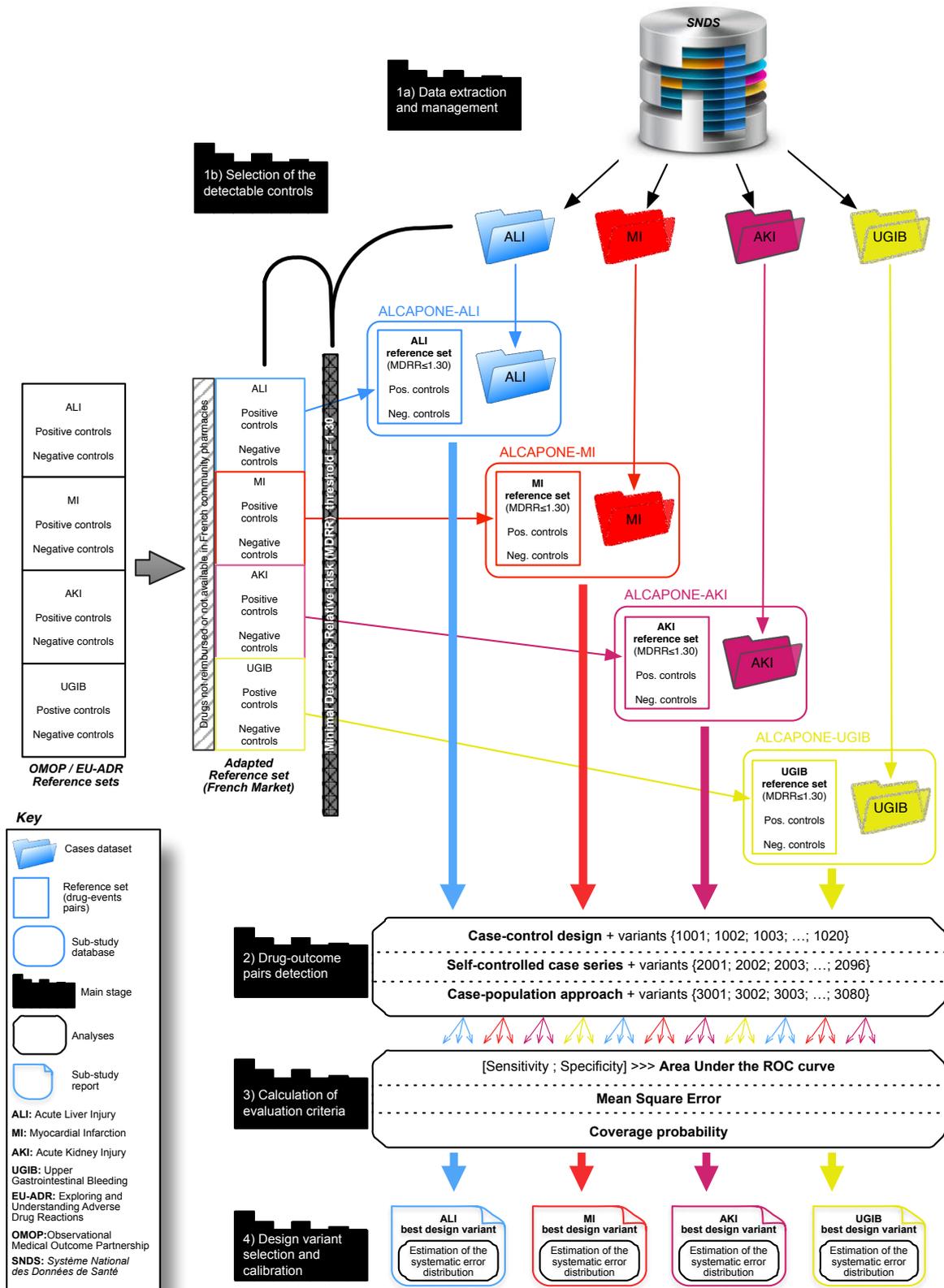


Figure 1. Overall process of the ALCAPONE project

APPENDICES

APPENDIX 1 HEALTH OUTCOME OF INTEREST ICD10 CODES

▪ ALCAPONE- ALI

Inclusion codes for acute liver injury

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

In order to take into account specificities of ALI, specific rules and exclusion criteria will apply:

- Outcomes presenting during hospitalizations with end dates between the 1st and the 30th day before their onset will be excluded;
- Outcomes presenting during hospitalizations with end dates on the same day as their onset, and corresponding hospitalization start date preceding the onset by more than 7 days, will be excluded;
- Hospitalizations starting and ending in a 7 days period prior to the outcome occurrence with concomitant stays separated by less than 1 day will be aggregated (this rule also applies to the hospitalization of interest);
- Outcomes with less than 182 days of observation prior to their onset will be excluded.
- Outcomes presenting with an associated diagnosis, related diagnosis or main diagnosis corresponding to the following codes in the 182 days prior to the index date will be excluded

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
Z94.4	Liver transplant status
T36-T50	Poisoning by drugs, medicaments and biological substances
T51-T65	Toxic effects of substances chiefly nonmedicinal as to source
S36.1	Injury of liver or gallbladder
B15	Acute hepatitis A
B16	Acute hepatitis B
B17	Other acute viral hepatitis
B19	Unspecified viral hepatitis
B24	Unspecified human immunodeficiency virus [HIV] disease
K77	Liver disorders in diseases classified elsewhere
B25.1	Cytomegaloviral hepatitis
K73	Chronic hepatitis, not elsewhere classified
K75.4	Autoimmune hepatitis
O26.6	Liver disorders in pregnancy, childbirth and the puerperium

■ ALCAPONE- MI

Inclusion codes for myocardial infarction

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

In order to minimize information bias:

- Outcomes with less than 182 days of observation prior to their onset will be excluded;
- Outcomes with a total of 15 cumulative days of hospitalization or more in the 30 days prior their onset will be excluded.

▪ **ALCAPONE- UGIB**

Inclusion codes for upper gastrointestinal bleeding

ICD10	Definition	Narrow definition	Broad definition
K25.0	Gastric ulcer ; Acute with haemorrhage	yes	yes
K25.2	Gastric ulcer ; Acute with both haemorrhage and perforation	yes	yes
K25.4	Gastric ulcer ; Chronic or unspecified with haemorrhage	yes	yes
K25.6	Gastric ulcer ; Chronic or unspecified with both haemorrhage and perforation	yes	yes
K26.0	Duodenal ulcer ; Acute with haemorrhage	yes	yes
K26.2	Duodenal ulcer ; Acute with both haemorrhage and perforation	yes	yes
K26.4	Duodenal ulcer ; Chronic or unspecified with haemorrhage	yes	yes
K26.6	Duodenal ulcer ; Chronic or unspecified with both haemorrhage and perforation	yes	yes
K27.0	Peptic ulcer, site unspecified ; Acute with haemorrhage	yes	yes
K27.2	Peptic ulcer, site unspecified ; Acute with both haemorrhage and perforation	yes	yes
K27.4	Peptic ulcer, site unspecified ; Chronic or unspecified with haemorrhage	yes	yes
K27.6	Peptic ulcer, site unspecified ; Chronic or unspecified with both haemorrhage and perforation	yes	yes
K28.0	Gastrojejunal ulcer ; Acute with haemorrhage	yes	yes
K28.2	Gastrojejunal ulcer ; Acute with both haemorrhage and perforation	yes	yes
K28.4	Gastrojejunal ulcer ; Chronic or unspecified with haemorrhage	yes	yes
K28.6	Gastrojejunal ulcer ; Chronic or unspecified with both haemorrhage and perforation	yes	yes
K29.0	Acute haemorrhagic gastritis	yes	yes
K92.0	Haematemesis	yes	yes
K92.1	Melaena	yes	yes
K92.2	Gastrointestinal haemorrhage, unspecified		yes

In order to minimize information bias:

- Outcomes with less than 182 days of observation prior to their onset will be excluded;
- Outcomes with a total of 15 cumulative days of hospitalization or more in the 30 days prior their onset will be excluded.

▪ **ALCAPONE- AKI**

Inclusion codes for acute kidney injury

ICD10	Definition	Narrow definition	Broad definition
N17.0	Acute kidney failure with tubular necrosis	yes	yes
N17.1	Acute kidney failure with acute cortical necrosis	yes	yes
N17.2	Acute kidney failure with medullary necrosis	yes	yes
N17.8	Other acute kidney failure		yes
N17.9	Acute kidney failure, unspecified		yes
N19	Unspecified kidney failure		yes

In order to minimize information bias:

- Outcomes with less than 182 days of observation prior to their onset will be excluded;
- Outcomes with a total of 15 cumulative days of hospitalization or more in the 30 days prior their onset will be excluded.
- Outcomes presenting with an associated diagnosis, related diagnosis or primary diagnosis corresponding to the following codes in the 182 days prior to the index date will be excluded.

Exclusion codes for acute kidney injury

I12	Hypertensive renal disease
N99.0	Postprocedural renal failure
P96.0	Congenital renal failure
N18	Chronic kidney disease
T79.5	Traumatic anuria
E14.2	Unspecified diabetes mellitus with renal complication
O90.4	Postpartum acute renal failure
N14.3	Nephropathy induced by heavy metals
N14.4	Toxic nephropathy, not elsewhere classified
N16	Renal tubulo-interstitial disorders in diseases classified elsewhere
N29	Other disorders of kidney and ureter in diseases classified elsewhere
N15	Other renal tubulo-interstitial diseases
Q61	Cystic kidney disease
N04.9	Nephrotic syndrome
Z94.0	Kidney transplant status
Z49	Care involving dialysis
N08	Glomerular disorders in diseases classified elsewhere
N28.0	Ischaemia and infarction of kidney
N26	Unspecified contracted kidney
N12	Tubulo-interstitial nephritis, not specified as acute or chronic
N10	Acute tubulo-interstitial nephritis
N11	Chronic tubulo-interstitial nephritis
T39	Poisoning by nonopioid analgesics, antipyretics and antirheumatics
T40	Poisoning by narcotics and psychodysleptics [hallucinogens]
T41	Poisoning by anaesthetics and therapeutic gases
T42	Poisoning by antiepileptic, sedative-hypnotic and antiparkinsonism drugs
T43	Poisoning by psychotropic drugs, not elsewhere classified
T44	Poisoning by drugs primarily affecting the autonomic nervous system
T45	Poisoning by primarily systemic and haematological agents, not elsewhere classified
T46	Poisoning by agents primarily affecting the cardiovascular system
T47	Poisoning by agents primarily affecting the gastrointestinal system
T48	Poisoning by agents primarily acting on smooth and skeletal muscles and the respiratory system
T49	Poisoning by topical agents primarily affecting skin and mucous membrane and by ophthalmological, otorhinolaryngological and dental drugs
T50	Poisoning by diuretics and other and unspecified drugs, medicaments and biological substances

APPENDIX 2 IT FRAMEWORK AND SOFTWARE USED

Analysis were run using 2 servers: a Nec® Flex server with two Intel® Xeon® e5645 (6 cores, 12 threads each), 64GB RAM and 18TB storage, and a Dell® T640 with a two Intel® Xeon® 4110 (8 cores, 16 threads each) 64GB RAM and 37TB storage.

Software version

- R version 3.5.1
- SAS version 3.4

R package version

- CaseControl 1.5.0
- Cyclops 1.3.4
- DatabaseConnector 2.2.0
- dplyr 0.7.6
- EmpiricalCalibration 1.3.6
- FeatureExtraction 2.1.5
- gsubfn 0.7
- Java 0.9.10
- MethodEvaluation 0.2.1
- ohdsiRTools 1.5.5
- RSQLite 2.1.1
- pRoc 1.10.0
- proto 1.0.0
- survival 2.42-3
- sqldf 0.4-11
- xlsx 0.6.1

2.4.2. Methodological aspects

2.4.2.1. Selection of detectable controls

Drug controls were considered as detectable or not detectable on the basis of their minimal detectable relative risks (MDRR). MDRRs were calculated according to the following formula:¹⁶⁹

$$MDRR = \left[1 + \frac{z_{\alpha} - z_{1-\beta}}{2\sqrt{E}} \right]^2$$

Where:

- α denotes the size of the test (type I error, in this case α was set to 0.05);
- $1-\beta$ denotes the power of the test (in this case $1-\beta$ was set to 80%);
- z_{α} denotes the cut-off point for the upper $100(\alpha/2)$ percentile (two-sided test) of a standard normal distribution, for $\alpha = 0.05$, $z_{\alpha} = 1.96$;
- $z_{1-\beta}$ denotes the cutoff point for the upper $100(1-\beta)$ percentile of a standard normal distribution, for $1-\beta = 0.8$, $z_{1-\beta} = -0.842$ (note, $z_{1-\beta}$ is negative for power greater than 50%);
- E is the number of HOIs expected in the cohort of interest.

However, as in ALCAPONE we were dealing with case-based extractions including the totality of the population cases, we decided to approximate E by the observed number of patients exposed to the drug of interest in the relevant HOI sub-population.

More accurate methods exist to calculate sample size requirements for SCCS and CC designs.^{170, 171} However, those formulas are design-specific. Controls may have been considered as detectable for one design and not detectable for another, which would have considerably complicated the comparison step. Here, the idea was to use a generic method applicable across all the designs to generate a common reference set for each HOI, as it was done in the OMOP experiment.^{44, 127}

2.4.2.2. Self-controlled case series

Options tested in the SCCS included pre-exposure window adding, age adjustment, seasonality adjustment and multiple drug exposure adjustment.

As an outcome may temporarily reduce (e.g. contra-indication, NSAID prescription after UGIB) or increase (e.g. indication: sucralfate prescription after UGIB) the probability of exposure to the drug of interest, a deficit or an excess of outcomes in the period preceding the exposure may be observed.¹⁷² This prescribing behaviour may lead to a biased relative incidence in this period just prior to exposure and affect the overall unexposed period. One solution to this problem is to introduce a separate risk window, the “pre-exposure window”, just prior to exposure, to separate it from the remaining unexposed time.

Both, the rate of exposure and outcomes may change with age, and can even depend on the season. This may lead to confounding and bias the estimates. To correct for this, age and/or season can be included in the model. The effect of both age and season are assumed to be constant within each calendar month, and the rate from one month to the next can be different, even if subsequent months have somewhat similar rates. This is implemented using cubic spline functions. All people that have the outcome are used to estimate the effect of age and seasonality on the outcome.¹²⁶

Initially, we planned to also use the “Considering event-dependent observation time” option of the OHDSI SCCS package.¹²⁶ This option seemed very interesting since it was supposed to address an issue often observed in SCCS: the violation of the independence between outcomes and observation period. However, it appeared that, at the moment of the project, this option was not stable.

2.4.2.3. Case-control design

One of the major challenges when implementing CC design in ALCAPONE was the way controls were selected. As previously mentioned, 4 extractions were received from the SNDS, one for each HOI. At the beginning, we envisioned to select controls for one HOI amongst the patients of the 3 other extractions. However, this idea was rapidly dismissed in view of the potential selection bias that would have been induced. Subjects included through the 3 other HOIs could have had a potentially higher probability to be exposed to the drug controls than the general population, inducing an admission bias, also called Berkson bias.¹⁷³ The alternative solution was to exclude patients presenting the relevant HOI from the 800 000 subjects counting the EGB, the 1/97th SNDS sample, and to select controls in the remaining population. That did not present any difficulty for ALI and KI patients. However, with nearly

180 000 and 600 000 patients respectively, it was more challenging for UGIB and MI, especially considering a 1 to 10 matching approach. EGB option was chosen however, and it was decided that a given patient could be used as control for several cases and that unmatched cases were removed. The sampling of the largest patient extractions also helped to address this issue.

2.4.2.4. Case-population approach

As noted above, CP had never been formerly tested and was not present in the OHDSI Methods library. A full script was developed in-house relying on the following principles for both the per-user and the person-time approach.

- **Per-user approach**

Risk estimates generated from CP per-user approach are based on the following contingency table (Table 1)

Table 1. Contingency table for CP per-user approach

	Cases	Population (persons)
Exposed	<i>a</i>	<i>e</i>
Not exposed	<i>c</i>	<i>f</i>
Total	<i>n</i>	<i>N</i>

Where, over the study period:

- *a* denotes the number of cases exposed to the drug of interest;
- *c* denotes the number of cases not exposed to the drug of interest;
- *n* denotes the total number of cases whether they are exposed or not;
- *e* denotes the total number of exposed subjects in the source population (*i.e.* the total number of subjects with at least one dispensing of the drug of interest over the study period in the overall population);
- *f* denotes the total number of unexposed subjects;
- *N* denotes the total number of subjects of the source population.

From this, the following elements can be derived:

- Case-population Ratio

$$CPR = \frac{a}{c} * \frac{f}{e} \approx \frac{a}{n} * \frac{N}{e}$$

- Confidence interval of CPR

$$CI_{CPR} = CPR * \exp\left(\pm z_{1-\alpha/2} \sqrt{\frac{1}{a} + \frac{1}{n}}\right)$$

- Predicted Relative Risk

$$pRR = CPR * \frac{1 - e/N}{1 - a/n}$$

- Confidence interval of pRR

$$CI_{pRR} = pRR * \exp\left(\pm z_{1-\alpha/2} \sqrt{\frac{1}{a} + \frac{1}{c}}\right)$$

With $z_{1-\alpha/2} = 1.96$

- **Person-time approach**

In the person-time approach, rate from the control group (*i.e.* the overall population) are measured in observation time per-person as presented in Table 2.

Table 2. Contingency table for CP person-time approach

	Cases	Population (person-time)
Exposed	<i>a</i>	PT_E
Not exposed	<i>c</i>	PT_{NE}
Total	<i>n</i>	PT_{POP}

Where, over the study period:

- *a* denotes the number of cases exposed to the drug of interest;
- *c* denotes the number of cases not exposed to the drug of interest;
- *n* denotes the total number of cases whether they are exposed or not;
- PT_E denotes the exposed person-times in the source population (*i.e.* the cumulative duration corresponding to the dispensings of the drug of interest over the study period);
- PT_{NE} denotes the unexposed person-times in the source population (*i.e.* $PT_{POP} - PT_E$);

- PT_{POP} denotes the overall person-times in the source population (*i.e.* the duration of the study period multiplied by the catchment population).

Only CPR and pRR formula are impacted by these changes:

- Case-population Ratio

$$CPR = \frac{a}{c} * \frac{PT_{NE}}{PT_E} \approx \frac{a}{n} * \frac{PT_{pop}}{PT_E}$$

- Predicted Relative Risk

$$pRR = CPR * \frac{1 - PT_E / PT_{pop}}{1 - a/n}$$

As mentioned in the article, the difference between CPR and pRR flows from the fact that in the former, the number of cases (a) and the exposure rate (e or PT_E) are considered as so small that the overall number of cases (n) and the overall population (N or PT_{POP}) can respectively approximate the number of unexposed cases (c) and the unexposed population (f or PT_{NE}). A theoretical demonstration showed that the rarer the exposure and the event rate, the better the CPR approximates the actual pRR.¹¹⁷ CPR might be useful when only relative exposures of cases and population are known without any information about the exact number of exposed and unexposed patients. Both estimates are used in the literature.^{117, 118, 121}

The influence of the different extrapolation options of the aggregated data from the EGB was also assessed. In the raw extrapolation, for each year of the study, a single coefficient of extrapolation was applied to the whole population, irregardless of the age class or the gender. In the extrapolation stratified on age and gender, one coefficient was used for each age and gender class. Those coefficients are shared every year by the *Institut National de la Statistique et des Études Économiques* (INSEE), the French National Institute of Statistics and Economic studies.

2.4.2.5. Assessment and comparison of design performances

- ***Discriminating ability***

The discriminating ability of the methods was assessed through the area under the receiver operating characteristic (ROC) curve (AUC). AUC tells how much the variant

is capable to distinguish between positive and negative controls. In our works ROC curve was plotted and AUC was computed using the R package pROC.¹⁷⁴ Considering Table 3, pROC plotted the ROC curve with the true positive rate [TP/(TP+FN)] against the false positive rate [FP/(TN+FP)] (*i.e.* sensitivity vs. 1-specificity), using different thresholds to classify controls according to the log of the estimate generated by a design variant. An excellent model has an AUC near 1, which means it is capable of correctly classifying positive and negative controls. When AUC is 0.5, it means that classification is random and that no discriminative capacity exists.

Table 3. Contingency table computed for each design variant of each health outcome of interest

	Positive detectable controls	Negative detectable controls
Drug-outcome pairs detected as positive	TP (True Positives)	FP (False Positives)
Drug-outcome pairs detected as negative	FN (False Negatives)	TN (True Negatives)

- **Accuracy of the estimates**

We were also interested in the accuracy of the point estimates we generated. Since the true value of the positive control estimates are not known, we were only able to presume of the negative control ones, which should be 1.

Mean square error was calculated as described above and represents the average squared difference between the log estimate (RR) and zero. The smaller the MSE is, the better the estimation is.

$$MSE = mean[[\log(RR_{est}) - \log(RR_{true})]^2]$$

In some OMOP papers, accuracy was also assessed using “the bias” that is the average difference between the log estimate and 0.^{42, 48} However, it appeared that a variant generating negative biases ($RR_{est} < 0$) and positive biases ($RR_{est} > 0$) in the same proportion get a bias near 0, the negative error cancelling the positive one, inducing bias in the bias calculation.

Coverage probability was also computed as metrics to assess the accuracy. It can be seen as the probability that the interval generated by a variant contain the parameter of interest. In ALCAPONE the coverage for a design variant corresponds to the

number of times the confidence interval generated for the negative control estimates included the true value, 1.

OHDSI proposed a methodology to measure the accuracy of estimates generated for positive controls (*i.e.* with true estimate >1). To do so, starting from negative control, synthetic outcomes are injected in target exposure period until a desired ratio.^{175, 176} However, such an approach could not be implemented in ALCAPONE as only case-based designs were considered.

2.4.2.6. Why did we sample from the population?

According to the ALCAPONE protocol, 96 SCCS, 20 CC and 80 CP were planned across 4 HOIs, each HOI presenting a specific and a sensitive definition. Considering the number of drug controls to test, the total number of analyses to run was 61 338. Generating such a number of estimates in a reasonable timeframe across the SNDS extractions (more than 825 000 patients) would have required large computing capacity. At that time no such resources were available in-house. It would have been necessary to fall back on cloud computing services such as Amazon Web Services.¹⁷⁷ However, by law, SNDS data extraction are attached to a data processor which is supposed to be fully controlling its IT environment, the former having been previously declared during the regulatory phase. We then made the choice to do a first assessment of the design performances in sampled populations before going through a complete calibration of the best design variant in the unsampled ones.

2.4.2.7. Parameters with major impact

This stage has been added to the methodology when first results were generated. Although it was clear that some approaches were performing better than others, metrics of some design variants were so close that it was not easy to distinguish them. Hence, for the best-performing approach of each HOI, we computed logistic analyses to screen for the parameters with major impact on the discriminative ability. The dependent variable was the probability that a variant had an AUC above the 70th percentile of the AUC distribution. The independent covariates included the parameters that were varied in the different design variants. For example, independent covariates used in the logistic analyses applied to SCCS would be multiple drug adjustment: yes / no; pre-exposure duration: 0, 7 or 30 days, seasonality adjustment: yes / no, *etc.* When it was possible, multivariate analyses

were then run with the parameters that appeared to be significant in the univariate one.

2.4.2.8. Calibration of the reference design

Calibration, through the use of negative controls, is particularly important in ALCAPONE since it allows to see if the SNDS is suitable for alert generation (*i.e.* if controls are actually detected) and to what extent generated estimates can be trusted (*i.e.* if controls are correctly detected). At a glance, distribution of the negative control estimates gives a first idea of the global accuracy of the methods. To go into more depth, we used OHDSI EmpiricalCalibration R package to characterize the systematic error generated by the application of the design variant in the SNDS and to calibrated the *p-value* accordingly.^{83, 178} In traditional significance testing, *p-value* indicates the probability that a study finding greater than or equal to the one observed could have arisen under the null hypothesis (*i.e.* the hypothesis of no effect). In other words, it is the probability that the difference observed results from random error. The theoretical null distribution is usually used to compute it. In ALCAPONE, the negative controls give an indication of what is really happening under the null hypothesis, uncovering a potential systematic bias. Considering their estimates for which a theoretical value of 1 was expected and based on how often *p-value* was below 0.05 while the null hypothesis was true, a Gaussian probability distribution was fitted to effect estimates, modelling the distribution of the residual bias under the null hypothesis. Parameters of this “empirical null distribution” (mean and the standard deviation) helped to appreciate the bias induced by the design in the SNDS and were then used to compute calibrated *p-values*. Under the strong assumption that residual bias in the effect estimates is drawn from the same distribution as the residual bias in the set of negative controls, the calibrated *p-value* is supposed to take into account both random and systematic errors inherent to the application of a design variant. However, since there is no guarantee that this assumption is met – negative and positive controls may not be affected by the same biases – calibrated *p-value* must be interpreted with caution. Simulation studies showed that while this calibration process can help to control type I error rate (wrongly concluding to an effect by rejecting the null), type II error rate often increases (wrongly concluding to an absence of effect by not rejecting the null): a few

highly biased negative controls may induce overestimated p -values, masking potentially true associations.¹⁷⁹

The EmpiricalCalibration package also enabled the calibration of confidence intervals.⁸¹ However, to be conducted efficiently, such an approach requires to generate synthetic positive controls in a cohort, injecting artificial outcomes during periods covered by negative controls. It was deemed not suitable for ALCAPONE case-based extractions.

2.5. Feasibility study

The previous section introduced the overall ALCAPONE methodology implemented in the SNDS extractions. However, prior to this step, a feasibility study was conducted in the EGB involving phases *1a) Data extraction and management*, and *1b) Selection of the detectable control* of the full process.

2.5.1. Assessment of the mapping of the SNDS to the OMOP CDM

A major part of the statistical analyses planned in the ALCAPONE project was based on R packages made available by OHDSI, requiring data formatted according to OMOP CDM specifications to run. Starting from the EGB extractions, an assessment of the works necessary to conform with OMOP CDM v5 was conducted.¹⁸⁰ The findings of this assessment were presented during the 1st OHDSI European symposium in Rotterdam, Netherlands, in March 2018.

2.5.1.1. Abstract presented at the 2018 European OHDSI Symposium, March 23th, Rotterdam, Netherlands

Name:	Nicolas Thurin
Affiliation:	Bordeaux PharmacoEpi, INSERM CIC1401, Université de Bordeaux, Bordeaux, France
Email:	nicolas.thurin@u-bordeaux.fr
Presentation type (select one):	Poster

Mapping the 66.6M-patient French Nationwide Healthcare Database (SNDS) to the OMOP CDM v5: a feasibility study

Nicolas Thurin^{1,2}, Marine Pénichon¹, Régis Lassalle¹, Patrick Blin¹, Cécile Droz-Perroteau¹, Martijn Schuemie³, Nicholas Moore^{1,2}

¹Bordeaux PharmacoEpi, INSERM CIC1401, Université de Bordeaux, Bordeaux, France; ²INSERM U1219, Bordeaux, France; ³Janssen R&D, Titusville, New Jersey, USA

Abstract

France has a nationwide healthcare insurance system database, the SNDS (*Système National des Données de Santé*), covering about 99% of the French population. A 1/97th sample – the EGB (*Echantillon général des bénéficiaires*) – is also available. Work in these claims databases usually calls for a long and complex data management step. Transforming the SNDS database to the OMOP CDM would result in saving time in future studies and facilitating the collaboration with other international research teams. We describe our challenges to apply the ETL process to the clinical part of the EGB database and the technical issues to be overcome to achieve a complete mapping.

Introduction

The *Système National des Données de Santé* (SNDS) is the French Nationwide Healthcare System Database covering about 99% of the French population (about 66.6 million persons) from birth (or immigration) to death (or emigration). The SNDS database merges anonymous information of reimbursed claims from all French health care insurance plans, linked to the national hospital-discharge summaries database system (PMSI) and the national death registry. The database includes demographic data; health care encounters such as physician or paramedical visits, medicines, medical devices, and lab tests (without results); chronic medical conditions (ICD10 codes); hospitalisations with ICD10 codes for primary, linked and associated diagnoses, date and duration, procedures, diagnostic-related groups, and cost coding; date but currently not cause of death. EGB (*Echantillon Généraliste de Bénéficiaires*) is the 1/97th random permanent representative sample of SNDS database (780 000 subjects), with planned 20-year longitudinal data¹. EGB is powerful enough to study common issues with widespread drugs. Furthermore its structure is similar to SNDS, which makes it a perfect platform to develop feasibility studies before their implementation in the SNDS. This makes EGB an ideal candidate for assessing the possibility to apply the OMOP CDM v5 to the SNDS.

Methods

An extraction of all the myocardial infarctions between 2009 and 2014 were conducted in the EGB. The 74 resulting datasets were data managed according to the Bordeaux PharmacoEpi standardized operating procedure to generate 14 transitional datasets including socio-demographic data, medical visits, chronic medical conditions, hospitalization diagnoses, drug dispensing, lab tests, and outpatient and inpatient procedures. A theoretical mapping plan from the EGB extraction to the OMOP CDM v5.0.1 was realized through WhiteRabbit and RabbitInAHat softwares (Figure 1)². Mapping from source to standardized vocabularies was assessed (Table 1)³. Treatment durations for the drug_exposure table were generated according to drug dispensing data and medians of treatment durations, taking into account 3 months drug packaging.

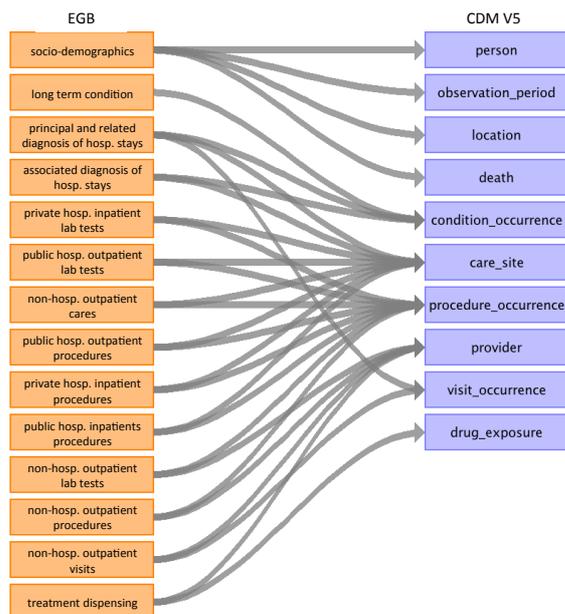


Figure 1. EGB Data Mapping Approach to CDM V5 (clinical part only)

Domains	EGB vocabularies	Presence is the OMOP vocabularies
Condition	ICD10	Yes
Drug	ATC	Yes
	CIP	No
Procedure	CCAM	No
	NABM	No

ICD: International Classification of diseases; **ATC:** Anatomical Therapeutic Chemical; **CIP:** Code identifiants de présentations; **CCAM:** Codage des actes médicaux; **NABM:** Nomenclature des actes de biologie médicale

Table 1. Presence of EGB vocabularies in the OMOP vocabularies

Results

To date, person, observation_period, location and drug_exposure tables have been generated. Death, condition_occurrence, care_site, provider and visit occurrence tables have not yet been created but do not seem to present technical barriers. However the utilization of national nomenclatures (CCAM and NABM) to code procedures and lab tests in the EGB prevents the utilization of the OMOP Standardized Vocabularies to generate the procedure_occurrence table.

Conclusion

Most of the standardized clinical data tables can be generated following the OHDSI Extract Transform and Load (ETL) processes. However, the mapping of the French terminologies for procedures and lab tests to standard vocabularies is required to complete fully the transformation of the clinical part of EGB and SNDS to the OMOP CDM, thus enabling the use of the whole set of OHDSI tools.

References

1. Bezin J, Duong M, Lassalle R, et al. The national healthcare system claims databases in France, SNIIRAM and EGB: Powerful tools for pharmacoepidemiology. *Pharmacoepidemiology and drug safety*. 2017; 26: 954-62.
2. OHDSI. WhiteRabbit.
3. Reich C, Ryan P, Belenkaya R, Natarajan K and Blacketer C. OMOP Common Data Model Specifications. OHDSI, 2017.

The poster displayed is available on OHDSI Europe website:

<https://www.ohdsi-europe.org/images/symposium-2018/posters/22-Nicolas-Thurin.pdf>

2.5.1.2. Discussion and consequences

As a significative part of the terminologies used in the SNDS were not available in the OMOP vocabulary, mapping its overall clinical part to the CDM would have required extra time and resources. As a result, in the frame of the ALCAPONE project, it was decided to only generate data tables required to perform the planned analyses: the tables *person*, *observation_period*, *location*, *drug_exposure*, *cohort*, *cohort_definition*, and *drug_era*. A reflection was thus launched on how to use SNDS dispensing data to construct exposure duration as required by the *drug_exposure* table of the CDM. It was decided to use median durations of treatment for non-chronic treatments. In France, for chronic treatment, if no 90-pill box are available, pharmacists have to dispense enough quantity of drug to cover 28 days, whatever the original size of the box.¹⁸¹ The patient has then to come monthly to the pharmacy to refill. Based on this observation, in ALCAPONE, once the 90-pill boxes identified, a duration of 1 month was set to all the dispensings of chronic drugs. Median durations of treatment were then used for dispensings of non-chronic drugs. The *drug_exposure* table was then used to generate the *drug_era* table, which is the one queried by the R packages to run the analyses. Area of treatment were built using a 30-day grace period. No extra duration resulting from potential pill storage was applied.

2.5.2. Adaption of the reference set and power consideration

2.5.2.1. Results

To be sure of their presence and their reimbursement on the French market, drugs contained in the original OMOP reference set have been mapped to Medic'AM by their international non-proprietary names.⁴⁷ Medic'AM contains aggregated data corresponding to the monthly dispensings of all the drugs reimbursed by the French national health insurance.¹⁸² Drugs that do not appear in this database were not retained. Drugs with more than 2 years without any reimbursed box between 2009 and 2014 were dismissed. Methotrexate, tetracosactide and neostigmine were ruled out because of their different usage pattern and the difficulty in estimating the treatment duration. Non-topical and non-ophthalmic ATC of the remaining drugs were extracted to constitute a reference set adapted to the French market. As presented in the Table 4, between 70% and 80% of the positive controls from the original OMOP

set have been conserved, however more than 40% of the negative controls have been dismissed.

Table 4. Percentage of loss of positive (CTR+) and negative (CTR-) controls resulting from the restriction of the OMOP reference to the drugs available on the French market.

		OMOP Reference set	Restricted Reference set*	Loss
ALI	CTR+	81	56	31%
	CTR-	37	19	49%
MI	CTR+	36	26	28%
	CTR-	66	37	44%
KI	CTR+	24	19	21%
	CTR-	64	32	50%
UGIB	CTR+	24	19	21%
	CTR-	66	38	42%

* taking into account the different administration routes for a same drug

We then checked if it was possible to detect an association ≤ 1.25 for these controls in the EGB. To do so, we computed MDRRs according to EGB settings for each one of the remaining drugs of the set. MDRR were then extrapolated to the SNDS. Results showed that the EGB was not powerful enough to detect associations ≤ 1.25 among a sufficient number of drug-outcome pairs, in particular when the outcome and/or the exposition was rare (Table 5). Nevertheless, the results extrapolated to the SNDS seemed to be conclusive. The low number of controls considered as detectable in the ALI and KI population could be explained by the small size of the extraction and the random error: only two exposed cases in the EGB are required to be considered as detectable in the SNDS.

Table 5. Calculated number of detectable controls in the EGB and estimated number of detectable controls in the SNDS

		Restricted reference set	Calculated number of detectable controls in the EGB		Estimated number of detectable controls in the SNDS	
			<i>CTR+ et CTR- with MDRR\leq1,25</i>		<i>CTR+ et CTR- with MDRR\leq1,25</i>	
			Narrow definition	Broad definition	Narrow definition	Broad definition
ALI	CTR+	56	0	0	15	18
	CTR-	19	0	0	1	2
MI	CTR+	26	3	5	23	23
	CTR-	37	1	5	29	31
KI	CTR+	19	0	3	11	18
	CTR-	32	0	0	5	16
UGIB	CTR+	19	5	7	18	19
	CTR-	38	1	1	30	31

These results were presented to the 33rd ICPE in Montréal, QC, Canada in August 2017.¹⁸³ The corresponding poster is available in Appendix 6.

2.5.3. Conclusion of the feasibility study

The feasibility study concluded that ALCAPONE overall process could not be conducted in the EGB, but that the SNDS seemed to have a sufficient size to implement such a process. It also highlighted that the number of detectable controls had to be confirmed in SNDS settings and that the addition of controls from EU-ADR reference set or the adjustment of the MDRR threshold should be considered if power issue remained.

As presented in the article about methodology, for the execution of ALCAPONE in the SNDS, EU-ADR drug controls were actually added to the reference set adapted to the French market. MDRR threshold was also re-evaluated and set to 1.30 instead of 1.25.

2.6. Preparation of SNDS data

2.6.1. Data extraction

To date 143 data tables with 3 714 variables are stored in the SNDS.¹⁸⁴ Every year, tables and variables can be added, modified or deleted. To be able to link between

the different universes for a same patient (dispensings, death, hospital data) 9 different variables have to be merged to form a join key. When data are requested to the CNAM for research purposes, extractions are performed at once but by year: according to the study settings, up to 134 tables and 2630 variables can be provided for each year of the study period.¹⁸⁵ ALCAPONE covers the 2009-2014 period for a total of 6 years. The corresponding extraction provided by the CNAM contained 612 data tables for a total 1.256To. Using the latest version of SAS®, data were checked for integrity and consistency according to the standardized operating procedures of the Bordeaux PharmacoEpi platform. They were then cleaned (e.g. deleting subject with age>120 or ghost care centres, etc.) and the yearly-varying version of a same table concatenated. The previously presented feasibility study helped us to identify the SNDS tables of interest for the project: the demographics, the outcome of interest and the drugs dispensed in primary care. From the SNDS tables (called “source” tables), the following were considered for mapping

- *IR_BEN_R*: the referential of the person covered by the national health insurance.
- *ER_PHA_F*: contains all the drugs dispensed by community pharmacies and reimbursed by the national health insurance. Drugs dispensed in hospital settings are not included
- *T_MCO_B*: contains hospital stay general information (dates, duration, hospital, location, etc.)
- *T_MCO_C*: table enabling to link hospital data to primary care data
- *T_MCO_D*: contains ICD10 diagnostic codes for associated health condition
- *T_MCO_UM*: contains primary and secondary diagnostics codes corresponding to the cause of the hospitalization.

2.6.2. Data mapping

A complete extract, transform, and load process (ETL) was designed to map the source tables to the OMOP CDM. Figure 1 displays an overview of this process. The full ETL, including the selected variables as well as the drug duration construction is available in Appendix 7.

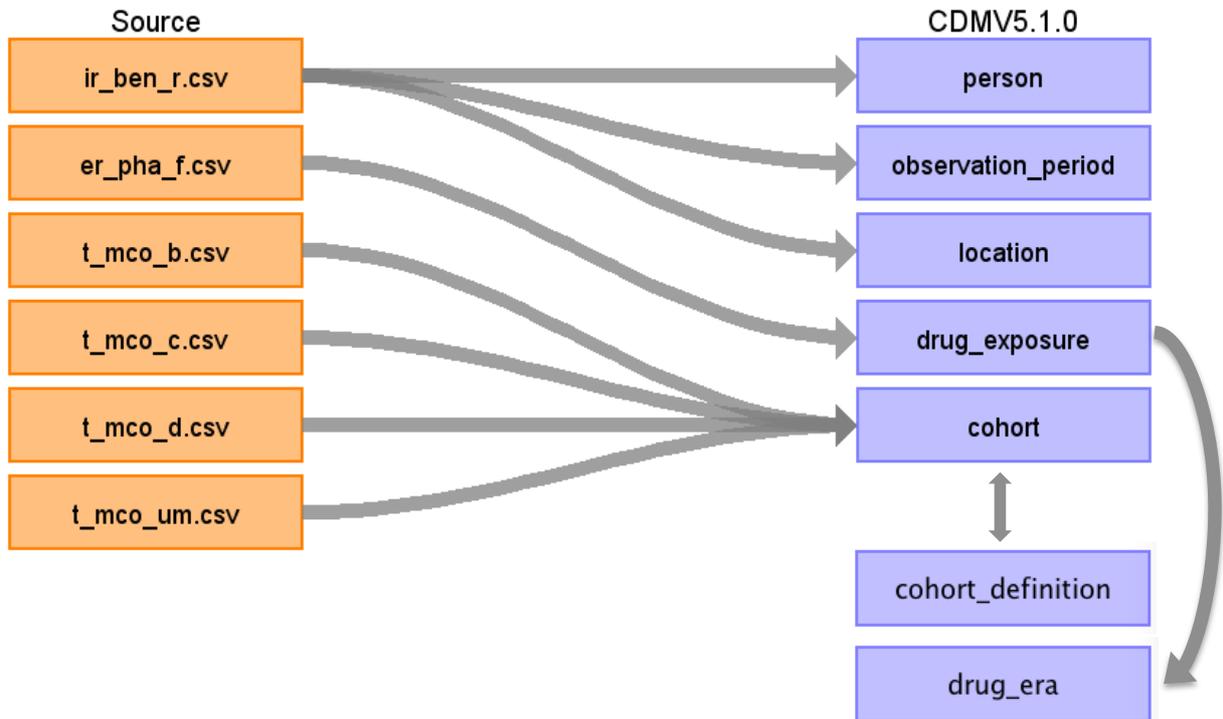


Figure 1. Overview of the Extract Transform and Load process from the SNDS to the OMOP Common Data Model (CDM) v5

OMOP CDM specifications define “standard vocabularies” that must be used in the CDM (e.g. SNOMED-CT for diagnostic codes, RxNorm for drug codes, etc.). This is particularly useful when different databases are involved in a same project. However, switching from the source vocabulary to the standard one call for time and resources, and sometimes impact information quality. In the case of ALCAPONE, we tried to avoid such process since no other partners were involved. Each definition (sensitive or specific) of each HOI (ALI, AKI, MI, UGIB) was matched to a `cohort_definition_id` running from 1001 to 1008 in the `cohort_definition` table. This `cohort_definition_id` was then used in the `cohort` table to flag corresponding patients. This way to process, allowed to skip the stage of implementation of the overall diagnostic codes of the patients in the database. Regarding drugs, ATC codes were conserved even if translated into `concept_ids`: in the CDM codes are under a numeric format, cross-files

are available to translate code from a defined classification to an OMOP *concept_id*. For example, the ATC code for ibuprofen is “M01AE01”, this code corresponds in the OMOP CDM to the *concept_id* “21603967”. Transformed data were then loaded into a PostgreSQL database schema to be queried using R software.

2.7. Results

2.7.1. General results

The first observation that can be drafted from ALCAPONE is that SCCS seems to performed better than CC and CP for the identification of drugs associated with ALI, MI, AKI and UGIB, in both the specific and the sensitive definition. Figure 2 and Figure 3 presented respectively for the specific and the sensitive definition of each HOI, the variant with the highest AUC and the lowest MSE of each approach (SCCS, CC and CP).

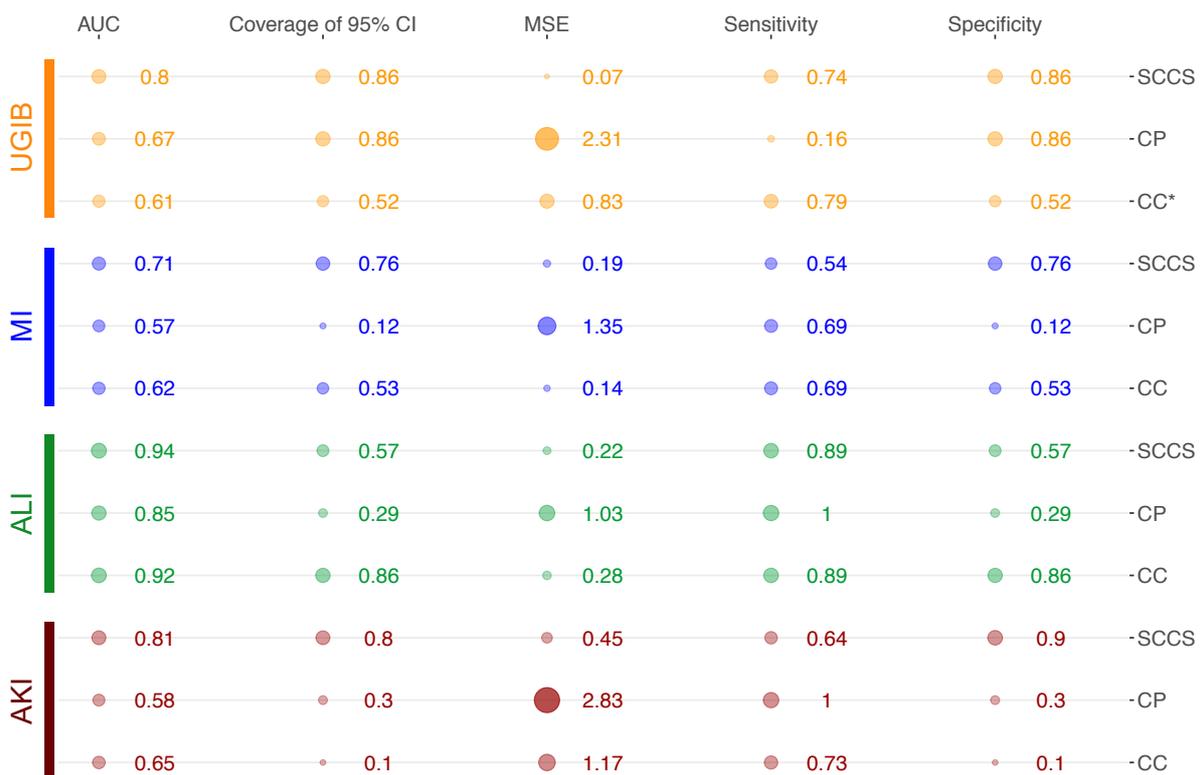


Figure 2. Overall performances of self-controlled case series (SCCS), case-population (CP) and case-control (CC) designs in the detection of drugs associated with upper-gastrointestinal bleeding (UGIB), myocardial infarction (MI), acute liver injury (ALI) and acute kidney injury (AKI), specific definitions

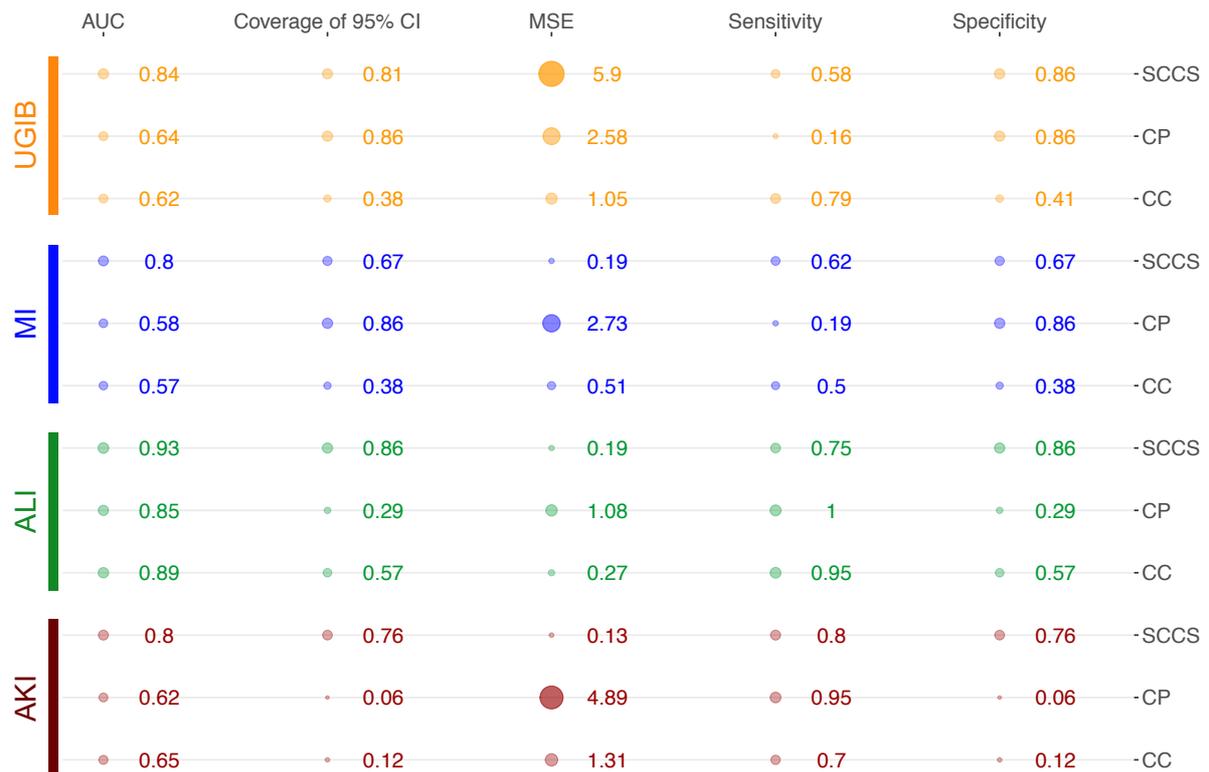


Figure 3. Overall performances of self-controlled case series (SCCS), case-population (CP) and case-control (CC) designs in the detection of drugs associated with upper-gastrointestinal bleeding (UGIB), myocardial infarction (MI), acute liver injury (ALI) and acute kidney injury (AKI), sensitive definitions

It clearly appears that in both definitions, most of the time, the best-performing SCCS variant had the highest discriminant ability with AUC values running from 0.7 to 0.94, and the best predictive accuracy with MSE up to 0.07. This difference was less marked for ALI than for the other HOIs. In most of the cases CC and CP AUC were relatively close but CP always showed significantly higher MSE.

Global results of ALCAPONE were presented during the 2nd European OHDSI Symposium in Rotterdam, Netherlands and during the 35th ICPE in Philadelphia, PA, USA in August 2019.^{186, 187} Links to oral presentation are provided in Appendix 1.

2.7.2. UGIB example, submitted article: Empirical assessment of case-based methods for identification of drugs associated with upper-gastrointestinal bleeding in the French National Healthcare System database (SNDS)

Nicolas H. Thurin, Régis Lassalle, Martijn Schuemie, Marine Pénichon, Joshua J. Gagne, Jeremy A. Rassen, Jacques Benichou, Alain Weill, Patrick Blin, Nicholas Moore, Cécile Droz-Perroteau, Pharmacoepidemiology and Drug safety (1st submission)

Empirical assessment of case-based methods for identification of drugs associated with upper gastrointestinal bleeding in the French National Healthcare System database (SNDS)

Short title: Case-based methods for UGIB-associated drug identification in the SNDS

Nicolas H Thurin^{1,2,*}, Régis Lassalle¹, Martijn Schuemie^{3,4}, Marine Pénichon¹, Joshua J Gagne⁵, Jeremy A Rassen⁶, Jacques Benichou^{7,8}, Alain Weill⁹, Patrick Blin¹, Nicholas Moore^{1,2}, Cécile Droz-Perroteau¹

¹Bordeaux PharmacoEpi, INSERM CIC1401, Université de Bordeaux, Bordeaux, France.

²INSERM U1219, Université de Bordeaux, Bordeaux, France.

³Epidemiology Analytics, Janssen Research and Development, Titusville, NJ, USA.

⁴Observational Health Data Sciences and Informatics (OHDSI), New York, NY, USA.

⁵Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA USA.

⁶Aetion, Inc., New York, NY, USA.

⁷Department of Biostatistics and Clinical Research, Rouen University Hospital Rouen, France.

⁸INSERM U1181

⁹Caisse Nationale de l'Assurance Maladie, Paris, France.

* Corresponding authors

nicolas.thurin@u-bordeaux.fr
Bordeaux PharmacoEpi
Université de Bordeaux
case 41 - 146 rue Léo Saignat
33076 Bordeaux cedex - France

Key words:

Upper gastrointestinal bleeding, self-controlled case series, case-control, case-population, calibration, claims database

Key points:

- Self-controlled case series approaches show the best performances for the identification of drug associated with upper gastrointestinal bleeding in the French National Healthcare System database (SNDS) with very small systematic error
- Adjusting for multiple drugs and using a risk window corresponding to the 30 first days of exposure seem crucial to consider when assessing upper gastrointestinal bleeding risk
- Using a specific definition of the disease can help reduce bias, especially in reduced samples
- Negative controls are useful to check for performances of the method
- Clinical expertise is necessary to ensure a correct interpretation of the results

Word count excluding abstract, tables, figures and references: 2994

Statement about prior postings and presentations, name(s) of any sponsor(s) of the research contained in the paper, along with grant number(s):

ALCAPONE (Alert generation using the case population approach) methodology and related results were presented during the 33rd and 35th International Conference on Pharmacoepidemiology & Therapeutic Risk Management and at the 1st and 2nd OHDSI European Symposium.

Pharmacoepidemiology and Drug safety is currently reviewing an article about ALCAPONE methodology (PDS-19-0255).

The ALCAPONE project is funded by the French Ministry of Health (PREPS 2014, 0635).

ABSTRACT

245/250 words

Purposes: Upper gastrointestinal bleeding (UGIB) is a severe event, which is frequently drug-related. In order to enable efficient drug safety alert generation in the French National Healthcare System database (SNDS), we assessed and calibrated empirically case-based designs to identify drug associated with UGIB risk.

Methods: All cases of UGIB were extracted from SNDS (2009-2014) using two definitions. Positive and negative drug controls were used to compare 196 self-controlled case series (SCCS), case-control (CC), and case-population (CP) design variants. Each variant was evaluated in a 1/10th population sample using area under the receiver operating curve (AUC) and mean square error (MSE). Parameters that had major impacts on results were identified through logistic regression. Optimum designs were replicated in the unsampled population.

Results: Using a specific UGIB definition, AUCs ranged from 0.64 to 0.80, 0.44 to 0.61 and 0.50 to 0.67, for SCCS, CC and CP, respectively. MSE ranged from 0.07 to 0.39, 0.83 to 1.33 and 1.96 to 4.6, respectively. Univariate regressions showed that high AUCs were achieved with SCCS with multiple drug adjustment and a 30-day risk window starting at exposure. The top-performing SCCS variant in the unsampled population yielded an AUC=0.84 and MSE=0.14, with 10/36 negative controls presenting significant estimates.

Conclusions: SCCS adjusting for multiple drugs and using a 30-day risk window showed good performances for the identification of UGIB in the SNDS. Negative control implementation highlighted that low systematic error was generated but that protopathic bias and confounding by indication remained unaddressed issues.

1 INTRODUCTION

Upper gastrointestinal bleeding (UGIB) is a serious medical emergency, related to bleeding from the esophagus, stomach, or duodenum.¹ The incidence of UGIB ranges from 36 to 172/100 000 inhabitants per year and has been declining in recent decades.²⁻⁴ This trend could be explained by the decreasing prevalence of *Helicobacter pylori* and the large increase in the use of proton pump inhibitors (PPI), especially with high-dose, long-term non-steroidal anti-inflammatory drugs (NSAIDs).^{1, 2}

UGIB leads to death in about 10% of cases.^{5, 6} The high incidence of UGIB and the role of drugs as potential causes, especially NSAIDs, has made UGIB an important focus of pharmacoepidemiology.⁷⁻¹⁰ UGIB was included in the 10 events studied in the Exploring and Understanding Adverse Drug Reactions by integrative mining of clinical records and biomedical knowledge (EU-ADR) project as well as among the health outcomes of interest screened in the Observational Medical Outcomes Partnership (OMOP) experiment.¹¹⁻¹⁴ NSAIDs and UGIB were also assessed in the Safety Of non-Steroidal anti-inflammatory drugs (SOS) project, an international initiative partially funded by the European Commission following rofecoxib withdrawal, through a meta-analysis and a multi-database nested case-control study.^{15, 16}

The French National Healthcare System database – the *Système National des données de Santé* (SNDS) – currently includes 66.6 million persons, capturing all reimbursed medical and paramedical encounters, including all dispensed drugs, hospital-discharge summaries and dates of death,¹⁷ was not available at the time of SOS project. With the whole French population, SNDS is twice as large as the combined databases in SOS without the same degree of heterogeneity, which makes it a valuable tool to address drug safety questions. Other databases in Europe were part of the OMOP experiment,¹⁸ which tested and calibrated various methods to measure drug-outcome associations, using reference lists of known associations and non-associations. ALCAPONE (Alert generation using the case population approach in the French databases), a project funded by the French Health Ministry (PREPS 2014, 0635), aimed to further the objectives of OMOP by empirically assessing and calibrating case-based methods – self-controlled case series (SCCS), case-control

(CC) and case-population (CP) – to determine the best-performing design and corresponding settings to assess associations between drugs and UGIB in the SNDS.

2 METHODS

The overall ALCAPONE methodology has been fully described elsewhere (manuscript under review PDS-19-0255).

Patients with UGIB were identified in the SNDS (2009-2014) using ICD10 codes from hospital discharge summaries based on either a specific or a sensitive definition. The specific definition included codes for gastric ulcer, duodenal ulcer, peptic ulcer, gastro-jejunal ulcer, acute hemorrhagic gastritis, hematemesis and melaena. The more sensitive definition also considered patients hospitalized for unspecified gastrointestinal hemorrhage (Appendix 1). The index date was the date of hospital admission. Since inpatient drug use is not captured in the database, outcomes with more than 15 days of hospitalization in the month preceding the index date were excluded to ensure completeness of exposure assessment. Outcomes with less than 182 days of observation prior to the index date were also excluded.

Drugs of interest were restricted to those with enough power to be detected in the population (minimum detectable relative risk ≤ 1.30).¹⁹ Positive controls were drugs with a known association with UGIB. Negative controls were those with no known association with UGIB. The full ALCAPONE reference set for UGIB is available in Appendix 2. Different settings, such as adjustment strategies, risk window lengths, *etc.*, were applied to the three case-based approaches (SCCS, CC and CP), forming different design variants. These variants were applied in two steps to generate point estimates (relative incidence for SCCS, odds ratio for CC and case population ratio or predicted relative risk for CP) between UGIB and each drug control, as described below. A total of 96 SCCS, 20 CC and 80 CP variants were tested. The exact settings of each design variant are described in Appendix 3.

In a first step, all design variants were run in a 1/10th sample of the case population to identify the best-performing approach based on area under the receiver operating characteristics curve (AUC), mean square error (MSE) and coverage probability. Only drug controls that had sufficient power to detect a relative risk ≤ 1.30 in this sampled population were included. MSE and coverage probability were estimated for negative

controls only, assuming no association (*i.e.*, true point estimate of 1 for the measure of relative association). The AUC (ranging from 0 to 1) measures discrimination between positive and negative controls, with higher values indicating better discrimination. The MSE quantifies the accuracy of an estimator, combining both bias and random error, with smaller values indicating better accuracy. The coverage probability is the proportion of the 95% confidence interval estimator that includes the true parameter value, 1 in our case. Values as close as possible to nominal 95% coverage are desirable.

Once the best performing case-based approach was identified, we conducted an univariable logistic regression analysis to screen for parameters that best discriminated the performance of the different design variants. The dependent variable was the probability that a variant had an AUC>0.75 with the specific definition (AUC>0.78 with the sensitive definition). AUC thresholds were selected as the 70th percentile of the AUC distributions of the variants. The independent covariates included the parameters that were varied in SCCS analyses (multiple drug adjustment: yes/no; pre-exposure length: 0, 7 or 30 days, *etc.*).

In the second step, the best-performing variant was applied to the full, unsampled case population.

Considering the estimates from the negative controls, for which a theoretical value of 1 was expected, we observed how often $p < 0.05$ while the null hypothesis was true, and we fitted distribution to the effect estimates, modeling the distribution of the residual bias under the null.²⁰⁻²² Estimated parameters of this “empirical null distribution” was then used to compute “calibrated” p -values, taking into account random and systematic error (*i.e.* the background noise) inherent to the application of a design variant to the SNDS.

3 RESULTS

3.1 POPULATION

The selection process of UGIB cases according to specific and sensitive definition is presented in Figure 1. Over 6 years, 139 172 patients with 156 057 UGIB episodes were included according to the specific UGIB definition; 1 661 patients presented with

more than one outcome. The sensitive definition added 50 120 patients with 54 547 gastrointestinal hemorrhages for a total of 204 442 outcomes. With both definitions, median age was 72 and 57% were men. From the 64 drugs of interest screened in the unsampled population, 58 and 59 presented a minimum detectable relative risk ≤ 1.30 and were then deemed detectable according to the specific and sensitive definition, respectively. Moving to the 1/10th sampled population, both specific and sensitive definitions enabled the detection of 19 positive drugs controls out of 22 and 22 negative controls out of 42.

3.2 BEST PERFORMING DESIGN VARIANT AND MAJOR IMPACT PARAMETERS

SCCS globally showed better discrimination (Figure 2) and MSE (Figure 3) than CC and CP. For the specific case definition, AUCs ranged from 0.64 to 0.80 for SCCS, from 0.44 to 0.61 for CC and from 0.50 to 0.67 for CP. MSE ranged from 0.07 to 0.39 for SCCS, from 0.83 to 1.33 for CC and from 1.96 to 4.6 for CP. For the sensitive definition, the same trends were observed with some AUC reaching 0.84 for SCCS. The lowest MSE was also observed for SCCS (MSE<0.3), although the distribution was more heterogeneous with some very large MSE values (e.g. >2000). Performances of the all the design variants tested in the 1/10th sampled population are available in Appendix 4.

For the specific definition, the design variants with the highest AUC (0.80) and smallest MSE (0.07) were the SCCS 2066 and 2068, with a coverage probability of 86% (Table 1). SCCS variant 2066 (see Appendix 3) only considered the first occurrence of the outcome, used a risk window corresponding to the first 30 days following the dispensing of the drug of interest and adjusted for multiple drugs. Variant 2068 further adjusted for seasonality. These design variants also performed quite well when applied to the sensitive definition with AUC ≥ 0.82 , MSE=0.29 and a coverage probability of 82%. Variant 2090 yielded a higher AUC (0.84) but at the expense of the MSE (MSE=5.9) (Table 2).

Although SCCS clearly appeared as the best performing approach, identifying why this design variant was optimal remained challenging. Univariable logistic regression analyses showed that, for both the specific and the sensitive definitions, the strongest determinant of a high AUC was multiple drug adjustment (Table 3 and Table 4). In the specific UGIB definition, considering the first 30 days after drug dispensing as the risk

window and considering only the first occurrence of UGIB were also associated with a high AUC (Table 3). Since adjusting for seasonality appeared to have no impact on SCCS discriminative ability, we determined SCCS 2066 to be the optimum variant.

3.3 CALIBRATION OF THE BEST DESIGN VARIANT

3.3.1 Specific case definition

In the unsampled specific case population, SCCS 2066 showed a better AUC (0.84 vs. 0.80) with slightly increased MSE and reduced coverage (respectively 0.14 vs. 0.07 and 75% vs. 86%) as compared to what was observed in the 1/10th sampled population. Out of 36 negative controls, 26 were not significantly associated with UGIB (Figure 4). Nine negative controls (i.e., miconazole, sucralfate, lactulose, sitagliptin, erythropoietin, nitrofurantoin, loratadine, methocarbamol and zopiclone) had the lower bounds of their confidence intervals above 1 whereas scopolamine had an upper bound below 1 (protective effect). Almost all the positive controls were significantly associated with UGIB except clindamycin, sulindac, etodolac and mefenamic acid (Figure 5).

An empirical null distribution ($\hat{\mu} = 0.12$; $\hat{\sigma} = 0.17$) was derived based on the negative control estimates and used to compute calibrated p -values for SCCS 2066. Using conventional p -values, 10 out of the 36 negative controls and 18 out of the 22 positive controls were significant. Using the calibrated p -values, only two negative controls were still significant (sucralfate and scopolamine) (Table 5), and 9 positive controls moved from significant to non-significant (potassium chloride, prednisolone, indomethacin, ibuprofen, fenoprofen, nabumetone, fluoxetine, citalopram, sertraline) (Table 6). Figure 6 illustrates this calibration process. The gray zone (below the dashed line) represents the area where the conventional p -value is smaller than 0.05. The orange zone shows the area where calibrated p -value is under 0.05. We can see that the 8 negative controls significantly associated with UGIB according to the conventional p -value are below the dashed line but out of the orange area, meaning that once the p -value calibrated they are no longer considered as significant.

3.3.2 Sensitive case definition

As compared to the 1/10th sampled population, an improvement in AUC (0.85 vs. 0.82) and MSE (0.14 vs. 0.29) for SCCS 2066 was observed after execution in the

unsampled sensitive population. The coverage probability decreased slightly (78% vs. 82%). SCCS 2066 generated estimates comparable to those found for specific definition for both negatives and positives controls (Figure 7 and Figure 8). The empirical null distribution was derived ($\hat{\mu} = 0.12$; $\hat{\sigma} = 0.17$) and used to compute calibrated p -values (Table 6). The same trends as those observed for the specific definition appeared in the calibration process, except for sertraline which remained significantly associated after calibration, and escitalopram which did not.

All estimates generated in the unsampled population are available in Appendix 5.

4 DISCUSSION

Overall, the SCCS clearly showed better results than CC and CP in terms of discrimination and accuracy of point estimates in this large-scale assessment in the SNDS. Using both specific and sensitive definitions for UGIB, adjustment for multiple drugs seemed to be the strategy with the largest impact on accurately classifying positive and negative controls. Restricting the risk window to the 30 first days from dispensing appeared to increase performance of SCCS when using the specific definition. This may be related to the exclusion of non-specific bleeding that happens long after treatment initiation and that is unrelated to the drugs of interest, or to a depletion of susceptibles.²³ Although the corresponding coefficient was not statistically significant in the regression model, restricting outcomes to incident events seemed to have positive effect. Such an approach is often used in SCCS when recurrences of an event are not independent, which may be the case for UGIB.²⁴⁻²⁶ The AUC obtained in the unsampled population for these best-performing designs (0.84 and 0.85 for the specific and the sensitive definitions, respectively) are consistent with what was observed in the original OMOP experiment assessing SCCS and UGIB in US databases (0.77 to 0.88), even if settings differed.²⁷ In addition, adjustment for multiple drugs and the restriction to the first 30 days from initial exposure were also the settings of the best performing SCCS (AUC = 0.84) in the European replication of the OMOP experiment, for UGIB.¹⁸

When considering CP, analyses 3001 to 3040 using the actual number of persons exposed (per user approach), clearly yielded a better AUC (>0.6) compared to analyses 3041-3080 using cumulated person-time exposure (person-time approach)

(<0.6). As with the SCCS 30-day risk window, this observation suggests that UGIB is globally a patient-dependent effect (*i.e.* type B idiosyncratic reaction) that occurs early after initiation of exposure.²⁸ Optimal conditions for CP are low exposure and event rates.²⁹ Given the high level of events observed in the population, the absence of adjustment may be the principal explanation for the relatively lower performance of the CP approach.

Even if less biased than CP, CC results showed poor performances with AUC<0.6 for most settings. Since we have access to all cases of interest in the SNDS, but not all non-cases, cases were matched to controls selected from a 1/97th sample of the database. The limited pool of controls available combined may have affected the performance of the CC approach.^{30, 31} Moreover, cases and controls were matched on age and sex only. More advanced methods for adjustment, such as propensity score or disease risk score matching or stratification, may have reduced residual confounding and increased method performance.^{32, 33}

Regardless of the population or the SCCS variant, some negative control drugs always appeared as significantly associated with UGIB. Some of these associations can be easily explained. For example, the protective effect of scopolamine is not surprising, since it is a strong anticholinergic agent.³⁴ SNDS captures outcomes through hospital discharge diagnosis codes and drugs such as sucralfate, a standard treatment of evolving gastric and duodenal ulcer, could have been dispensed to the patients following a general practitioner visit for the initial symptoms of UGIB prior to the hospitalization (protopathic bias).³⁵ Such bias could also be observed with erythropoietin, which is indicated to manage anemia, potentially resulting from bleeding. Lactulose is indicated to prevent hepatic encephalopathy by lowering ammonia concentrations in the digestive tract after bleeding. Hepatic encephalopathy usually results from liver failure, including cirrhosis.³⁶ Patients with cirrhosis are well known to have higher risk of esophageal varices. Thus, the association between lactulose and UGIB may be due to confounding by indication.³⁷ Similarly, arthritic patients are often medicated with methocarbamol, a muscle relaxant, in addition to NSAIDs during flares. Since NSAIDs carry a major risk of UGIB, methocarbamol's association with UGIB could also result from confounding. Lastly, azole antifungals are well known inhibitors of cytochrome P450 enzymes (miconazole and CYP2C9, ketoconazole and CYP3A4). Vitamin K antagonists and direct oral anticoagulants are

metabolized respectively by CYP2C9 and CYP3A4, respectively. By inhibiting these enzymes, drug plasma level may increase, increasing risk of major bleeding, leading to an association due to confounding.

All of these examples show that even the optimal SCCS variant is not always able to address protopathic bias and confounding by indication without further consideration and additional forms of adjustment. Clinical and pharmacological inputs are essential to interpret final results.

Estimates from NSAIDs included in the reference set were consistent with the results from the SOS project for ibuprofen, ketoprofen, naproxen and meloxicam, and lower for piroxicam and indomethacin.¹⁵ We had insufficient power to properly assess sulindac, etodolac and mefenamic acid. However, these drugs are rarely used and are therefore of less public health importance. We also observed consistency in results for clopidogrel and selective serotonin reuptake inhibitors (SSRIs), which is reassuring since SSRIs and clopidogrel facilitate UGIB by the same mechanism – inhibition of platelet aggregation.

Given the huge statistical power afforded by the SNDS (>66 million people), one may question the relevance of the p -value since random error approaches 0 with increasing sample size and even small clinically irrelevant findings can become statistically significant. Under the strong assumption that residual bias in the effect estimate is drawn from the same distribution as the residual bias in the set of negative controls, calibrating the p -value can further account for systematic error. Systematic error seems to be small when certain SCCS variants are applied to investigate UGIB in the SNDS. However calibrated p -values have to be interpreted carefully since the distribution of residual bias described in effect estimates of negative controls may differ from those of the drugs of interest (i.e., the positive controls). Moreover, simulation studies showed that type II error rates often increase with this calibration process – few highly biased negative controls may lead to overestimated p -values, masking potential true associations, which can be problematic in the context of signal detection.²² Although the use of p -value calibration is debated, we think that the implementation of negative controls provides some reassurance about the performance of the SCCS method.

The SCCS had the best performance for the identification of drug-related UGIB in SNDS. Adjusting for multiple drugs and considering the initial period of treatment seemed to be important features of this design. However, not all the possibilities of

SCCS have been assessed here and specific design adjustment may be required in the context of a particular study. Outcome definitions must be carefully selected to ensure good accuracy of the method. The calibration process showed that low systematic error was generated by SCCS in the SNDS when applied to UGIB. However, the analysis of negative controls indicated that some biases such as protopathic bias and confounding by indication remained unaddressed and indicate a need for a clinical expert input to ensure a correct interpretation of the results.

ACKNOWLEDGMENTS

The ALCAPONE project was funded by the French Ministry of Health (PREPS 2014, 0635).

The authors wish to thank the SNDS database management team, and ADERA, the non-profit organization that provides managerial and human resource support to Bordeaux PharmacoEpi.

REFERENCES

1. Stanley AJ, Laine L. Management of acute upper gastrointestinal bleeding. *BMJ* 2019, 364:l536.
2. Hreinsson JP, Kalaitzakis E, Gudmundsson S, Bjornsson ES. Upper gastrointestinal bleeding: incidence, etiology and outcomes in a population-based setting. *Scand J Gastroenterol* 2013, 48:439-447.
3. Laine L, Yang H, Chang SC, Datto C. Trends for incidence of hospitalization and death due to GI complications in the United States from 2001 to 2009. *Am J Gastroenterol* 2012, 107:1190-1195; quiz 1196.
4. Wuerth BA, Rockey DC. Changing Epidemiology of Upper Gastrointestinal Hemorrhage in the Last Decade: A Nationwide Analysis. *Dig Dis Sci* 2018, 63:1286-1293.
5. Hearnshaw SA, Logan RF, Lowe D, Travis SP, Murphy MF, Palmer KR. Acute upper gastrointestinal bleeding in the UK: patient characteristics, diagnoses and outcomes in the 2007 UK audit. *Gut* 2011, 60:1327-1335.
6. Nahon S, Hagege H, Latrive JP, Rosa I, Nalet B, Bour B, Faroux R, Gower P, Arpurt JP, Denis J, et al. Epidemiological and prognostic factors involved in upper gastrointestinal bleeding: results of a French prospective multicenter study. *Endoscopy* 2012, 44:998-1008.
7. Hernandez-Diaz S, Rodriguez LA. Association between nonsteroidal anti-inflammatory drugs and upper gastrointestinal tract bleeding/perforation: an overview of epidemiologic studies published in the 1990s. *Arch Intern Med* 2000, 160:2093-2099.
8. Moore N, Pollack C, Butkerait P. Adverse drug reactions and drug-drug interactions with over-the-counter NSAIDs. *Ther Clin Risk Manag* 2015, 11:1061-1075.
9. Wilcox CM, Alexander LN, Cotsonis G. A prospective characterization of upper gastrointestinal hemorrhage presenting with hematochezia. *Am J Gastroenterol* 1997, 92:231-235.
10. Henry DA, Johnston A, Dobson A, Duggan J. Fatal peptic ulcer complications and the use of non-steroidal anti-inflammatory drugs, aspirin, and corticosteroids. *Br Med J (Clin Res Ed)* 1987, 295:1227-1229.
11. Coloma PM, Avillach P, Salvo F, Schuemie MJ, Ferrajolo C, Pariente A, Fourrier-Réglat A, Molokhia M, Patadia V, van der Lei J, et al. A Reference Standard for Evaluation of Methods for Drug Safety Signal Detection Using Electronic Healthcare Record Databases. *Drug Safety* 2013, 36:13-23.
12. Trifiro G, Pariente A, Coloma PM, Kors JA, Polimeni G, Miremont-Salame G, Catania MA, Salvo F, David A, Moore N, et al. Data mining on electronic health record databases for signal detection in pharmacovigilance: which events to monitor? *Pharmacoepidemiol Drug Saf* 2009, 18:1176-1184.
13. Ryan PB, Madigan D, Stang PE, Overhage JM, Racoosin JA, Hartzema AG. Empirical assessment of methods for risk identification in healthcare data: results from the experiments of the Observational Medical Outcomes Partnership. *Statistics in medicine* 2012, 31:4401-4415.

14. Ryan PB, Schuemie MJ, Welebob E, Duke J, Valentine S, Hartzema AG. Defining a reference set to support methodological research in drug safety. *Drug Saf* 2013, 36 Suppl 1:S33-47.
15. Castellsague J, Riera-Guardia N, Calingaert B, Varas-Lorenzo C, Fourier-Reglat A, Nicotra F, Sturkenboom M, Perez-Gutthann S, Safety of Non-Steroidal Anti-Inflammatory Drugs P. Individual NSAIDs and upper gastrointestinal complications: a systematic review and meta-analysis of observational studies (the SOS project). *Drug Saf* 2012, 35:1127-1146.
16. Valkhoff V. Use and Safety of Non-Steroidal Anti-Inflammatory Drugs and Aspirin. 2012.
17. Bezin J, Duong M, Lassalle R, Droz C, Pariente A, Blin P, Moore N. The national healthcare system claims databases in France, SNIIRAM and EGB: Powerful tools for pharmacoepidemiology. *Pharmacoepidemiol Drug Saf* 2017, 26:954-962.
18. Schuemie MJ, Gini R, Coloma PM, Straatman H, Herings RM, Pedersen L, Innocenti F, Mazzaglia G, Picelli G, van der Lei J, et al. Replication of the OMOP experiment in Europe: evaluating methods for risk identification in electronic health record databases. *Drug safety : an international journal of medical toxicology and drug experience* 2013, 36 Suppl 1:S159-169.
19. Armstrong B. A simple estimator of minimum detectable relative risk, sample size, or power in cohort studies. *Am J Epidemiol* 1987, 126:356-358.
20. Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. Robust empirical calibration of p-values using observational data. *Stat Med* 2016, 35:3883-3888.
21. Schuemie MJ, Ryan PB, DuMouchel W, Suchard MA, Madigan D. Interpreting observational studies: why empirical calibration is needed to correct p-values. *Stat Med* 2014, 33:209-218.
22. Gruber S, Tchetgen Tchetgen E. Limitations of empirical calibration of p-values using observational data. *Statistics in Medicine* 2016, 35:3869-3882.
23. Yola M, Lucien A. Evidence of the depletion of susceptibles effect in non-experimental pharmacoepidemiologic research. *Journal of Clinical Epidemiology* 1994, 47:731-737.
24. Petersen I, Douglas I, Whitaker H. Self controlled case series methods: an alternative to standard epidemiological study designs. *BMJ* 2016, 354:i4515.
25. Terdiman JP, Ostroff JW. Risk of persistent or recurrent and intractable upper gastrointestinal bleeding in the era of therapeutic endoscopy. *Am J Gastroenterol* 1997, 92:1805-1811.
26. Whitaker HJ, Steer CD, Farrington CP. Self-controlled case series studies: Just how rare does a rare non-recurrent outcome need to be? *Biometrical Journal* 2018, 60:1110-1120.
27. 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 safety : an international journal of medical toxicology and drug experience* 2013, 36 Suppl 1:S83-93.

28. Moore N, Gulmez SE, Larrey D, Pageaux GP, Lignot S, Lassalle R, Jove J, Pariente A, Blin P, Benichou J, 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. *Pharmacoepidemiology and drug safety* 2013, 22:160-167.
29. Moore N, Gulmez SE, Blin P, Lassalle R, Jove J, Theophile H, Begaud B, Larrey D, Benichou J. Relative risks from case-population data. *Epidemiology* 2013, 24:935-936.
30. Schuemie MJ, Ryan PB, Man KKC, Wong ICK, Suchard MA, Hripcsak G. A plea to stop using the case-control design in retrospective database studies. *Stat Med* 2019, 38:4199-4208.
31. Schneeweiss S, Suissa S. Discussion of Schuemie et al: "A plea to stop using the case-control design in retrospective database studies". *Stat Med* 2019, 38:4209-4212.
32. Cook EF, Goldman L. Performance of tests of significance based on stratification by a multivariate confounder score or by a propensity score. *J Clin Epidemiol* 1989, 42:317-324.
33. Sturmer T, Schneeweiss S, Brookhart MA, Rothman KJ, Avorn J, Glynn RJ. Analytic strategies to adjust confounding using exposure propensity scores and disease risk scores: nonsteroidal antiinflammatory drugs and short-term mortality in the elderly. *Am J Epidemiol* 2005, 161:891-898.
34. Clissold SP, Heel RC. Transdermal hyoscine (Scopolamine). A preliminary review of its pharmacodynamic properties and therapeutic efficacy. *Drugs* 1985, 29:189-207.
35. Horwitz RI, Feinstein AR. The problem of "protopathic bias" in case-control studies. *Am J Med* 1980, 68:255-258.
36. Al Sibae MR, McGuire BM. Current trends in the treatment of hepatic encephalopathy. *Ther Clin Risk Manag* 2009, 5:617-626.
37. Strom BL. *Pharmacoepidemiology*. 4th ed. Chichester ; Hoboken, NJ: J. Wiley; 2005.

TABLES

Table 1. Performances of the 10 most discriminant self-controlled case series variants for UGIB specific definition (1/10th sampled population)

Analysis ID	Sensitivity	Specificity	AUC	MSE	Coverage probability
2066	0.737	0.864	0.801	0.074	86%
2068	0.737	0.864	0.799	0.074	86%
2020	0.737	0.909	0.797	0.093	91%
2018	0.737	0.909	0.794	0.092	91%
2060	0.737	0.818	0.792	0.092	82%
2024	0.737	0.909	0.792	0.098	86%
2070	0.737	0.864	0.789	0.080	86%
2058	0.737	0.864	0.789	0.092	82%
2022	0.737	0.909	0.789	0.098	86%
2072	0.737	0.864	0.787	0.080	86%

A drug control was considered as positive for a given variant, when the left bound of its confidence interval was >1

AUC = area under the receiver operating characteristics curve

MSE = mean square error

Table 2. Performances of the 10 most discriminant self-controlled case series variants for UGIB sensitive definition (1/10th sampled population)

Analysis ID	Sensitivity	Specificity	AUC	MSE	Coverage probability
2090	0.579	0.864	0.837	5.905	81%
2092	0.579	0.864	0.837	69.639	81%
2068	0.684	0.818	0.828	0.291	82%
2094	0.632	0.818	0.828	2181.224	81%
2096	0.632	0.818	0.825	2201.157	81%
2066	0.684	0.818	0.823	0.291	82%
2082	0.579	0.864	0.823	5.911	86%
2084	0.579	0.864	0.823	69.637	81%
2058	0.684	0.818	0.818	0.301	82%
2060	0.684	0.818	0.818	0.301	82%

A drug control was considered as positive for a given variant, when the left bound of its confidence interval was >1

AUC = area under the receiver operating characteristics curve

MSE = mean square error

Table 3. Univariate logistic regression analysis of self-controlled case series parameters influencing on the discriminating power (UGIB specific definition, 1/10th sampled population)

	Variants with low AUC n=59	Variants with high AUC n=37	High vs. Low AUC OR [IC à 95%]	p	AUC of the univariate model
Age				0.8375	0.51
No	30 (50.8)	18 (48.6)	1		
Yes	29 (49.2)	19 (51.4)	1.09 [0.48 - 2.48]		
Seasonality				0.8375	0.51
No	30 (50.8)	18 (48.6)	1		
Yes	29 (49.2)	19 (51.4)	1.09 [0.48 - 2.48]		
Outcome				0.0087	0.64
All occurrences	36 (61.0)	12 (32.4)	1		
First occurrence	23 (39.0)	25 (67.6)	3.17 [1.34 - 7.50]		
Multiple drugs				<0.0001	0.80
No	43 (72.9)	5 (13.5)	1		
Yes	16 (27.1)	32 (86.5)	15.58 [5.30 - 45.77]		
Pre-Exposure Window				0.1404	0.62
No	16 (27.1)	16 (43.2)	1		
7 days	19 (32.2)	13 (35.1)	0.69 [0.26 - 1.86]		
30 days	24 (40.7)	8 (21.6)	0.35 [0.12 - 0.99]		
Risk window				<0.0001	0.73
Period of dispensing	40 (67.8)	8 (21.6)	1		
30 days from dispensing first day	19 (32.2)	29 (78.4)	7.21 [2.80 - 18.54]		

AUC = area under the receiver operating characteristics curve;
A high AUC was defined as an AUC≥0.75

Table 4. Univariate logistic regression analysis of self-controlled case series parameters influencing on the discriminating power (UGIB sensitive definition, 1/10th sampled population)

	Variants with low AUC n=56	Variants with high AUC n=40	High vs. Low AUC OR [IC à 95%]	p	AUC of the univariate model
Age				0.4183	0.54
No	26 (46.4)	22 (55.0)	1		
Yes	30 (53.6)	18 (45.0)	0.71 [0.32 - 1.60]		
Seasonality				1.0000	0.50
No	28 (50.0)	20 (50.0)	1		
Yes	28 (50.0)	20 (50.0)	1.00 [0.45 - 2.24]		
Outcome				0.1067	0.59
All occurrences	32 (57.1)	16 (40.0)	1		
First occurrence	24 (42.9)	24 (60.0)	1.97 [0.88 - 4.51]		
Multiple drugs				<0.0001	0.93
No	48 (85.7)	0 (0.0)	1		
Yes	8 (14.3)	40 (100.0)	462.25 [56.08 - >999.99]		
Pre-Exposure Window				0.3292	0.59
No	16 (28.6)	16 (40.0)	1		
7 days	18 (32.1)	14 (35.0)	0.78 [0.30 - 2.07]		
30 days	22 (39.3)	10 (25.0)	0.47 [0.17 - 1.26]		
Risk window				0.4183	0.54
Period of dispensing	26 (46.4)	22 (55.0)	1		
30 days from dispensing first day	30 (53.6)	18 (45.0)	0.71 [0.32 - 1.60]		

AUC = area under the receiver operating characteristics curve;
A high AUC was defined as an AUC≥0.78

Table 5. . Negative controls with significant association with UGIB according to traditional and calibrated p-value (SCCS variant 2066, unsampled population)

ATC	INN	UGIB specific definition		UGIB sensitive definition	
		Traditional p-value	Calibrated p-value	Traditional p-value	Calibrated p-value
A01AB09	MICONAZOLE	0.046	0.476	0.167	0.780
A02BX02	SUCRALFATE	<0.001	0.018	<0.001	0.034
A04AD01	SCOPOLAMINE	0.011	0.009	0.001	0.001
A06AD11	LACTULOSE	<0.001	0.100	<0.001	0.097
A10BH01	SITAGLIPTINE	<0.001	0.359	<0.001	0.257
B03XA01	ERYTHROPOIETIN	<0.001	0.092	<0.001	0.038
J01XE01	NITROFURANTOINE	<0.001	0.418	<0.001	0.523
M03BA03	METHOCARBAMOL	<0.001	0.416	<0.001	0.383
N05CF01	ZOPICLONE	<0.001	0.920	<0.001	0.908
R06AX13	LORATADINE	0.007	0.775	0.008	0.845

ATC = Anatomical Therapeutic and Chemical classification

INN = International nonproprietary name

Table 6. Traditional and calibrated p-value for positive controls according to UGIB definition (SCCS variant 2066, unsampled population)

ATC	INN	UGIB specific definition		UGIB sensitive definition	
		Traditional p-value	Calibrated p-value	Traditional p-value	Calibrated p-value
A12BA01	POTASSIUM CHLORIDE	<0.001	0.318	<0.001	0.303
B01AB01	HEPARIN	<0.001	<0.001	<0.001	<0.001
B01AC04	CLOPIDOGREL	<0.001	0.044	<0.001	0.034
H02AB06	PREDNISOLONE	<0.001	0.133	<0.001	0.111
J01FF01	CLINDAMYCIN	0.412	0.991	0.639	0.801
M01AB01	INDOMETHACIN	0.044	0.285	0.011	0.175
M01AB02	SULINDAC	0.549	0.782	0.122	0.293
M01AB08	ETODOLAC	0.856	0.724	0.792	0.735
M01AC01	PIROXICAM	<0.001	<0.001	<0.001	<0.001
M01AC06	MELOXICAM	<0.001	0.002	<0.001	0.002
M01AE01	IBUPROFEN	<0.001	0.125	<0.001	0.114
M01AE02	NAPROXEN	<0.001	<0.001	<0.001	<0.001
M01AE03	KETOPROFEN	<0.001	<0.001	<0.001	<0.001
M01AE04	FENOPROFEN	0.043	0.194	0.037	0.196
M01AE09	FLURBIPROFEN	<0.001	<0.001	<0.001	<0.001
M01AG01	MEFENAMIC ACID	0.140	0.109	0.518	0.344
M01AX01	NABUMETONE	<0.001	0.247	<0.001	0.299
N02BA01	ACETYLSALICYLIC ACID	<0.001	0.011	<0.001	0.009
N06AB03	FLUOXETINE	<0.001	0.182	<0.001	0.137
N06AB04	CITALOPRAM	<0.001	0.357	<0.001	0.339
N06AB06	SERTRALINE	<0.001	0.058	<0.001	0.034
N06AB10	ESCITALOPRAM	<0.001	0.038	<0.001	0.073

ATC = Anatomical Therapeutic and Chemical classification

INN = International nonproprietary name

FIGURES

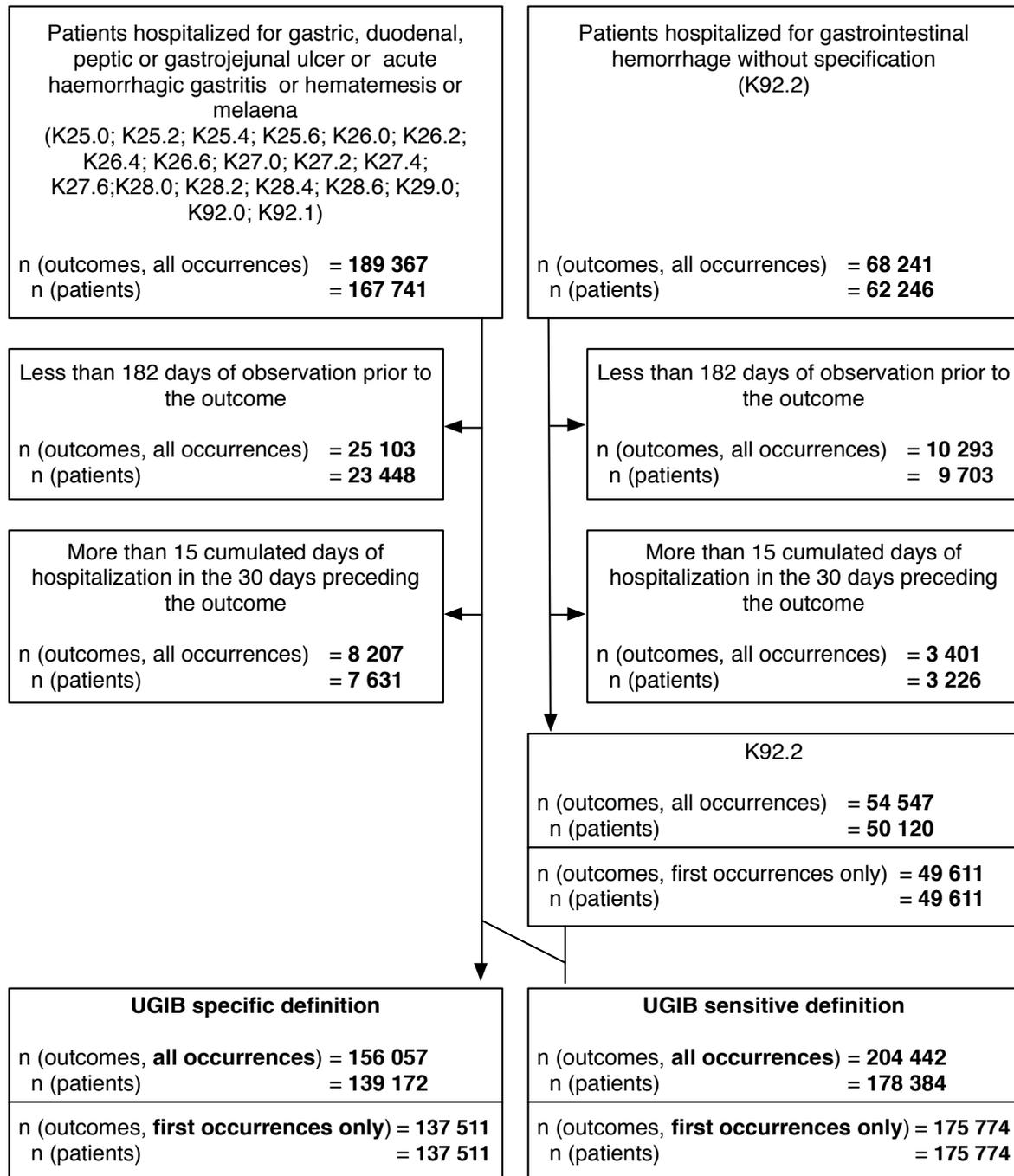


Figure 1. Selection of patients with upper-gastrointestinal bleeding (UGIB) according to sensitive and specific definitions

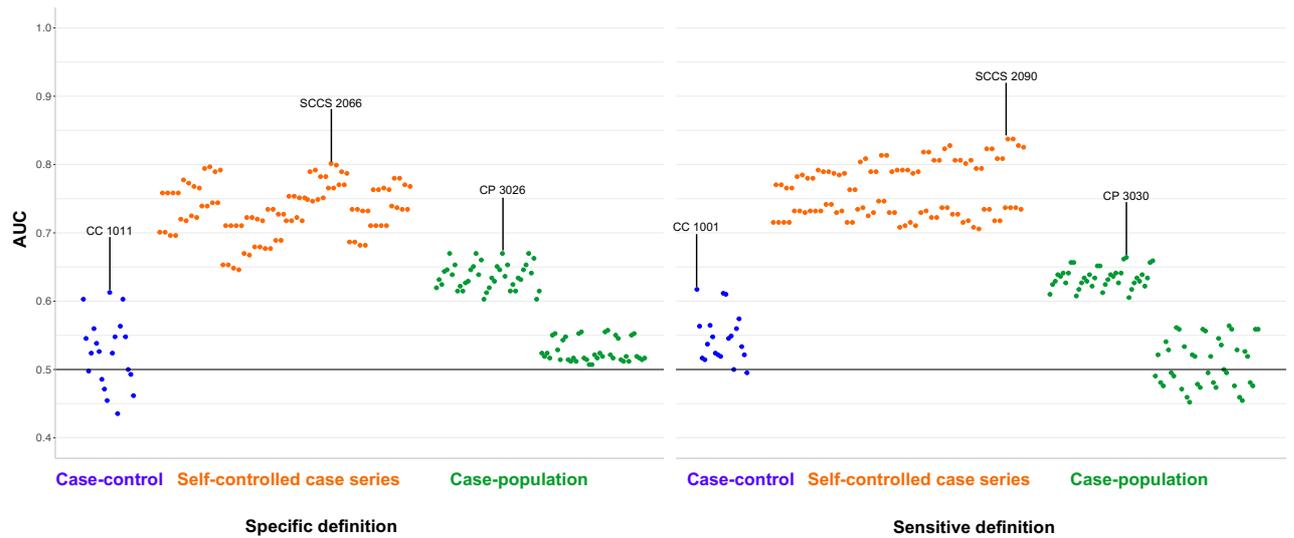


Figure 2. Area under the receiver operating characteristics curve (AUC) for case-control (CC), self-controlled case series (SCCS) and case-population (CP) approaches in the 1/10th sampled population according to upper-gastro-intestinal bleeding definition. For each approach, variant with the highest AUC is mentioned.

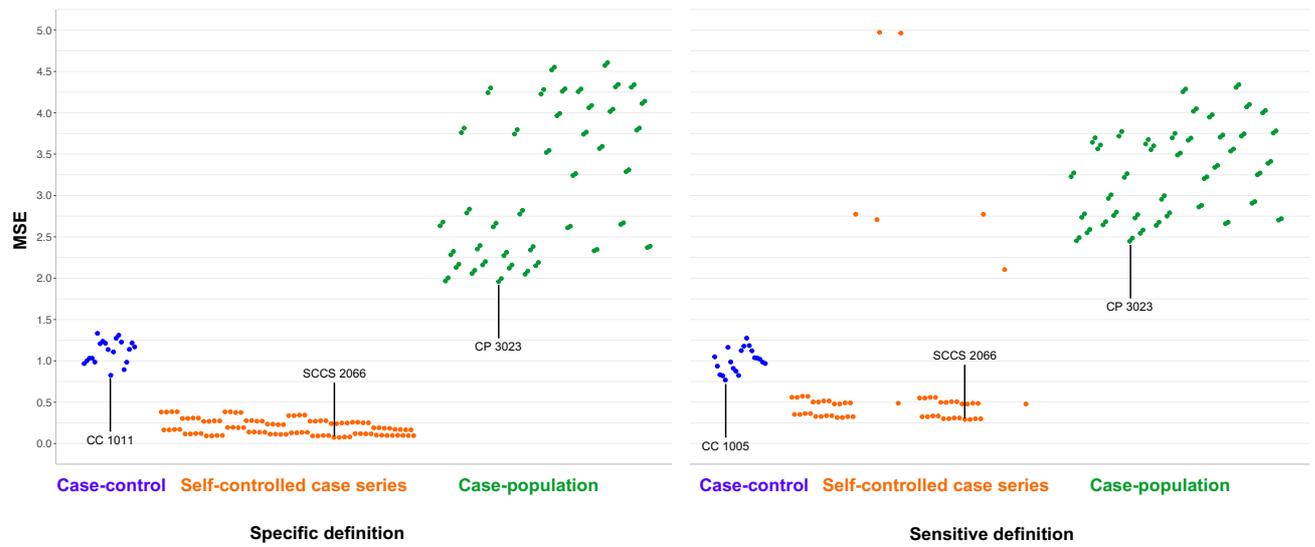


Figure 3. Means square error (MSE) for case-control (CC), self-controlled case series (SCCS) and case-population (CP) approaches in the 1/10th sampled population according to upper-gastro-intestinal bleeding definition. For each approach, variant with the lowest MSE is mentioned.

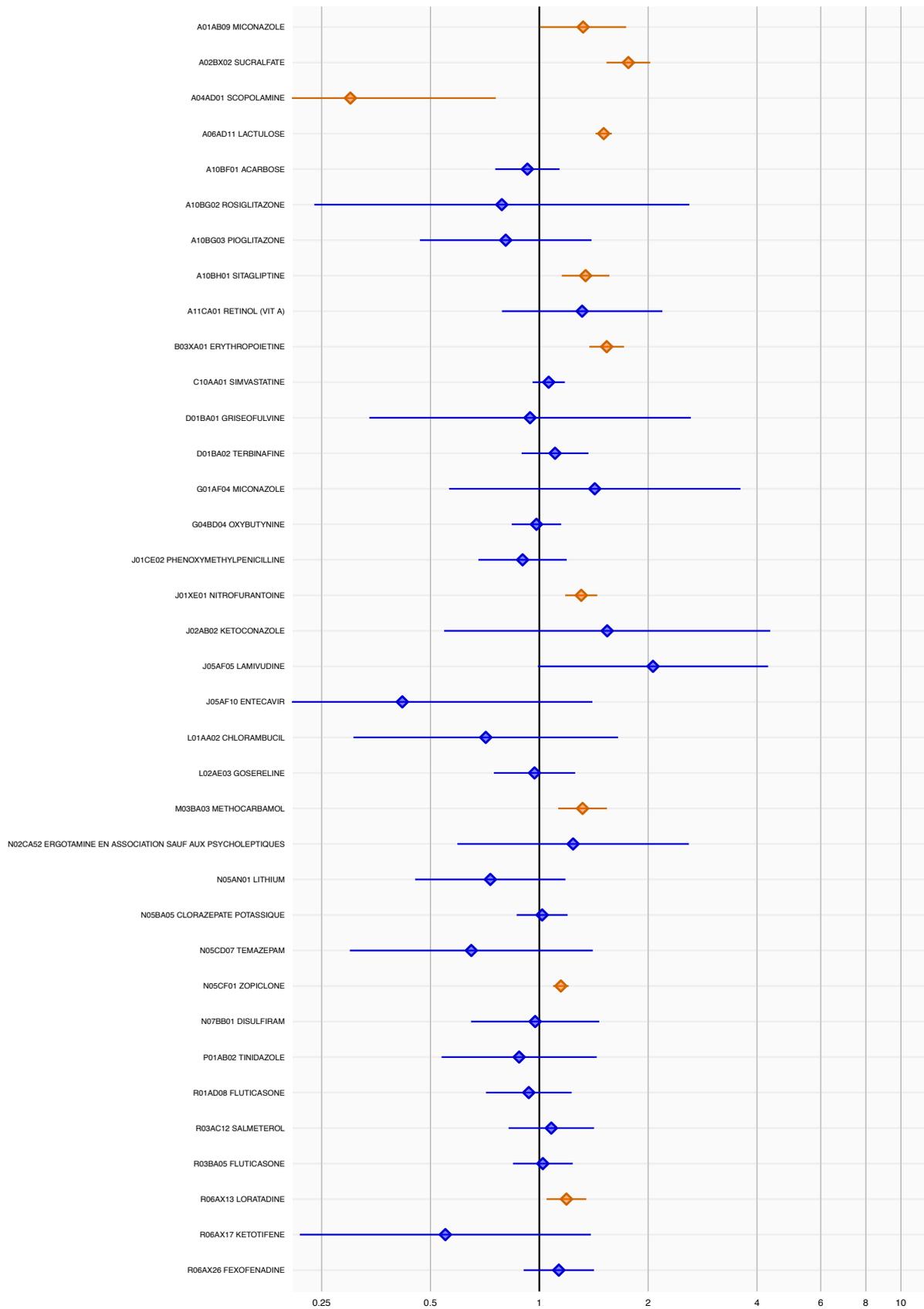


Figure 4. Point estimates of negative controls for the specific definition of upper-gastro-intestinal bleeding (self-controlled case series variant 2066, unsampled population). Estimates that are significantly different from 1 ($\alpha = 0.05$) are marked in orange, others are marked in blue.

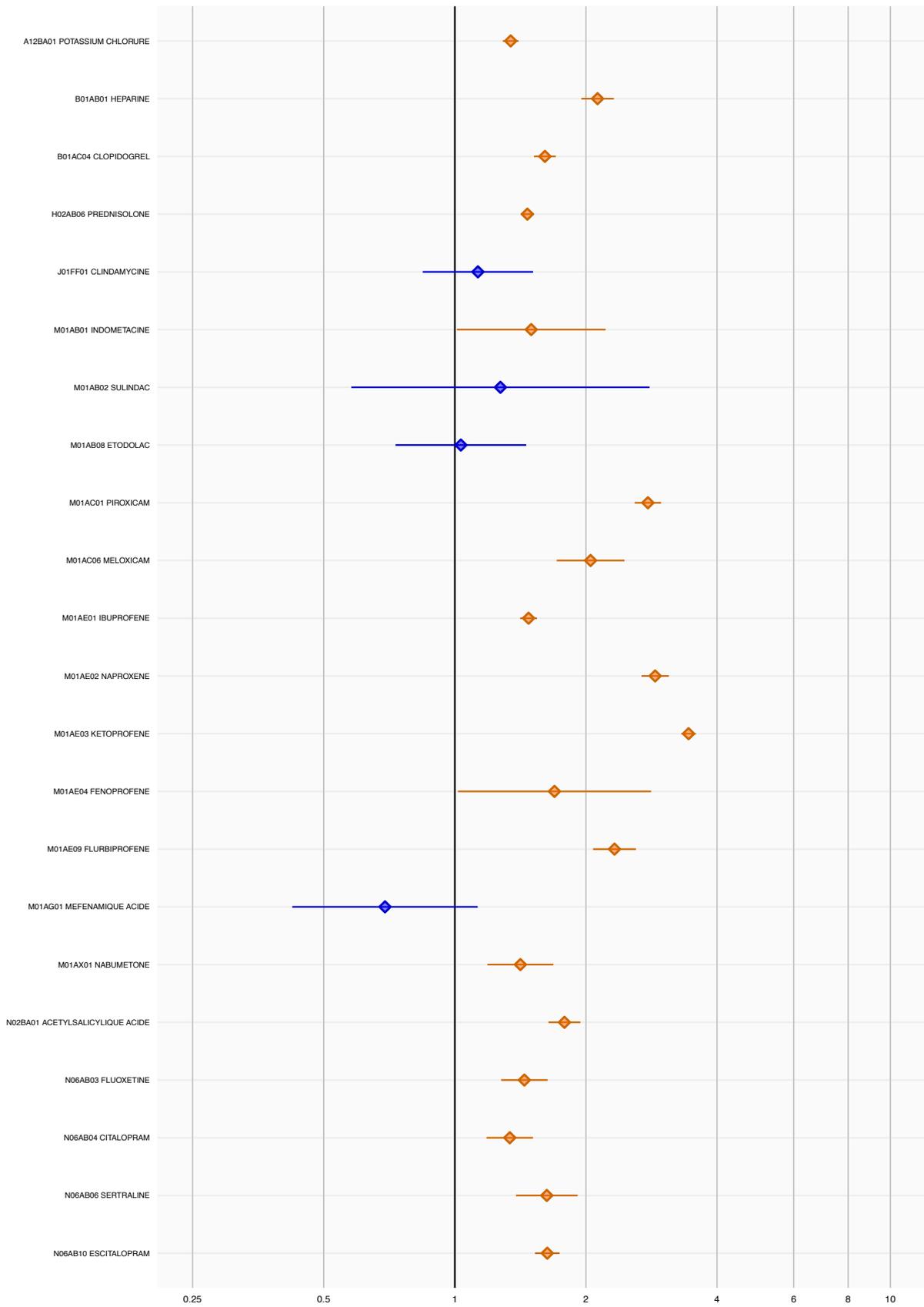


Figure 5. Point estimates of positive controls for the specific definition of upper-gastro-intestinal bleeding (self-controlled case series variant 2066, unsampled population). Estimates that are significantly different from 1 ($\alpha = 0.05$) are marked in orange, others are marked in blue.

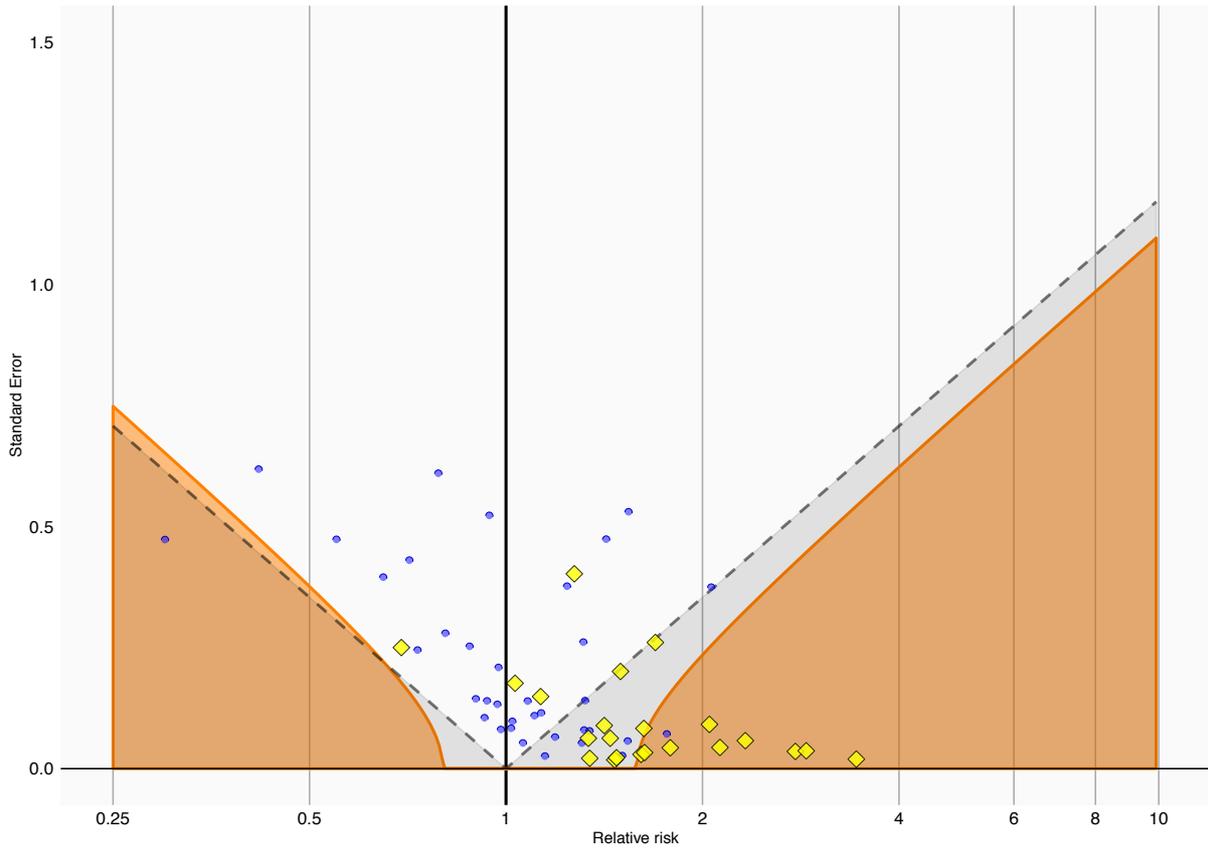


Figure 6. Point estimates from self-controlled case series variant 2066, (upper-gastro-intestinal bleeding specific definition, unsampled population). Estimates below the dashed line have $p < 0.05$ using traditional p -value calculation. Estimates in the orange area have $p < 0.05$ using calibrated p -value calculation. Blue dots indicate negative controls. Yellow diamonds indicate positive controls.

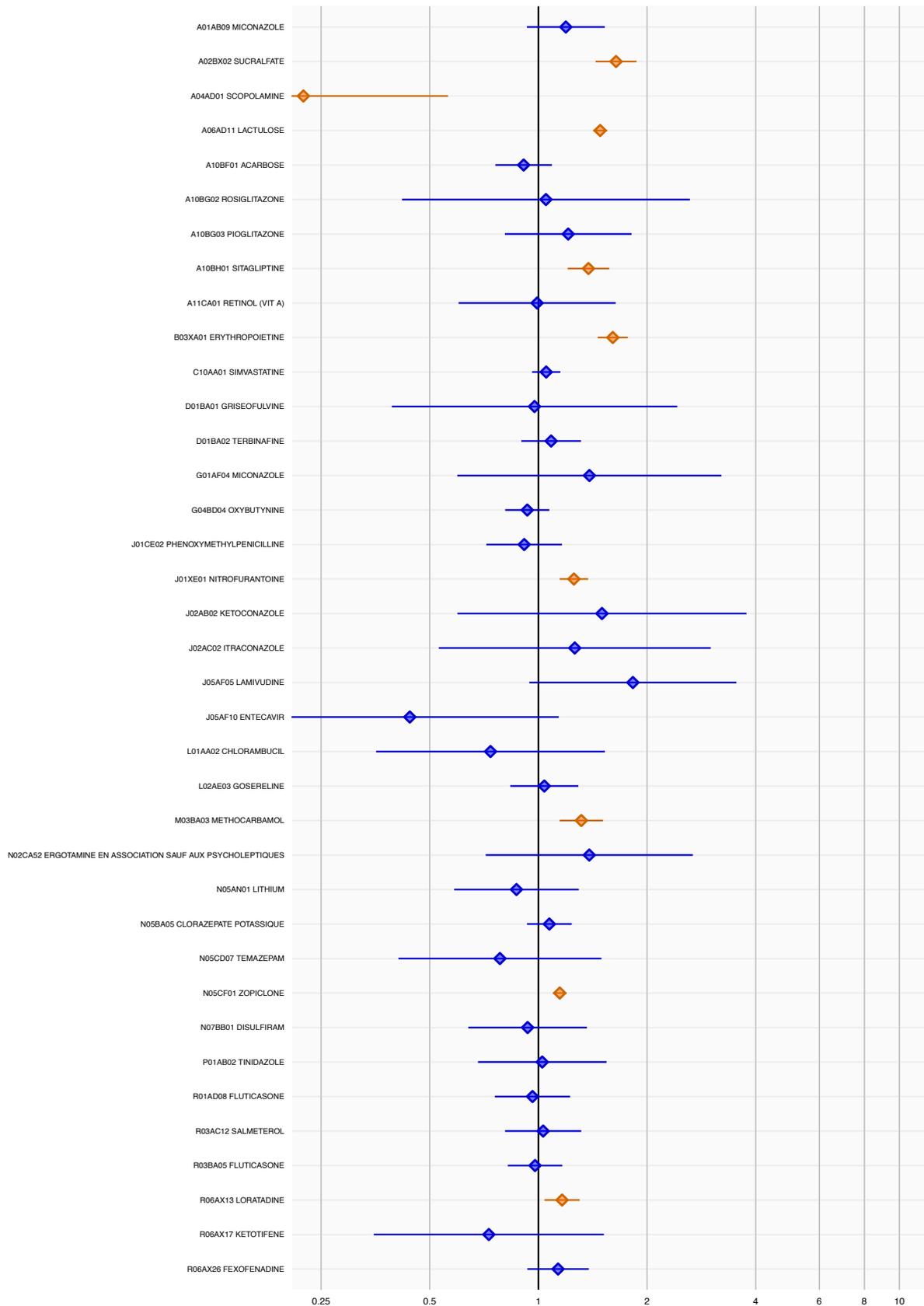


Figure 7. Point estimates of negative controls for the sensitive definition of upper-gastro-intestinal bleeding (self-controlled case series variant 2066, unsampled population). Estimates that are significantly different from 1 ($\alpha = 0.05$) are marked in orange, others are marked in blue.

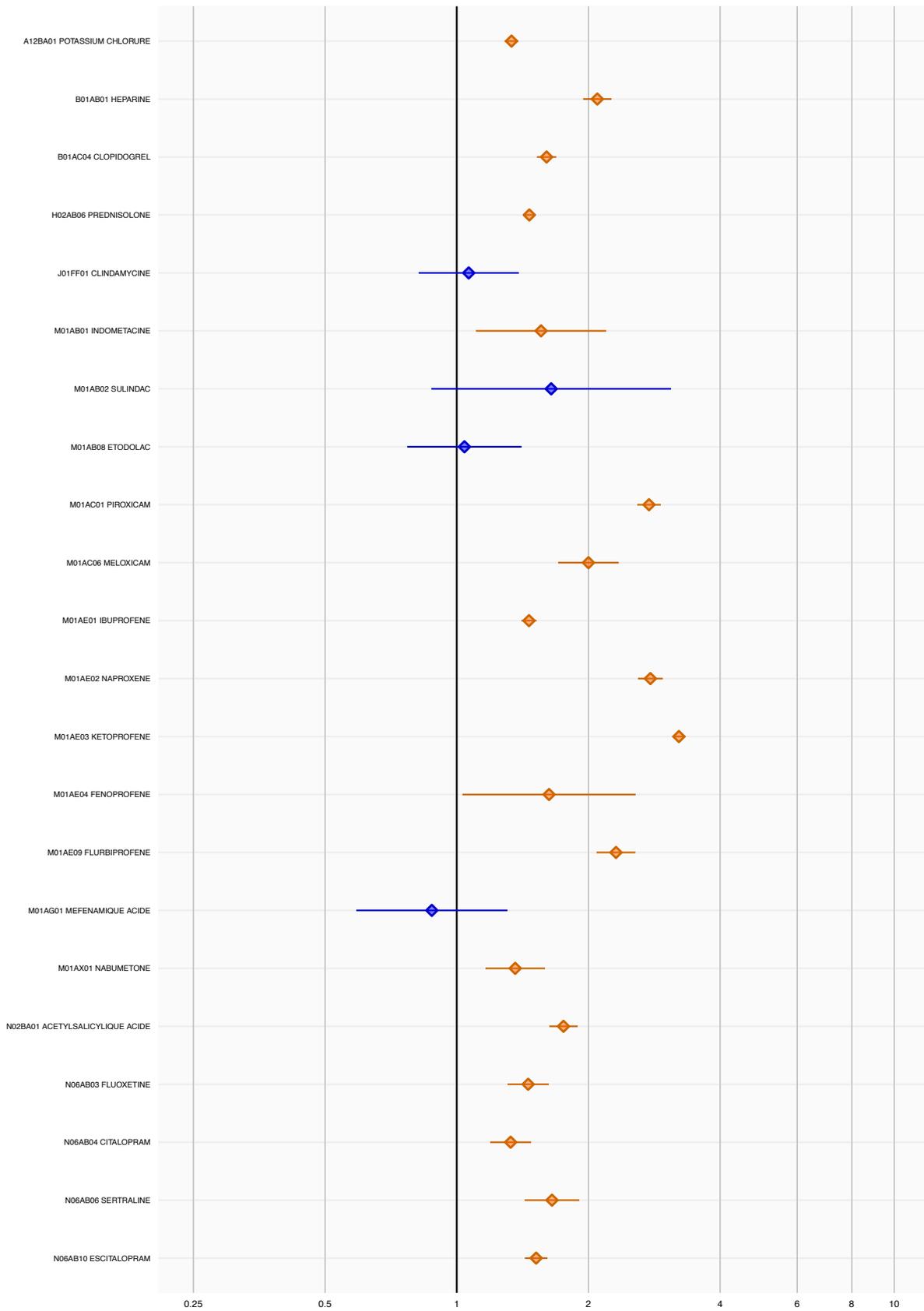


Figure 8. Point estimates of positive controls for the sensitive definition of upper-gastro-intestinal bleeding (self-controlled case series variant 2066, unsampled population). Estimates that are significantly different from 1 ($\alpha = 0.05$) are marked in orange, others are marked in blue.

APPENDICES

APPENDIX 1 SPECIFIC AND SENSITIVE DEFINITIONS OF UPPER GASTROINTESTINAL BLEEDING (UGIB)

ICD10	Definition	Narrow definition	Broad definition
K25.0	Gastric ulcer; Acute with haemorrhage	yes	yes
K25.2	Gastric ulcer; Acute with both haemorrhage and perforation	yes	yes
K25.4	Gastric ulcer; Chronic or unspecified with haemorrhage	yes	yes
K25.6	Gastric ulcer; Chronic or unspecified with both haemorrhage and perforation	yes	yes
K26.0	Duodenal ulcer; Acute with haemorrhage	yes	yes
K26.2	Duodenal ulcer; Acute with both haemorrhage and perforation	yes	yes
K26.4	Duodenal ulcer; Chronic or unspecified with haemorrhage	yes	yes
K26.6	Duodenal ulcer; Chronic or unspecified with both haemorrhage and perforation	yes	yes
K27.0	Peptic ulcer, site unspecified; Acute with haemorrhage	yes	yes
K27.2	Peptic ulcer, site unspecified; Acute with both haemorrhage and perforation	yes	yes
K27.4	Peptic ulcer, site unspecified; Chronic or unspecified with haemorrhage	yes	yes
K27.6	Peptic ulcer, site unspecified; Chronic or unspecified with both haemorrhage and perforation	yes	yes
K28.0	Gastrojejunal ulcer; Acute with haemorrhage	yes	yes
K28.2	Gastrojejunal ulcer; Acute with both haemorrhage and perforation	yes	yes
K28.4	Gastrojejunal ulcer; Chronic or unspecified with haemorrhage	yes	yes
K28.6	Gastrojejunal ulcer; Chronic or unspecified with both haemorrhage and perforation	yes	yes
K29.0	Acute haemorrhagic gastritis	yes	yes
K92.0	Haematemesis	yes	yes
K92.1	Melaena	yes	yes
K92.2	Gastrointestinal haemorrhage, unspecified		yes

APPENDIX 2 DETECTABLE POSITIVE AND NEGATIVE CONTROLS WITH CORRESPONDING MINIMUM DETECTABLE RELATIVE RISK

ATC code	International nonproprietary name	Source reference set	Type of control	Specific definition		Sensitive definition	
				MDRR unsampled population	MDRR 1/10 th sampled population	MDRR unsampled population	MDRR 1/10 th sampled population
A01AB09	MICONAZOLE	OMOP	Negative	1.09	1.30	1.07	1.23
A02BX02	SUCRALFATE	OMOP	Negative	1.04	1.12	1.04	1.12
A04AD01	SCOPOLAMINE	OMOP	Negative	1.18	1.64	1.15	1.53
A06AD11	LACTULOSE	OMOP	Negative	1.02	1.05	1.01	1.04
A10BF01	ACARBOSE	OMOP	Negative	1.05	1.16	1.04	1.14
A10BG02	ROSIGLITAZONE	OMOP	Negative	1.15	1.60	1.13	1.46
A10BG03	PIOGLITAZONE	OMOP	Negative	1.08	1.26	1.07	1.22
A10BH01	SITAGLIPTINE	OMOP	Negative	1.04	1.13	1.03	1.11
A11CA01	RETINOL (VIT A)	OMOP	Negative	1.15	1.46	1.13	1.47
A12BA01	POTASSIUM CHLORURE	OMOP	Positive	1.01	1.04	1.01	1.04
B01AB01	HEPARINE	EU-ADR	Positive	1.03	1.09	1.02	1.08
B01AC04	CLOPIDOGREL	OMOP	Positive	1.02	1.05	1.01	1.05
B03XA01	ERYTHROPOIETINE	OMOP	Negative	1.04	1.11	1.03	1.10
C09AA13	MOEXIPRIL	OMOP	Negative	1.69		1.61	2.47
C10AA01	SIMVASTATINE	EU-ADR	Negative	1.03	1.08	1.02	1.07
D01BA01	GRISEOFULVINE	OMOP	Negative	1.17	1.48	1.15	1.54
D01BA02	TERBINAFINE	OMOP	Negative	1.04	1.15	1.04	1.12
G01AF04	MICONAZOLE	OMOP	Negative	1.21	1.67	1.18	1.64
G04BD04	OXYBUTYNNINE	OMOP	Negative	1.04	1.13	1.04	1.12
H02AB06	PREDNISOLONE	EU-ADR	Positive	1.01	1.04	1.01	1.03
J01CE02	PHENOXYMETHYLPENICILLINE	OMOP	Negative	1.07	1.25	1.06	1.22
J01FF01	CLINDAMYCINE	OMOP	Positive	1.07	1.23	1.06	1.19
J01XE01	NITROFURANTOINE	OMOP	Negative	1.03	1.09	1.02	1.08
J02AB02	KETOCONAZOLE	OMOP	Negative	1.23	1.75	1.19	1.72
J02AC02	ITRACONAZOLE	OMOP	Negative	1.32	2.08	1.28	2.47
J05AF01	ZIDOVUDINE	OMOP	Negative	1.97		1.77	5.76
J05AF04	STAVUDINE	OMOP	Negative	2.34	3.27	2.02	3.96
J05AF05	LAMIVUDINE	OMOP	Negative	1.25	1.85	1.21	1.65

J05AF06	ABACAVIR	OMOP	Negative	1.41	3.27	1.33	2.65
J05AF10	ENTECAVIR	OMOP	Negative	1.21	1.69	1.18	1.54
J05AG01	NEVIRAPINE	OMOP	Negative	1.45	2.34	1.35	1.65
L01AA02	CHLORAMBUCIL	OMOP	Negative	1.18	1.72	1.16	1.49
L02AE03	GOSERELINE	EU-ADR	Negative	1.11	1.38	1.10	1.33
M01AB01	INDOMETACINE	OMOP	Positive	1.11	1.39	1.10	1.31
M01AB02	SULINDAC	OMOP	Positive	1.18	1.80	1.16	1.61
M01AB08	ETODOLAC	OMOP	Positive	1.07	1.24	1.06	1.21
M01AC01	PIROXICAM	OMOP	Positive	1.02	1.07	1.02	1.06
M01AC06	MELOXICAM	OMOP	Positive	1.05	1.17	1.05	1.14
M01AE01	IBUPROFENE	OMOP	Positive	1.01	1.04	1.01	1.04
M01AE02	NAPROXENE	OMOP	Positive	1.02	1.08	1.02	1.07
M01AE03	KETOPROFENE	OMOP	Positive	1.01	1.05	1.01	1.04
M01AE04	FENOPROFENE	OMOP	Positive	1.13	1.42	1.12	1.40
M01AE09	FLURBIPROFENE	OMOP	Positive	1.04	1.11	1.03	1.10
M01AG01	MEFENAMIQUE ACIDE	OMOP	Positive	1.09	1.27	1.08	1.27
M01AX01	NABUMETONE	OMOP	Positive	1.04	1.13	1.03	1.11
M03BA03	METHOCARBAMOL	OMOP	Negative	1.04	1.12	1.03	1.10
N02BA01	ACETYLSALICYLIQUE ACIDE	EU-ADR	Positive	1.02	1.08	1.02	1.07
N02CA52	ERGOTAMINE EN ASSOCIATION SAUF AUX PSYCHOLEPTIQUES	OMOP	Negative	1.20	1.65	1.18	1.67
N05AN01	LITHIUM	OMOP	Negative	1.14	1.47	1.12	1.40
N05BA05	CLORAZEPATE POTASSIQUE	OMOP	Negative	1.04	1.12	1.04	1.12
N05CD07	TEMAZEPAM	OMOP	Negative	1.13	1.37	1.11	1.42
N05CF01	ZOPICLONE	EU-ADR	Negative	1.02	1.05	1.01	1.04
N06AB03	FLUOXETINE	OMOP	Positive	1.03	1.11	1.03	1.10
N06AB04	CITALOPRAM	OMOP	Positive	1.03	1.11	1.03	1.09
N06AB06	SERTRALINE	OMOP	Positive	1.04	1.14	1.04	1.12
N06AB10	ESCITALOPRAM	OMOP	Positive	1.02	1.06	1.02	1.05
N07BB01	DISULFIRAM	OMOP	Negative	1.09	1.27	1.08	1.25
P01AB02	TINIDAZOLE	OMOP	Negative	1.09	1.29	1.08	1.30
R01AD08	FLUTICASONNE	OMOP	Negative	1.06	1.19	1.05	1.18
R03AC12	SALMETEROL	OMOP	Negative	1.07	1.23	1.06	1.21
R03BA05	FLUTICASONNE	OMOP	Negative	1.05	1.15	1.04	1.13
R06AX13	LORATADINE	OMOP	Negative	1.03	1.10	1.03	1.08
R06AX17	KETOTIFENE	OMOP	Negative	1.15	1.49	1.13	1.44
R06AX26	FEXOFENADINE	EU-ADR	Negative	1.05	1.16	1.04	1.13

ATC = Anatomical Therapeutic and Chemical classification
MDRR = Minimum Detectable Relative Risk

Detectable positive controls	22	19	22	19
Detectable negative controls	36	22	37	22

APPENDIX 3 DESIGN VARIANT DESCRIPTION

Design approach	Analysis_ID	Controls per case	Required observation time prior to outcome	Outcomes to include	Risk windows	Case-control matching strategy	Remove unmatched case	Metric
CC	1001	up to 2	182d	All occurrences	7d and 0d	Age and gender	Yes	Odds ratio
CC	1002	up to 2	182d	All occurrences	30d and 0d	Age and gender	Yes	Odds ratio
CC	1003	up to 2	182d	All occurrences	60d and 0d	Age and gender	Yes	Odds ratio
CC	1004	up to 2	182d	All occurrences	60d and 7d	Age and gender	Yes	Odds ratio
CC	1005	up to 2	182d	All occurrences	60d and 15d	Age and gender	Yes	Odds ratio
CC	1006	up to 10	182d	All occurrences	7d and 0d	Age and gender	Yes	Odds ratio
CC	1007	up to 10	182d	All occurrences	30d and 0d	Age and gender	Yes	Odds ratio
CC	1008	up to 10	182d	All occurrences	60d and 0d	Age and gender	Yes	Odds ratio
CC	1009	up to 10	182d	All occurrences	60d and 7d	Age and gender	Yes	Odds ratio
CC	1010	up to 10	182d	All occurrences	60d and 15d	Age and gender	Yes	Odds ratio
CC	1011	up to 2	182d	First occurrence	7d and 0d	Age and gender	Yes	Odds ratio
CC	1012	up to 2	182d	First occurrence	30d and 0d	Age and gender	Yes	Odds ratio
CC	1013	up to 2	182d	First occurrence	60d and 0d	Age and gender	Yes	Odds ratio
CC	1014	up to 2	182d	First occurrence	60d and 7d	Age and gender	Yes	Odds ratio
CC	1015	up to 2	182d	First occurrence	60d and 15d	Age and gender	Yes	Odds ratio
CC	1016	up to 10	182d	First occurrence	7d and 0d	Age and gender	Yes	Odds ratio
CC	1017	up to 10	182d	First occurrence	30d and 0d	Age and gender	Yes	Odds ratio
CC	1018	up to 10	182d	First occurrence	60d and 0d	Age and gender	Yes	Odds ratio
CC	1019	up to 10	182d	First occurrence	60d and 7d	Age and gender	Yes	Odds ratio
CC	1020	up to 10	182d	First occurrence	60d and 15d	Age and gender	Yes	Odds ratio

Design approach	Analysis_ID	Outcomes to include	Naive period	Risk window	Pre-exposure window	Include age into the model	Include seasonality into the model	Adjusting for multiple drugs
SCCS	2001	All occurrences	182d	30d from dispensing first day	30d	No	No	No
SCCS	2002	All occurrences	182d	30d from dispensing first day	30d	No	No	Yes
SCCS	2003	All occurrences	182d	30d from dispensing first day	30d	No	Yes	No
SCCS	2004	All occurrences	182d	30d from dispensing first day	30d	No	Yes	Yes
SCCS	2005	All occurrences	182d	30d from dispensing first day	30d	Yes	No	No
SCCS	2006	All occurrences	182d	30d from dispensing first day	30d	Yes	No	Yes
SCCS	2007	All occurrences	182d	30d from dispensing first day	30d	Yes	Yes	No
SCCS	2008	All occurrences	182d	30d from dispensing first day	30d	Yes	Yes	Yes
SCCS	2009	All occurrences	182d	30d from dispensing first day	7d	No	No	No
SCCS	2010	All occurrences	182d	30d from dispensing first day	7d	No	No	Yes
SCCS	2011	All occurrences	182d	30d from dispensing first day	7d	No	Yes	No
SCCS	2012	All occurrences	182d	30d from dispensing first day	7d	No	Yes	Yes
SCCS	2013	All occurrences	182d	30d from dispensing first day	7d	Yes	No	No
SCCS	2014	All occurrences	182d	30d from dispensing first day	7d	Yes	No	Yes
SCCS	2015	All occurrences	182d	30d from dispensing first day	7d	Yes	Yes	No
SCCS	2016	All occurrences	182d	30d from dispensing first day	7d	Yes	Yes	Yes
SCCS	2017	All occurrences	182d	30d from dispensing first day	No	No	No	No
SCCS	2018	All occurrences	182d	30d from dispensing first day	No	No	No	Yes
SCCS	2019	All occurrences	182d	30d from dispensing first day	No	No	Yes	No
SCCS	2020	All occurrences	182d	30d from dispensing first day	No	No	Yes	Yes
SCCS	2021	All occurrences	182d	30d from dispensing first day	No	Yes	No	No
SCCS	2022	All occurrences	182d	30d from dispensing first day	No	Yes	No	Yes
SCCS	2023	All occurrences	182d	30d from dispensing first day	No	Yes	Yes	No
SCCS	2024	All occurrences	182d	30d from dispensing first day	No	Yes	Yes	Yes
SCCS	2025	All occurrences	182d	Period of dispensing	30d	No	No	No
SCCS	2026	All occurrences	182d	Period of dispensing	30d	No	No	Yes
SCCS	2027	All occurrences	182d	Period of dispensing	30d	No	Yes	No
SCCS	2028	All occurrences	182d	Period of dispensing	30d	No	Yes	Yes
SCCS	2029	All occurrences	182d	Period of dispensing	30d	Yes	No	No
SCCS	2030	All occurrences	182d	Period of dispensing	30d	Yes	No	Yes
SCCS	2031	All occurrences	182d	Period of dispensing	30d	Yes	Yes	No
SCCS	2032	All occurrences	182d	Period of dispensing	30d	Yes	Yes	Yes
SCCS	2033	All occurrences	182d	Period of dispensing	7d	No	No	No
SCCS	2034	All occurrences	182d	Period of dispensing	7d	No	No	Yes
SCCS	2035	All occurrences	182d	Period of dispensing	7d	No	Yes	No
SCCS	2036	All occurrences	182d	Period of dispensing	7d	No	Yes	Yes

SCCS	2037	All occurrences	182d	Period of dispensing	7d	Yes	No	No
SCCS	2038	All occurrences	182d	Period of dispensing	7d	Yes	No	Yes
SCCS	2039	All occurrences	182d	Period of dispensing	7d	Yes	Yes	No
SCCS	2040	All occurrences	182d	Period of dispensing	7d	Yes	Yes	Yes
SCCS	2041	All occurrences	182d	Period of dispensing	No	No	No	No
SCCS	2042	All occurrences	182d	Period of dispensing	No	No	No	Yes
SCCS	2043	All occurrences	182d	Period of dispensing	No	No	Yes	No
SCCS	2044	All occurrences	182d	Period of dispensing	No	No	Yes	Yes
SCCS	2045	All occurrences	182d	Period of dispensing	No	Yes	No	No
SCCS	2046	All occurrences	182d	Period of dispensing	No	Yes	No	Yes
SCCS	2047	All occurrences	182d	Period of dispensing	No	Yes	Yes	No
SCCS	2048	All occurrences	182d	Period of dispensing	No	Yes	Yes	Yes
SCCS	2049	First occurrence	182d	30d from dispensing first day	30d	No	No	No
SCCS	2050	First occurrence	182d	30d from dispensing first day	30d	No	No	Yes
SCCS	2051	First occurrence	182d	30d from dispensing first day	30d	No	Yes	No
SCCS	2052	First occurrence	182d	30d from dispensing first day	30d	No	Yes	Yes
SCCS	2053	First occurrence	182d	30d from dispensing first day	30d	Yes	No	No
SCCS	2054	First occurrence	182d	30d from dispensing first day	30d	Yes	No	Yes
SCCS	2055	First occurrence	182d	30d from dispensing first day	30d	Yes	Yes	No
SCCS	2056	First occurrence	182d	30d from dispensing first day	30d	Yes	Yes	Yes
SCCS	2057	First occurrence	182d	30d from dispensing first day	7d	No	No	No
SCCS	2058	First occurrence	182d	30d from dispensing first day	7d	No	No	Yes
SCCS	2059	First occurrence	182d	30d from dispensing first day	7d	No	Yes	No
SCCS	2060	First occurrence	182d	30d from dispensing first day	7d	No	Yes	Yes
SCCS	2061	First occurrence	182d	30d from dispensing first day	7d	Yes	No	No
SCCS	2062	First occurrence	182d	30d from dispensing first day	7d	Yes	No	Yes
SCCS	2063	First occurrence	182d	30d from dispensing first day	7d	Yes	Yes	No
SCCS	2064	First occurrence	182d	30d from dispensing first day	7d	Yes	Yes	Yes
SCCS	2065	First occurrence	182d	30d from dispensing first day	No	No	No	No
SCCS	2066	First occurrence	182d	30d from dispensing first day	No	No	No	Yes
SCCS	2067	First occurrence	182d	30d from dispensing first day	No	No	Yes	No
SCCS	2068	First occurrence	182d	30d from dispensing first day	No	No	Yes	Yes
SCCS	2069	First occurrence	182d	30d from dispensing first day	No	Yes	No	No
SCCS	2070	First occurrence	182d	30d from dispensing first day	No	Yes	No	Yes
SCCS	2071	First occurrence	182d	30d from dispensing first day	No	Yes	Yes	No
SCCS	2072	First occurrence	182d	30d from dispensing first day	No	Yes	Yes	Yes
SCCS	2073	First occurrence	182d	Period of dispensing	30d	No	No	No
SCCS	2074	First occurrence	182d	Period of dispensing	30d	No	No	Yes
SCCS	2075	First occurrence	182d	Period of dispensing	30d	No	Yes	No
SCCS	2076	First occurrence	182d	Period of dispensing	30d	No	Yes	No
SCCS	2077	First occurrence	182d	Period of dispensing	30d	Yes	No	Yes

SCCS	2078	First occurrence	182d	Period of dispensing	30d	No	Yes	No	Yes
SCCS	2079	First occurrence	182d	Period of dispensing	30d	Yes	Yes	Yes	No
SCCS	2080	First occurrence	182d	Period of dispensing	30d	Yes	Yes	Yes	Yes
SCCS	2081	First occurrence	182d	Period of dispensing	7d	No	No	No	No
SCCS	2082	First occurrence	182d	Period of dispensing	7d	No	No	No	Yes
SCCS	2083	First occurrence	182d	Period of dispensing	7d	No	No	Yes	No
SCCS	2084	First occurrence	182d	Period of dispensing	7d	No	Yes	Yes	Yes
SCCS	2085	First occurrence	182d	Period of dispensing	7d	Yes	Yes	No	No
SCCS	2086	First occurrence	182d	Period of dispensing	7d	Yes	Yes	No	Yes
SCCS	2087	First occurrence	182d	Period of dispensing	7d	Yes	Yes	Yes	No
SCCS	2088	First occurrence	182d	Period of dispensing	7d	Yes	Yes	Yes	Yes
SCCS	2089	First occurrence	182d	Period of dispensing	No	No	No	No	No
SCCS	2090	First occurrence	182d	Period of dispensing	No	No	No	No	Yes
SCCS	2091	First occurrence	182d	Period of dispensing	No	No	No	Yes	No
SCCS	2092	First occurrence	182d	Period of dispensing	No	No	No	Yes	Yes
SCCS	2093	First occurrence	182d	Period of dispensing	No	No	Yes	No	Yes
SCCS	2094	First occurrence	182d	Period of dispensing	No	No	Yes	No	Yes
SCCS	2095	First occurrence	182d	Period of dispensing	No	No	Yes	Yes	No
SCCS	2096	First occurrence	182d	Period of dispensing	No	Yes	Yes	Yes	Yes

Design Approach	Analysis_ID	Required observation time prior to outcome	Outcome to include	Exposure windows and exclusion period	Approach	Extrapolation of the aggregated control data	Metrics
CP	3001	182d	All occurrences	30d and 0d	Per-user approach	Raw (no stratification)	CPR
CP	3002	182d	All occurrences	30d and 0d	Per-user approach	Raw (no stratification)	RR predicted
CP	3003	182d	All occurrences	60d and 0d	Per-user approach	Raw (no stratification)	CPR
CP	3004	182d	All occurrences	60d and 0d	Per-user approach	Raw (no stratification)	RR predicted
CP	3005	182d	All occurrences	60d and 15d	Per-user approach	Raw (no stratification)	CPR
CP	3006	182d	All occurrences	60d and 15d	Per-user approach	Raw (no stratification)	RR predicted
CP	3007	182d	All occurrences	60d and 7d	Per-user approach	Raw (no stratification)	CPR
CP	3008	182d	All occurrences	60d and 7d	Per-user approach	Raw (no stratification)	RR predicted
CP	3009	182d	All occurrences	7d and 0d	Per-user approach	Raw (no stratification)	CPR
CP	3010	182d	All occurrences	7d and 0d	Per-user approach	Raw (no stratification)	RR predicted
CP	3011	182d	First occurrence	30d and 0d	Per-user approach	Raw (no stratification)	CPR
CP	3012	182d	First occurrence	30d and 0d	Per-user approach	Raw (no stratification)	RR predicted
CP	3013	182d	First occurrence	60d and 0d	Per-user approach	Raw (no stratification)	CPR
CP	3014	182d	First occurrence	60d and 0d	Per-user approach	Raw (no stratification)	RR predicted
CP	3015	182d	First occurrence	60d and 15d	Per-user approach	Raw (no stratification)	CPR
CP	3016	182d	First occurrence	60d and 15d	Per-user approach	Raw (no stratification)	RR predicted
CP	3017	182d	First occurrence	60d and 7d	Per-user approach	Raw (no stratification)	CPR
CP	3018	182d	First occurrence	60d and 7d	Per-user approach	Raw (no stratification)	RR predicted
CP	3019	182d	First occurrence	7d and 0d	Per-user approach	Raw (no stratification)	CPR
CP	3020	182d	First occurrence	7d and 0d	Per-user approach	Raw (no stratification)	RR predicted
CP	3021	182d	All occurrences	30d and 0d	Per-user approach	Stratified on age and gender	CPR
CP	3022	182d	All occurrences	30d and 0d	Per-user approach	Stratified on age and gender	RR predicted
CP	3023	182d	All occurrences	60d and 0d	Per-user approach	Stratified on age and gender	CPR
CP	3024	182d	All occurrences	60d and 0d	Per-user approach	Stratified on age and gender	RR predicted
CP	3025	182d	All occurrences	60d and 15d	Per-user approach	Stratified on age and gender	CPR
CP	3026	182d	All occurrences	60d and 15d	Per-user approach	Stratified on age and gender	RR predicted
CP	3027	182d	All occurrences	60d and 7d	Per-user approach	Stratified on age and gender	CPR
CP	3028	182d	All occurrences	60d and 7d	Per-user approach	Stratified on age and gender	RR predicted
CP	3029	182d	All occurrences	7d and 0d	Per-user approach	Stratified on age and gender	CPR
CP	3030	182d	All occurrences	7d and 0d	Per-user approach	Stratified on age and gender	RR predicted
CP	3031	182d	First occurrence	30d and 0d	Per-user approach	Stratified on age and gender	CPR
CP	3032	182d	First occurrence	30d and 0d	Per-user approach	Stratified on age and gender	RR predicted
CP	3033	182d	First occurrence	60d and 0d	Per-user approach	Stratified on age and gender	CPR
CP	3034	182d	First occurrence	60d and 0d	Per-user approach	Stratified on age and gender	RR predicted

CP	3035	182d	First occurrence	60d and 15d	Per-user approach	Stratified on age and gender	CPR
CP	3036	182d	First occurrence	60d and 15d	Per-user approach	Stratified on age and gender	RR predicted
CP	3037	182d	First occurrence	60d and 7d	Per-user approach	Stratified on age and gender	CPR
CP	3038	182d	First occurrence	60d and 7d	Per-user approach	Stratified on age and gender	RR predicted
CP	3039	182d	First occurrence	7d and 0d	Per-user approach	Stratified on age and gender	CPR
CP	3040	182d	First occurrence	7d and 0d	Per-user approach	Stratified on age and gender	RR predicted
CP	3041	182d	All occurrences	30d and 0d	Person-time approach	Raw (no stratification)	CPR
CP	3042	182d	All occurrences	30d and 0d	Person-time approach	Raw (no stratification)	RR predicted
CP	3043	182d	All occurrences	60d and 0d	Person-time approach	Raw (no stratification)	CPR
CP	3044	182d	All occurrences	60d and 0d	Person-time approach	Raw (no stratification)	RR predicted
CP	3045	182d	All occurrences	60d and 15d	Person-time approach	Raw (no stratification)	CPR
CP	3046	182d	All occurrences	60d and 15d	Person-time approach	Raw (no stratification)	RR predicted
CP	3047	182d	All occurrences	60d and 7d	Person-time approach	Raw (no stratification)	CPR
CP	3048	182d	All occurrences	60d and 7d	Person-time approach	Raw (no stratification)	RR predicted
CP	3049	182d	All occurrences	7d and 0d	Person-time approach	Raw (no stratification)	CPR
CP	3050	182d	All occurrences	7d and 0d	Person-time approach	Raw (no stratification)	RR predicted
CP	3051	182d	First occurrence	30d and 0d	Person-time approach	Raw (no stratification)	CPR
CP	3052	182d	First occurrence	30d and 0d	Person-time approach	Raw (no stratification)	RR predicted
CP	3053	182d	First occurrence	60d and 0d	Person-time approach	Raw (no stratification)	CPR
CP	3054	182d	First occurrence	60d and 0d	Person-time approach	Raw (no stratification)	RR predicted
CP	3055	182d	First occurrence	60d and 15d	Person-time approach	Raw (no stratification)	CPR
CP	3056	182d	First occurrence	60d and 15d	Person-time approach	Raw (no stratification)	RR predicted
CP	3057	182d	First occurrence	60d and 7d	Person-time approach	Raw (no stratification)	CPR
CP	3058	182d	First occurrence	60d and 7d	Person-time approach	Raw (no stratification)	RR predicted
CP	3059	182d	First occurrence	7d and 0d	Person-time approach	Raw (no stratification)	CPR
CP	3060	182d	First occurrence	7d and 0d	Person-time approach	Raw (no stratification)	RR predicted
CP	3061	182d	All occurrences	30d and 0d	Person-time approach	Stratified on age and gender	CPR
CP	3062	182d	All occurrences	30d and 0d	Person-time approach	Stratified on age and gender	RR predicted
CP	3063	182d	All occurrences	60d and 0d	Person-time approach	Stratified on age and gender	CPR
CP	3064	182d	All occurrences	60d and 0d	Person-time approach	Stratified on age and gender	RR predicted
CP	3065	182d	All occurrences	60d and 15d	Person-time approach	Stratified on age and gender	CPR
CP	3066	182d	All occurrences	60d and 15d	Person-time approach	Stratified on age and gender	RR predicted
CP	3067	182d	All occurrences	60d and 7d	Person-time approach	Stratified on age and gender	CPR
CP	3068	182d	All occurrences	60d and 7d	Person-time approach	Stratified on age and gender	RR predicted
CP	3069	182d	All occurrences	7d and 0d	Person-time approach	Stratified on age and gender	CPR
CP	3070	182d	All occurrences	7d and 0d	Person-time approach	Stratified on age and gender	RR predicted
CP	3071	182d	First occurrence	30d and 0d	Person-time approach	Stratified on age and gender	CPR

CP	3072	182d	First occurrence	30d and 0d	Person-time approach	Stratified on age and gender	RR predicted
CP	3073	182d	First occurrence	60d and 0d	Person-time approach	Stratified on age and gender	CPR
CP	3074	182d	First occurrence	60d and 0d	Person-time approach	Stratified on age and gender	RR predicted
CP	3075	182d	First occurrence	60d and 15d	Person-time approach	Stratified on age and gender	CPR
CP	3076	182d	First occurrence	60d and 7d	Person-time approach	Stratified on age and gender	RR predicted
CP	3077	182d	First occurrence	60d and 7d	Person-time approach	Stratified on age and gender	CPR
CP	3078	182d	First occurrence	60d and 7d	Person-time approach	Stratified on age and gender	RR predicted
CP	3079	182d	First occurrence	7d and 0d	Person-time approach	Stratified on age and gender	CPR
CP	3080	182d	First occurrence	7d and 0d	Person-time approach	Stratified on age and gender	RR predicted

**APPENDIX 4 PERFORMANCES OF OVERALL DESIGN VARIANTS TESTED IN
1/10TH SAMPLED POPULATION**

AlcaponeUgib_Appendix4_v0.1.xlsx

**APPENDIX 5 CONTROL ESTIMATES FROM OPTIMUM DESIGN VARIANTS
TESTED IN THE UNSAMPLED POPULATION**

AlcaponeUgib_Appendix5_v0.1.xlsx

2.7.3. ALI example, submitted article: Empirical assessment of case-based methods for identification of drugs associated with acute liver injury in the French National Healthcare System database (SNDS)

Nicolas H. Thurin, Régis Lassalle, Martijn Schuemie, Marine Pénichon, Joshua J. Gagne, Jeremy A. Rassen, Jacques Benichou, Alain Weill, Patrick Blin, Nicholas Moore, Cécile Droz-Perroteau, Pharmacoepidemiology and Drug safety (1st submission)

Empirical assessment of case-based methods for identification of drugs associated with acute liver injury in the French National Healthcare System database (SNDS)

Short title: Case-based methods for ALI-associated drugs identification in the SNDS

Nicolas H Thurin^{1,2,*}, Régis Lassalle¹, Martijn Schuemie^{3,4}, Marine Pénichon¹, Joshua J Gagne⁵, Jeremy A Rassen⁶, Jacques Benichou^{7,8}, Alain Weill⁹, Patrick Blin¹, Nicholas Moore^{1,2}, Cécile Droz-Perroteau¹

¹Bordeaux PharmacoEpi, INSERM CIC1401, Université de Bordeaux, Bordeaux, France.

²INSERM U1219, Université de Bordeaux, Bordeaux, France.

³Epidemiology Analytics, Janssen Research and Development, Titusville, NJ, USA.

⁴Observational Health Data Sciences and Informatics (OHDSI), New York, NY, USA.

⁵Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA USA.

⁶Aetion, Inc., New York, NY, USA.

⁷Department of Biostatistics and Clinical Research, Rouen University Hospital Rouen, France.

⁸INSERM U1181

⁹Caisse Nationale de l'Assurance Maladie, Paris, France.

* Corresponding authors

nicolas.thurin@u-bordeaux.fr
Bordeaux PharmacoEpi
Université de Bordeaux
case 41 - 146 rue Léo Saignat
33076 Bordeaux cedex - France

Key words:

Acute liver injury, self-controlled case series, case-control, case-population, calibration, claims database

Key points:

- Self-controlled case series yielded the best performance for the identification of drugs associated with acute liver injury in the French National Healthcare database (SNDS) with very small systematic error
- Using a risk window corresponding to the period covered by drug dispensings is important to consider when assessing acute hepatocellular liver injury risk
- Adjustment for multiple drug use helps improve the true negative rate
- Careful definition of inclusion criteria can help reduce bias, especially confounding by indication

Word count excluding abstract, tables, figures and references: 2977

Statement about prior postings and presentations, name(s) of any sponsor(s) of the research contained in the paper, along with grant number(s):

ALCAPONE (Alert generation using the case population approach) methodology and related results were presented during the 33rd and 35th International Conference on Pharmacoepidemiology & Therapeutic Risk Management and at the 1st and 2nd OHDSI European Symposium.

Pharmacoepidemiology and Drug safety is currently reviewing two articles about ALCAPONE methodology (PDS-19-0255) and UGIB results (PDS-19-0409)

The ALCAPONE project is funded by the French Ministry of Health (PREPS 2014, 0635).

ABSTRACT

244/250 words

Objectives: To assess empirically in the French National claims database (SNDS) case-based designs to identify drugs associated with acute liver injury (ALI).

Background: Drug induced ALI is a frequent cause of liver failure. A calibrated method for ALI-associated drugs identification would enable efficient drug safety alert generation from the SNDS.

Methods: All cases of ALI were extracted from SNDS (2009-2014) using specific and sensitive definitions. Positive and negative drug controls were used to compare 196 self-controlled case series (SCCS), case-control (CC), and case-population (CP) design variants, using area under the receiver operating curve (AUC), mean square error (MSE) and coverage probability. Parameters that had major impacts on results were identified through logistic regression.

Results: Using a specific ALI definition, AUCs ranged from 0.78 to 0.94, 0.64 to 0.92 and 0.48 to 0.85, for SCCS, CC and CP, respectively. MSE ranged from 0.12 to 0.40, 0.22 to 0.39 and 1.03 to 5.29, respectively. Variants adjusting for multiple drug use had higher coverage probabilities. Univariate regressions showed that high AUCs were achieved with SCCS using exposed time as the risk window. The top SCCS variant yielded an AUC=0.93 and MSE=0.22 and coverage=86%, with 1/7 negative and 13/18 positive controls presenting significant estimates.

Conclusions: SCCS adjusting for multiple drugs and using exposed time as risk window showed good performances for the identification of ALI-associated drugs in the SNDS. Specific adjustments may be required in the context of particular studies, especially when evaluating the risk related to non-hepatocellular ALI.

1 INTRODUCTION

Acute liver injury (ALI) can range from simple elevation of liver enzymes to acute liver failure leading to liver transplantation or death.^{1, 2} In more than half of cases, severe ALI is drug induced, making medicines the most frequent cause of liver failure in most Western countries.³⁻⁵ Because of its incidence (14 to 19 cases per 100 000 persons),^{6, 7} its challenging management,⁸ and its potentially fatal consequences, drug-induced ALI is a major clinical burden and cause of regulatory action related to medications.⁹⁻¹² Nearly all drug classes can lead to ALI.^{13, 14} Even if some are well known for their proven hepatotoxicity (e.g. antimycobacterial agents, paracetamol, *etc.*), further investigations are needed to explore the potential of other drugs.^{8, 15, 16} As demonstration of the importance of ALI, it was included among the top 10 events studied in the Exploring and Understanding Adverse Drug Reactions by integrative mining of clinical records and biomedical knowledge (EU-ADR) project,^{17, 18} among the health outcomes of interest screened in the Observational Medical Outcomes Partnership (OMOP) experiment,^{19, 20} and was included as a key adverse event investigated by the Pharmacoepidemiological Research on Outcomes of Therapeutics by a European Consortium (PROTECT) project coordinated by the European Medicine Agency.²¹⁻²⁴

The French National Healthcare System database – *Système National des données de Santé* (SNDS) currently includes 66.6 million persons, capturing all reimbursed medical and paramedical encounters, including all dispensed drugs, hospital-discharge summaries and dates of death.²⁵ The database not included in the aforementioned international consortiums but has been useful in the identification of severe ALI and acute liver failure.^{15, 26} ALCAPONE (Alert generation using the case population approach in the French databases), a project funded by the French Health Ministry (PREPS 2014, 0635), was designed to further the common objective of these different collaborations in the SNDS: to evaluate methodological standards in real life settings applicable to drug safety issues. ALCAPONE aimed to leverage the wealth of data in SNDS to empirically assess case-based methods – self-controlled case series (SCCS), case-control (CC) and case-population (CP) – to determine and calibrate the top design for the identification of drugs associated with ALI.

2 METHODS

The overall ALCAPONE methodology has been fully described elsewhere (manuscript under review PDS-19-0255).

Patients with ALI were identified in the SNDS (2009-2014) using ICD10 codes from hospital discharge summaries based on either a specific or a sensitive definition. The specific definition included codes for toxic liver disease with hepatic necrosis or hepatitis or acute hepatitis (ICD10 = K71.1, K71.2, K71.6), and acute or subacute hepatic failure (K72.0). A more sensitive definition also considered patients hospitalized for unspecified toxic liver disease (K71.9). The index date was set to the hospital admission date. Consecutive hospital stays separated by less than 24 hours were aggregated.

To ensure good clinical characterization, the following outcomes were excluded (Appendix 1):

- Outcomes preceded by a hospital stay with a diagnosis of liver transplant status within 182 days
- Outcomes presenting concomitantly with, or preceded within 182 days by, any diagnosis code for non-drug-related liver disease or potential liver injury
- Outcomes presenting concomitantly with a diagnosis of poisoning, acute viral hepatitis, traumatic injury of liver, or pregnancy-related liver disorder

To ensure the completeness of data, and of drug exposure, in particular, outcomes presenting at with least one of the following were also excluded:

- Less than 182 days of observation prior to the index date
- Hospital stays (whatever the diagnosis) ending within 30 days prior to or starting more than 7 days before and ending on the index date

Drugs of interest were restricted to those with enough power to be detected in the population (minimum detectable relative risk ≤ 1.30).²⁷ Positive controls were drugs with a known association with ALI. Negative controls were those with no known association with ALI. The full ALCAPONE reference set for ALI is available in Appendix 2. Three case-based approaches were considered: SCCS which consists of comparing each case to itself, CC which compares the exposure distribution prior to outcomes in

cases with the distribution in individual patients at risk for the outcome, and CP which compares the exposure distribution in cases with the distribution in the overall population. Different settings, such as adjustment strategies, risk window lengths, *etc.*, were applied to these three case-based approaches forming different design variants. These variants were applied in two steps to generate point estimates (relative incidence for SCCS, odds ratio for CC and case population ratio or predicted relative risk for CP) between ALI and each drug control, as described below. A total of 96 SCCS, 20 CC and 80 CP variants were tested. The exact settings of each design variant are described in Appendix 3.

All design variants were run in both the specific and the sensitive case population to identify the best-performing variant based on area under the receiver operating characteristics curve (AUC), mean square error (MSE) and coverage probability. MSE and coverage probability were estimated for negative controls only, assuming no association (*i.e.*, true point estimate of 1 for the measure of relative association). The AUC (ranging from 0 to 1) measures discrimination between positive and negative controls, with higher values indicating better discrimination. The MSE quantifies the accuracy of an estimator, combining both bias and random error, with smaller values indicating better accuracy. The coverage probability is the proportion of the 95% confidence interval estimates that include the true parameter value, 1 in our case. Values as close as possible to nominal 95% coverage are desirable.

We also conducted an univariable logistic regression analysis to identify for the most influential parameters that best discriminated the performance of the design variants. The dependent variable was the probability that a variant had an AUC above the 70th percentile of the AUC distribution, here 0.90. The independent covariates included the parameters that were varied in the design variant considered. For example, for the SCCS, the independent covariates were: multiple drug adjustment: yes / no; pre-exposure length: 0, 7 or 30 days, *etc.*

Using the estimates from the negative controls, for which a theoretical value of 1 was expected, we observed how often $p < 0.05$ while the null hypothesis was true, and we fitted a distribution to the effect estimates, modeling the distribution of the residual bias under the null.²⁸⁻³⁰ Estimated parameters of this “empirical null distribution” were then used to compute “calibrated” p -values, taking into account random and systematic error (*i.e.* the background noise) inherent to the application of a design variant to the

SNDS. Bayesian 95% credible intervals (CI) were then computed using Markov Chain Monte Carlo.

3 RESULTS

3.1 POPULATION

The selection process of ALI cases according to specific and sensitive definitions is presented in Figure 1. Over 6 years, 5 225 ALI episodes among 5 152 patients were included according to the specific ALI definition; 20 patients presented more than one outcome. The sensitive definition added 368 outcomes from 363 patients for a total of 5 580 outcomes. Median age was 54 and 52 for the specific and sensitive definitions, respectively; 60% were male. From the 81 drugs of interest screened in the unsampled population, 25 and 27 presented a minimum detectable relative risk ≤ 1.30 and were deemed detectable according to the specific and sensitive definition, respectively. The specific definition enabled the detection of 18 positive drugs controls out of 58 and 7 negative controls out of 23. Sensitive definition enabled the detection of 20 positive controls out of 58 and 7 negative controls out of 23 (see Appendix 2).

3.2 BEST PERFORMING DESIGN VARIANTS AND MAJOR IMPACT PARAMETERS

SCCS globally showed better discrimination (Figure 2) and MSE (Figure 3) than CC and CP. This difference was stronger with the sensitive definition than the specific. For the specific case definition, AUCs ranged from 0.78 to 0.94 for SCCS, from 0.64 to 0.92 for CC and from 0.48 to 0.85 for CP. MSE ranged from 0.12 to 0.40 for SCCS, from 0.22 to 0.39 for CC and from 1.03 to 5.29 for CP. For the sensitive definition, almost no changes were observed in AUC of SCCS and CP whereas the top-CC AUC was slightly inferior (0.89). The MSE distribution followed the same trend as with the specific definition, with lower values for SCCS (0.17-0.46) and CC (0.20-0.54) than for CP (1.08-4.59). The performances of all the tested design variants are available in Appendix 4. Case-population variants using cumulated person-time exposure (person-time approach) showed better performance than those using the actual number of persons exposed (per user approach) ($AUC \geq 0.76$ and $MSE \leq 1.9$ vs. $AUC < 0.53$ and $MSE > 2.3$). No such clear trend appeared in CC performance even though variants that used an exclusion period prior to the outcome onset obtained lower AUC.

For the specific definition, the design variants with the highest AUC (0.94) and smallest MSE (0.22) were the SCCS 2089, with a coverage probability of 57% (Table 1). SCCS variant 2089 only considered the first occurrence of the outcome and used the period covered by drug dispensing as the risk window without using any adjustment methods (see Appendix 3). SCCS 2090 and 2092 used the same parameters of SCCS 2089 but also adjusted for multiple drug use (SCCS 2090) and multiple drug use and seasonality (SCCS 2092), resulting in a slightly lower AUC (0.93) but in substantially higher coverage probability (86%).

SCCS 2090 and 2092 also performed well with the sensitive definition with AUC = 0.92, MSE = 0.19 and coverage probability = 86%. However, a slightly better AUC (0.93) was obtained with SCCS 2046, 2048, 2094 and 2096 with comparable values for MSE and coverage (Table 2).

Univariable logistic regression analyses showed that, for both the specific and the sensitive definitions, the strongest significant determinant of a high AUC was the utilization of a risk window corresponding to the period covered by drug dispensing and not a fixed 30-day window (Table 3 and Table 4). In the overall results (Appendix 4), higher coverage probabilities were obtained adjusting for multiple drug use.

As adjusting for seasonality appeared to have no impact on SCCS discriminative ability, we determined SCCS 2090 to be the top-performing variant owing to its good overall performances with both the specific and the sensitive definitions.

3.3 CALIBRATION OF THE OPTIMUM DESIGN VARIANT

With the specific definition, out of 7 negative controls, 6 were not significantly associated with ALI (Figure 4) according to SCCS 2090. Only sitagliptin had a confidence interval excluding 1 (1.03 - 2.58) with a relative incidence of 1.64. Of the 18 positive controls 13 were significantly associated with ALI. Diltiazem, terbinafine, erythromycin, piroxicam and naproxen were not (Figure 5).

The empirical null distribution with mean 0.15, CI = [-0.15 – 0.38] and precision ($=1/\hat{\sigma}^2$) of 198, CI = [3 – 915] was derived based on the negative control estimates and used to compute calibrated *p*-values. Using the calibrated *p*-values, sitagliptin was no longer significant but inhaled fluticasone was (Table 5), and two positive controls, fluconazole and celecoxib, moved from significant to non-significant (Table 6). Figure 6 illustrates this calibration process. The gray zone (below the dashed line) represents

the area where the conventional p -value is smaller than 0.05. The orange zone shows the area where calibrated p -value is under 0.05. The red band represents the uncertainty in the calibration process. We can see that the only negative control (blue dot) significantly associated with ALI according to the conventional p -value is below the dashed line but out of the orange area, meaning that once the p -value is calibrated it is no longer considered as significant.

Using the sensitive definition, estimates generated for the drug controls were consistent with those observed in the specific definition. Two extra positive controls, rifampicin and etodolac, showed a significant association and no association with ALI, respectively. The empirical null distribution was derived (mean = 0.16, CI = [-0.12 – 0.40]; precision = 198 CI = [3 – 957]) and used to compute calibrated p -values (Table 6). The same trends as those observed for the specific definition appeared in the calibration process for negative controls. However more positive controls moved from significant to not significant: erythromycin, fluconazole, celecoxib, valproic acid.

All estimates generated using the different design variants are available in Appendix 5.

4 DISCUSSION

Overall, the self-controlled case series showed better results than CC and CP in terms of discrimination and accuracy of point estimates in this large-scale assessment of ALI in the SNDS. Using both specific and sensitive definitions for ALI, the utilization of the period covered by drug dispensings as risk window seemed to be the key parameter in the generation of a high AUC, which could be explained by the nature of the selected positive controls and the inclusion criteria. Drug-induced liver injury is typically classified as either direct or idiosyncratic.¹⁶ Most of the positive drug controls included in ALCAPONE belong to this second category. Idiosyncratic hepatotoxicity appears in a variable timeframe from days to weeks, sometimes months, from initial exposure.^{31, 32} This difference in latency is related to liver injury phenotypes: first symptoms of cholestatic hepatitis do not appear for weeks, whereas those resulting from acute hepatocellular hepatitis, the most common and serious manifestation of idiosyncratic liver injury, may occur within a few days.^{6, 32, 33} The majority of the positive controls included in our reference set can lead either to hepatocellular or cholestatic hepatitis. However, since ALCAPONE focused on serious events, toxic liver disease with cholestasis (ICD = K71.0) was not included. Thus, most of the ALIs analyzed in this study probably resulted from hepatocellular toxicity. The better performance observed in SCCS using the actual period covered by drug dispensings as the risk window compared to those considering a fixed period of 30 days, even in non-chronic drugs, suggests that such events are more likely to occur during the exposure period and that non-exposed time may be less at risk. Similar conclusions were observed in the PROTECT project.^{22, 23} However, it is important to stress that restricting the risk window to the period of treatment does not allow assessment of the risk related to more delayed hepatocellular ALI for non-chronic drugs. This may explain the non-significant estimates of some positive controls, such as piroxicam or naproxen for which hepatotoxic mechanism remains unclear.^{34, 35}

Although no direct impact was observed on the AUC, perhaps because of the low number of negative controls included, multiple drug adjustment appeared to have a beneficial effect on the coverage probability, improving the true negative rate. This approach has already demonstrated its greater ability to discriminate between positive and negative drug-outcome pairs, compared to unadjusted SCCS.³⁶

In this study, AUC values obtained through the top SCCS variants ($AUC > 0.92$) were higher than those observed in the OMOP experiment ($AUC \leq 0.70$).³⁷ These discrepancies may result from the difference in our respective drug reference set. It is also possible that the care taken to carefully define the ALI population, ruling out potential alcohol or viral-related outcomes as well as non-acute hepatitis, and excluding patients for whom it was impossible to assess previous drug exposures, may have contributed to the improved performance. This process may have reduced the rate of misclassification, increasing the accuracy of the results,³⁸ or reduced the susceptibility to confounding.^{39, 40} For example, we might have observed a significant association between lactulose and ALI, had alcohol-related liver injury and cirrhosis not been excluded: lactulose is most likely given to patients at risk for hepatic encephalopathy, which most of the time results from cirrhosis.^{41, 42} The same rationale could explain the good performance observed across the best-performing CC variants ($AUC \geq 0.89$), which had higher AUCs compared to those from the original OMOP experiment ($AUC \leq 0.60$ for 4 out to 5 databases tested).

When considering CP, analyses 3041 to 3080 using person-time approach, yielded better AUC (> 0.75) compared to analyses 3041-3080 using per user approach (< 0.55). Such results diverge from previous conclusions.⁴³ Since most of the controls were supposed to lead to idiosyncratic ALI without dose-related effect, we were expecting that an approach that does not take into account cumulated-exposure, such as the per-user approach, would generate the best results. However, given the low number of negative controls, small discrepancies in their classification may have an important impact in overall variant performances measurements. Further investigation with more drugs controls, and an advanced classification regarding their exposure typology and expected related effect would be necessary to better understand their association with ALIs.

The calibration process, including the derivation of the empirical null distribution, tended to show that systematic error inherent to the application of SCCS to investigate ALI in the SNDS was small. Even if the uncertainty resulting from the restricted number of subjects and negative controls is large, credible intervals of calibrated p -values usually included the traditional p -values. Positive control estimates, especially for antibiotics, were of the same order as those generated by SCCS in PROTECT project.²³ As described above, false negatives may result from the non-detection of

delayed effect. The absence of association for erythromycin could also be the consequence of lack of power, which at the national level would mean that this drug would not represent an important public health safety concern. No explanation was found for the significant estimate generated for ALI and sitagliptin. This association could either result from random error or could turn out to be a potential safety issue, since two cases were published in 2018.⁴⁴

The SCCS had the best performance for the identification of ALI-associated drugs in SNDS. Adjusting for multiple drugs and considering the actual period covered by drug dispensings appeared to be important features of this design. The careful selection of the study population seemed to limit residual bias. The calibration process showed that minimal systematic error was generated by the optimum SCCS in the SNDS when applied to ALI. However, not all possible SCCS implementations have been assessed and this conclusion mainly apply to idiosyncratic hepatocellular ALI. Specific adjustments may be required in the context of particular studies, especially when evaluating the risk related to other liver injury phenotypes.

ACKNOWLEDGMENTS

The ALCAPONE project was funded by the French Ministry of Health (PREPS 2014, 0635).

The authors wish to thank the SNDS database management team, and ADERA, the non-profit organization that provides managerial and human resource support to Bordeaux PharmacoEpi.

REFERENCES

1. Zimmerman HJ. The spectrum of hepatotoxicity. *Perspect Biol Med* 1968, 12:135-161.
2. Zimmerman HJ. Drug-induced liver disease. *Drugs* 1978, 16:25-45.
3. Hadem J, Tacke F, Bruns T, Langgartner J, Strnad P, Denk GU, Fikatas P, Manns MP, Hofmann WP, Gerken G, et al. Etiologies and Outcomes of Acute Liver Failure in Germany. *Clinical Gastroenterology and Hepatology* 2012, 10:664-669.e662.
4. Ostapowicz G, Fontana RJ, Schiødt FV, Larson A, Davern TJ, Han SHB, McCashland TM, Shakil AO, Hay JE, Hynan L, et al. Results of a Prospective Study of Acute Liver Failure at 17 Tertiary Care Centers in the United States. *Annals of Internal Medicine* 2002, 137:947-954.
5. Wei G, Bergquist A, Broome U, Lindgren S, Wallerstedt S, Almer S, Sangfelt P, Danielsson A, Sandberg-Gertzen H, Loof L, et al. Acute liver failure in Sweden: etiology and outcome. *J Intern Med* 2007, 262:393-401.
6. Bjornsson ES, Bergmann OM, Bjornsson HK, Kvaran RB, Olafsson S. Incidence, presentation, and outcomes in patients with drug-induced liver injury in the general population of Iceland. *Gastroenterology* 2013, 144:1419-1425, 1425.e1411-1413; quiz e1419-1420.
7. Sgro C, Clinard F, Ouazir K, Chanay H, Allard C, Guilleminet C, Lenoir C, Lemoine A, Hillon P. Incidence of drug-induced hepatic injuries: a French population-based study. *Hepatology* 2002, 36:451-455.
8. Navarro VJ, Senior JR. Drug-Related Hepatotoxicity. *New England Journal of Medicine* 2006, 354:731-739.
9. Chen M, Zhang J, Wang Y, Liu Z, Kelly R, Zhou G, Fang H, Borlak J, Tong W. The Liver Toxicity Knowledge Base: A Systems Approach to a Complex End Point. *Clinical Pharmacology & Therapeutics* 2013, 93:409-412.
10. Gulmez SE, Larrey D, Pageaux GP, Bernuau J, Bissoli F, Horsmans Y, Thorburn D, McCormick PA, Stricker B, Toussi M, et al. Liver transplant associated with paracetamol overdose: results from the seven-country SALT study. *Br J Clin Pharmacol* 2015, 80:599-606.
11. Gulmez SE, Larrey D, Pageaux GP, Lignot-Maleyran S, de Vries C, Sturkenboom M, Perez-Gutthann S, Benichou J, Bissoli F, Horsmans Y, 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, 69:605-616.
12. Watkins P. Drug Safety Sciences and the Bottleneck in Drug Development. *Clinical Pharmacology & Therapeutics* 2011, 89:788-790.
13. David S, Hamilton JP. Drug-induced Liver Injury. *US gastroenterology & hepatology review* 2010, 6:73-80.
14. Reuben A, Koch DG, Lee WM. Drug-induced acute liver failure: results of a U.S. multicenter, prospective study. *Hepatology* 2010, 52:2065-2076.

15. Moore N, Duret S, Grolleau A, Lassalle R, Barbet V, Duong M, Thurin N, Droz-Perroteau C, Gulmez SE. Previous Drug Exposure in Patients Hospitalised for Acute Liver Injury: A Case-Population Study in the French National Healthcare Data System. *Drug Saf* 2018.
16. Zimmerman HJ. *Hepatotoxicity : the adverse effects of drugs and other chemicals on the liver*. 2nd ed. Philadelphia: Lippincott Williams & Wilkins; 1999.
17. Coloma PM, Avillach P, Salvo F, Schuemie MJ, Ferrajolo C, Pariente A, Fourrier-Réglat A, Molokhia M, Patadia V, van der Lei J, et al. A Reference Standard for Evaluation of Methods for Drug Safety Signal Detection Using Electronic Healthcare Record Databases. *Drug Safety* 2013, 36:13-23.
18. Trifiro G, Pariente A, Coloma PM, Kors JA, Polimeni G, Miremont-Salame G, Catania MA, Salvo F, David A, Moore N, et al. Data mining on electronic health record databases for signal detection in pharmacovigilance: which events to monitor? *Pharmacoepidemiol Drug Saf* 2009, 18:1176-1184.
19. Ryan PB, Madigan D, Stang PE, Overhage JM, Racoosin JA, Hartzema AG. Empirical assessment of methods for risk identification in healthcare data: results from the experiments of the Observational Medical Outcomes Partnership. *Statistics in medicine* 2012, 31:4401-4415.
20. Ryan PB, Schuemie MJ, Welebob E, Duke J, Valentine S, Hartzema AG. Defining a reference set to support methodological research in drug safety. *Drug Saf* 2013, 36 Suppl 1:S33-47.
21. Abbing-Karahagopian V, Kurz X, de Vries F, van Staa TP, Alvarez Y, Hesse U, Hasford J, Dijk Lv, de Abajo FJ, Weil JG, et al. Bridging differences in outcomes of pharmacoepidemiological studies: design and first results of the PROTECT project. *Current clinical pharmacology* 2014, 9:130-138.
22. Brauer R, Douglas I, Garcia Rodriguez LA, Downey G, Huerta C, de Abajo F, Bate A, Feudjo Tepie M, de Groot MCH, Schlienger R, et al. Risk of acute liver injury associated with use of antibiotics. Comparative cohort and nested case-control studies using two primary care databases in Europe. *Pharmacoepidemiology and Drug Safety* 2016, 25:29-38.
23. Brauer R, Ruigómez A, Klungel O, Reynolds R, Feudjo Tepie M, Smeeth L, Douglas I. The risk of acute liver injury among users of antibiotic medications: a comparison of case-only studies. *Pharmacoepidemiology and drug safety* 2016, 25 Suppl 1:39-46.
24. Udo R, Tcherny-Lessenot S, Brauer R, Dolin P, Irvine D, Wang Y, Auclert L, Juhaeri J, Kurz X, Abenheim L, et al. The risk of acute liver injury associated with the use of antibiotics—evaluating robustness of results in the pharmacoepidemiological research on outcomes of therapeutics by a European consortium (PROTECT) project. *Pharmacoepidemiology and Drug Safety* 2016, 25:47-55.
25. Bezin J, Duong M, Lassalle R, Droz C, Pariente A, Blin P, Moore N. The national healthcare system claims databases in France, SNIIRAM and EGB: Powerful tools for pharmacoepidemiology. *Pharmacoepidemiol Drug Saf* 2017, 26:954-962.

26. Gulmez SE, Unal US, Lassalle R, Chartier A, Grolleau A, Moore N. Risk of hospital admission for liver injury in users of NSAIDs and nonoverdose paracetamol: Preliminary results from the EPIHAM study. *Pharmacoepidemiol Drug Saf* 2018, 27:1174-1181.
27. Armstrong B. A simple estimator of minimum detectable relative risk, sample size, or power in cohort studies. *Am J Epidemiol* 1987, 126:356-358.
28. Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. Robust empirical calibration of p-values using observational data. *Stat Med* 2016, 35:3883-3888.
29. Schuemie MJ, Ryan PB, DuMouchel W, Suchard MA, Madigan D. Interpreting observational studies: why empirical calibration is needed to correct p-values. *Stat Med* 2014, 33:209-218.
30. Gruber S, Tchetgen Tchetgen E. Limitations of empirical calibration of p-values using observational data. *Statistics in Medicine* 2016, 35:3869-3882.
31. Andrade RJ, Aithal GP, Björnsson ES, Kaplowitz N, Kullak-Ublick GA, Larrey D, Karlsen TH. EASL Clinical Practice Guidelines: Drug-induced liver injury. *Journal of Hepatology* 2019, 70:1222-1261.
32. Hoofnagle JH, Björnsson ES. Drug-Induced Liver Injury — Types and Phenotypes. *New England Journal of Medicine* 2019, 381:264-273.
33. Chalasani N, Bonkovsky HL, Fontana R, Lee W, Stolz A, Talwalkar J, Reddy KR, Watkins PB, Navarro V, Barnhart H, et al. Features and Outcomes of 899 Patients With Drug-Induced Liver Injury: The DILIN Prospective Study. *Gastroenterology* 2015, 148:1340-1352 e1347.
34. Bessone F. Non-steroidal anti-inflammatory drugs: What is the actual risk of liver damage? *World journal of gastroenterology* 2010, 16:5651-5661.
35. Schmeltzer PA, Kosinski AS, Kleiner DE, Hoofnagle JH, Stolz A, Fontana RJ, Russo MW, Drug-Induced Liver Injury N. Liver injury from nonsteroidal anti-inflammatory drugs in the United States. *Liver international : official journal of the International Association for the Study of the Liver* 2016, 36:603-609.
36. Simpson SE, Madigan D, Zorych I, Schuemie MJ, Ryan PB, Suchard MA. Multiple self-controlled case series for large-scale longitudinal observational databases. *Biometrics* 2013, 69:893-902.
37. 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 safety : an international journal of medical toxicology and drug experience* 2013, 36 Suppl 1:S83-93.
38. Copeland KT, Checkoway H, McMichael AJ, Holbrook RH. Bias due to misclassification in the estimation of relative risk. *Am J Epidemiol* 1977, 105:488-495.
39. Secret MH, Platt RW, Dormuth CR, Chateau D, Targownik L, Nie R, Doyle CM, Dell'Aniello S, Fillion KB. Extreme restriction design as a method for reducing confounding by indication in pharmacoepidemiologic research. *Pharmacoepidemiol Drug Saf* 2019.

40. Timmer A, de Sordi D, Kappen S, Kohse KP, Schink T, Perez-Gutthann S, Jacquot E, Deltour N, Pladevall M. Validity of hospital ICD-10-GM codes to identify acute liver injury in Germany. *Pharmacoepidemiol Drug Saf* 2019.
41. Al Sibae MR, McGuire BM. Current trends in the treatment of hepatic encephalopathy. *Ther Clin Risk Manag* 2009, 5:617-626.
42. Grønbaek L, Watson H, Vilstrup H, Jepsen P. Benzodiazepines and risk for hepatic encephalopathy in patients with cirrhosis and ascites. *United European gastroenterology journal* 2018, 6:407-412.
43. Moore N, Gulmez SE, Larrey D, Pageaux GP, Lignot S, Lassalle R, Jove J, Pariente A, Blin P, Benichou J, 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. *Pharmacoepidemiology and drug safety* 2013, 22:160-167.
44. Shahbaz A, Aziz K, Umair M, Sharifzadeh M, Sachmechi I. Acute Liver Injury Induced by Sitagliptin: Report of Two Cases and Review of Literature. *Cureus* 2018, 10:e2776-e2776.

TABLES

Table 1. Performances of the 10 most discriminant self-controlled case series variants for acute liver injury specific definition with the corresponding settings

Analysis ID	Sensitivity	Specificity	AUC	MSE	Coverage probability	Variant settings
2089	0.889	0.571	0.937	0.221	57%	Occurrence considered: first, Risk window: period of dispensation
2045	0.889	0.571	0.929	0.215	57%	Occurrence considered: all, Risk window: period of dispensation, Adjusted on age
2047	0.889	0.571	0.929	0.215	57%	Occurrence considered: all, Risk window: period of dispensation, Adjusted on age and seasonality
2041	0.889	0.571	0.929	0.216	57%	Occurrence considered: all, Risk window: period of dispensation
2043	0.889	0.571	0.929	0.216	57%	Occurrence considered: all, Risk window: period of dispensation, Adjusted on seasonality
2095	0.889	0.571	0.929	0.221	57%	Occurrence considered: first, Risk window: period of dispensation, Adjusted on age and seasonality
2093	0.889	0.571	0.929	0.221	57%	Occurrence considered: first, Risk window: period of dispensation, Adjusted on age
2091	0.889	0.571	0.929	0.221	57%	Occurrence considered: first, Risk window: period of dispensation, Adjusted on seasonality
2092	0.722	0.857	0.929	0.222	86%	Occurrence considered: first, Risk window: period of dispensation, Adjusted on seasonality and multiple drugs
2090	0.722	0.857	0.929	0.224	86%	Occurrence considered: first, Risk window: period of dispensation, Adjusted on multiple drugs

A drug control was considered as positive for a given variant, when the left bound of its confidence interval was >1

AUC = area under the receiver operating characteristics curve; MSE = mean square error

Table 2. Performances of the 10 most discriminant case series variants for acute liver injury sensitive definition with the corresponding settings

Analysis ID	Sensitivity	Specificity	AUC	MSE	Coverage probability	Variant settings
2048	0.750	0.857	0.929	0.185	86%	Occurrence considered: all, Risk window: period of dispensation, Adjusted on age, seasonality and multiple drugs
2046	0.750	0.857	0.929	0.186	86%	Occurrence considered: all, Risk window: period of dispensation, Adjusted on age and multiple drugs
2096	0.750	0.857	0.929	0.188	86%	Occurrence considered: first, Risk window: period of dispensation, Adjusted on age, seasonality and multiple drugs
2094	0.750	0.857	0.929	0.189	86%	Occurrence considered: first, Risk window: period of dispensation, Adjusted on age and multiple drugs
2044	0.750	0.857	0.921	0.191	86%	Occurrence considered: all, Risk window: period of dispensation, Adjusted on seasonality and multiple drugs
2042	0.750	0.857	0.921	0.192	86%	Occurrence considered: all, Risk window: period of dispensation, Adjusted on multiple drugs
2092	0.750	0.857	0.921	0.194	86%	Occurrence considered: first, Risk window: period of dispensation, Adjusted on seasonality and multiple drugs
2090	0.750	0.857	0.921	0.195	86%	Occurrence considered: first, Risk window: period of dispensation, Adjusted on multiple drugs
2040	0.750	0.857	0.914	0.197	86%	Occurrence considered: all, 7-day pre-exposure window, Risk window: period of dispensation, Adjusted on age, seasonality and multiple drugs
2088	0.750	0.857	0.914	0.199	86%	Occurrence considered: first, 7-day pre-exposure window, Risk window: period of dispensation, Adjusted on age, seasonality and multiple drugs

A drug control was considered as positive for a given variant, when the left bound of its confidence interval was >1

AUC = area under the receiver operating characteristics curve; MSE = mean square error

Table 3. Univariate logistic regression analysis of self-controlled case series parameters influencing on the discriminating power, acute liver injury specific definition

	Variants with low AUC n=51	Variants with high AUC n=45	High vs. Low AUC OR [95% Confidence Interval]	p	AUC of the univariate model
Age				0.841	0.51
No	26 (51.0)	22 (48.9)	1		
Yes	25 (49.0)	23 (51.1)	1.09 [0.49 - 2.41]		
Seasonality				0.841	0.51
No	26 (51.0)	22 (48.9)	1		
Yes	25 (49.0)	23 (51.1)	1.09 [0.49 - 2.41]		
Outcome				0.841	0.51
All occurrences	26 (51.0)	22 (48.9)	1		
First occurrence	25 (49.0)	23 (51.1)	1.09 [0.49 - 2.41]		
Multiple drugs				0.548	0.53
No	24 (47.1)	24 (53.3)	1		
Yes	27 (52.9)	21 (46.7)	0.78 [0.35 - 1.73]		
Pre-Exposure Window				0.703	0.54
No	16 (31.4)	16 (35.6)	1		
7 days	16 (31.4)	16 (35.6)	1.00 [0.38 - 2.64]		
30 days	19 (37.3)	13 (28.9)	0.69 [0.26 - 1.83]		
Risk window				<0.001	0.97
Period of dispensing	3 (5.9)	45 (100.0)	1		
30 days from dispensing first day	48 (94.1)	0 (0.0)	0.00 [0.00 - 0.01]		

AUC = area under the receiver operating characteristics curve; A high AUC was defined as an AUC \geq 0.90

Table 4. Univariate logistic regression analysis of self-controlled case series parameters influencing on the discriminating power, acute liver injury sensitive definition

	Variants with low AUC n=56	Variants with high AUC n=40	High vs. Low AUC OR [95% Confidence Interval]	p	AUC of the univariate model
Age				0.418	0.54
No	30 (53.6)	18 (45.0)	1		
Yes	26 (46.4)	22 (55.0)	1.40 [0.62 - 3.16]		
Seasonality				1.000	0.5
No	28 (50.0)	20 (50.0)	1		
Yes	28 (50.0)	20 (50.0)	1.00 [0.45 - 2.24]		
Outcome				0.418	0.54
All occurrences	26 (46.4)	22 (55.0)	1		
First occurrence	30 (53.6)	18 (45.0)	0.71 [0.32 - 1.60]		
Multiple drugs				0.418	0.54
No	30 (53.6)	18 (45.0)	1		
Yes	26 (46.4)	22 (55.0)	1.40 [0.63 - 3.16]		
Pre-Exposure Window				0.084	0.61
No	16 (28.6)	16 (40.0)	1		
7 days	16 (28.6)	16 (40.0)	1.00 [0.38 - 2.64]		
30 days	24 (42.9)	8 (20.0)	0.35 [0.12 - 0.96]		
Risk window				<0.001	0.93
Period of dispensing	8 (14.3)	40 (100.0)	1		
30 days from dispensing first day	48 (85.7)	0 (0.0)	0.00 [0.00 - 0.02]		

AUC = area under the receiver operating characteristics curve; A high AUC was defined as an AUC \geq 0.90

Table 5. Traditional and calibrated *p*-value for acute liver injury (ALI) negative controls, SCCS variant 2090

ATC	INN	ALI specific definition		ALI sensitive definition	
		Traditional <i>p</i> -value	Calibrated <i>p</i> -value [95% CI]	Traditional <i>p</i> -value	Calibrated <i>p</i> -value [95% CI]
C01DA02	NITROGLYCERIN	0.171	0.632 [0.170 - 0.982]	0.203	0.711 [0.242 - 0.988]
A06AD11	LACTULOSE	0.197	0.693 [0.190 - 0.985]	0.157	0.711 [0.204 - 0.988]
A10BH01	SITAGLIPTIN	0.033	0.216 [0.040 - 0.722]	0.016	0.171 [0.027 - 0.680]
G04BD04	OXYBUTYNIN	0.672	0.405 [0.136 - 0.930]	0.760	0.441 [0.153 - 0.924]
R03AC13	FORMOTEROL	0.840	0.759 [0.375 - 0.986]	0.660	0.811 [0.417 - 0.993]
R03BA05	FLUTICASONE	0.082	0.048 [0.016 - 0.309]	0.056	0.031 [0.010 - 0.221]
R01AD08	FLUTICASONE	0.580	0.493 [0.380 - 0.700]	0.989	0.835 [0.630 - 0.988]

ATC = Anatomical Therapeutic and Chemical classification; CI = Credible interval; INN = International nonproprietary name

Table 6. Traditional and calibrated *p*-value for acute liver injury (ALI) positive controls, SCCS variant 2090

ATC	INN	ALI specific definition		ALI sensitive definition	
		Traditional <i>p</i> -value	Calibrated <i>p</i> -value [95% CI]	Traditional <i>p</i> -value	Calibrated <i>p</i> -value [95% CI]
C08DB01	DILTIAZEM	0.004	0.028 [0.005 - 0.200]	0.002	0.021 [0.003 - 0.236]
C09AA05	RAMIPRIL	0.859	0.289 [0.016 - 0.949]	0.985	0.401 [0.032 - 0.961]
D01BA02	TERBINAFINE	0.173	0.437 [0.153 - 0.882]	0.139	0.428 [0.153 - 0.903]
J01CR02	AMOXICILLIN AND β - LACTAMASE INHIBITOR	<0.001	<0.001 [<0.021]	<0.001	<0.001 [<0.062]
J01FA01	ERYTHROMYCIN	0.108	0.192 [0.096 - 0.356]	0.020	0.056 [0.022 - 0.188]
J01MA01	OFLOXACIN	<0.001	<0.001 [<0.006]	<0.001	<0.001 [<0.031]
J01MA02	CIPROFLOXACIN	<0.001	<0.001 [<0.001]	<0.001	<0.001 [<0.011]
J01MA06	NORFLOXACIN	<0.001	<0.001 [<0.060]	<0.001	<0.001 [<0.121]
J01MA12	LEVOFLOXACIN	<0.001	<0.001 [<0.009]	<0.001	<0.001 [<0.031]
J01XE01	NITROFURANTOIN	<0.001	<0.001 [<0.024]	<0.001	<0.001 [<0.100]
J02AC01	FLUCONAZOLE	0.008	0.055 [0.008 - 0.317]	0.005	0.065 [0.009 - 0.439]
J04AB02	RIFAMPICIN			<0.001	0.008 [0.001 - 0.163]
M01AB08	ETODOLAC			0.334	0.523 [0.310 - 0.813]
M01AC01	PIROXICAM	0.118	0.462 [0.113 - 0.954]	0.116	0.536 [0.152 - 0.961]
M01AE01	IBUPROFEN	<0.001	<0.001 [<0.030]	<0.001	<0.001 [<0.089]
M01AE02	NAPROXEN	0.613	0.741 [0.287 - 0.989]	0.596	0.743 [0.272 - 0.987]
M01AH01	CELECOXIB	0.013	0.123 [0.014 - 0.583]	0.007	0.127 [0.014 - 0.642]
M04AA01	ALLOPURINOL	<0.001	0.008 [<0.311]	<0.001	0.015 [<0.434]
N03AF01	CARBAMAZEPINE	<0.001	<0.001 [<0.00]	<0.001	<0.001 [<0.004]
N03AG01	VALPROIC ACID	0.003	0.046 [0.004 - 0.396]	0.006	0.115 [0.012 - 0.628]

ATC = Anatomical Therapeutic and Chemical classification; CI = Credible interval; INN = International nonproprietary name

FIGURES

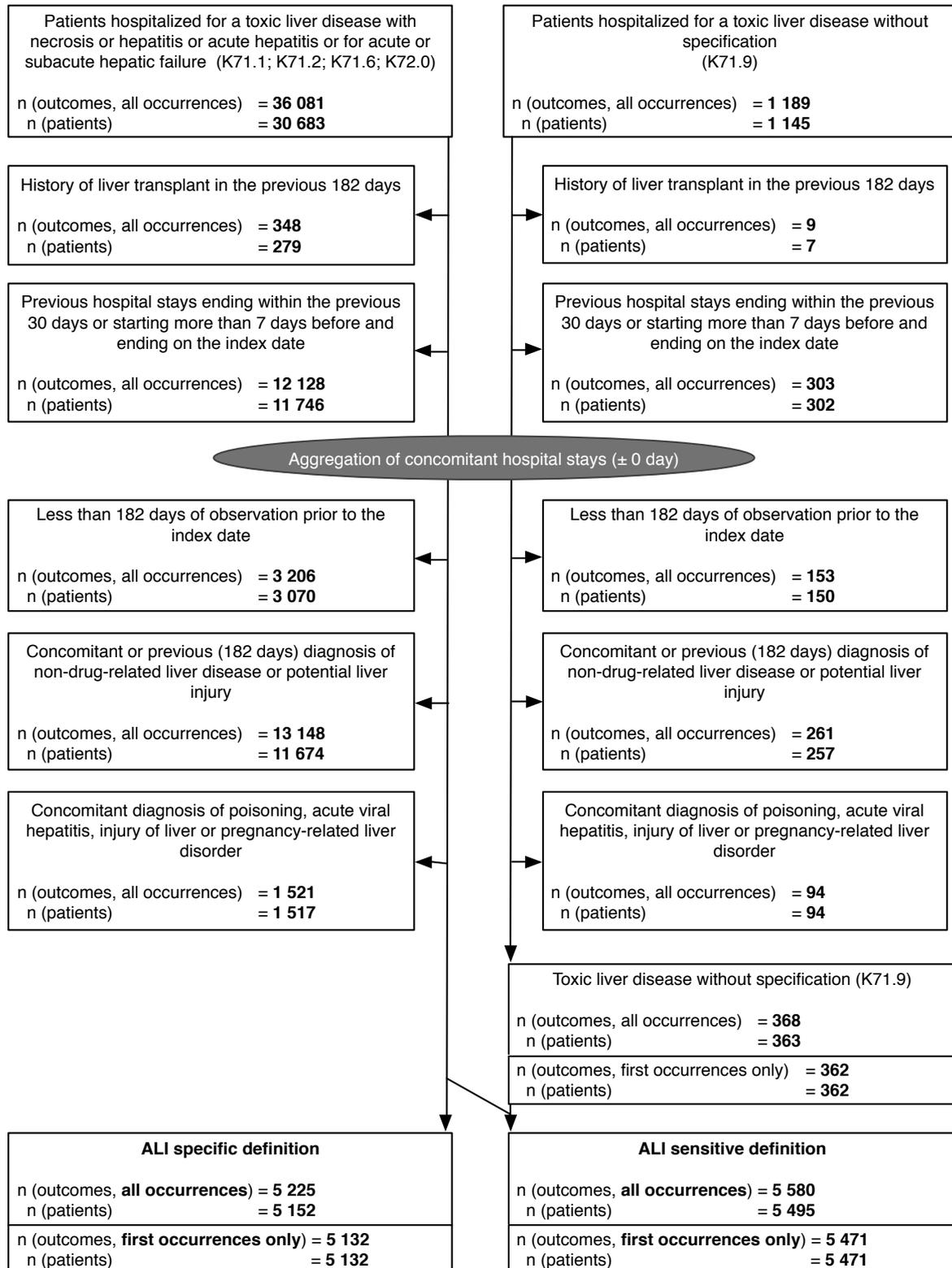


Figure 1. Selection of acute liver injury (ALI) outcomes according to sensitive and specific definitions

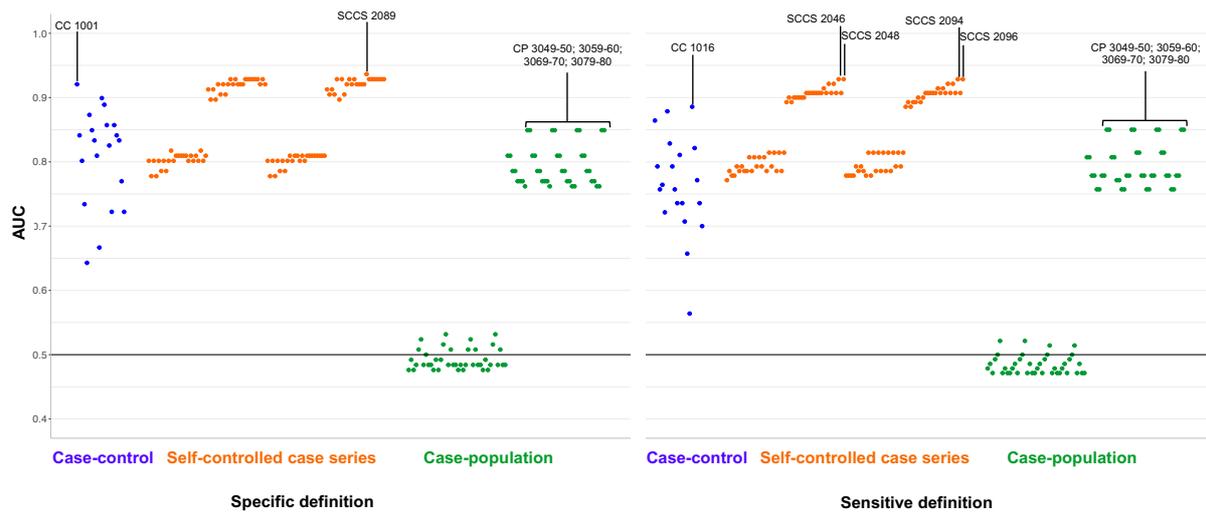


Figure 2. Area under the receiver operating characteristics curve (AUC) for case-control (CC), self-controlled case series (SCCS) and case-population (CP) approaches for acute liver injury specific and sensitive definitions. For each approach, variants with the highest AUC are mentioned.

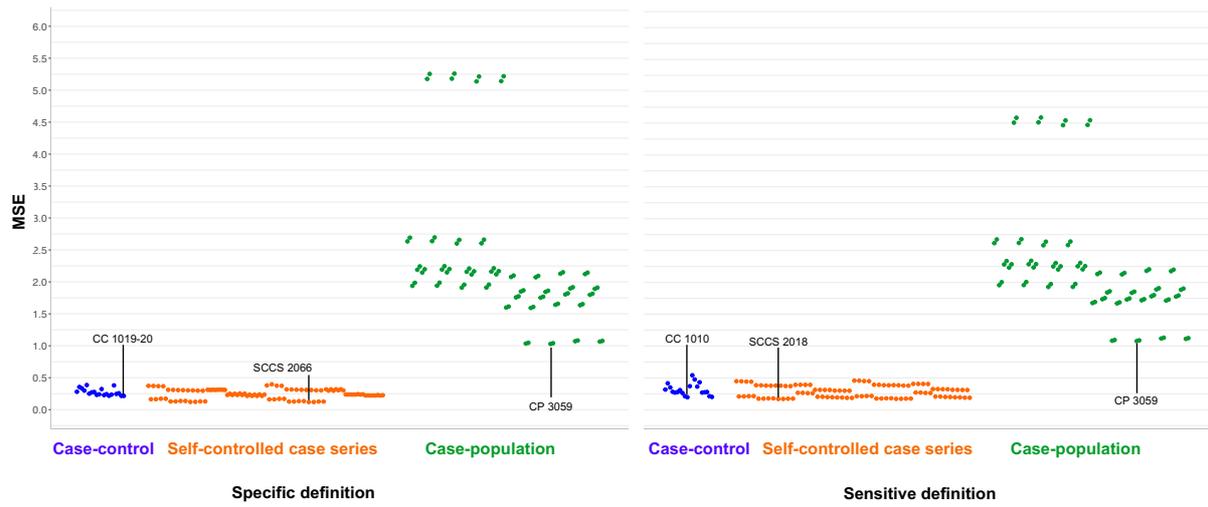


Figure 3. Means square error (MSE) for case-control (CC), self-controlled case series (SCCS) and case-population (CP) approaches for acute liver injury specific and sensitive definitions. For each approach, variants with the lowest MSE are mentioned.

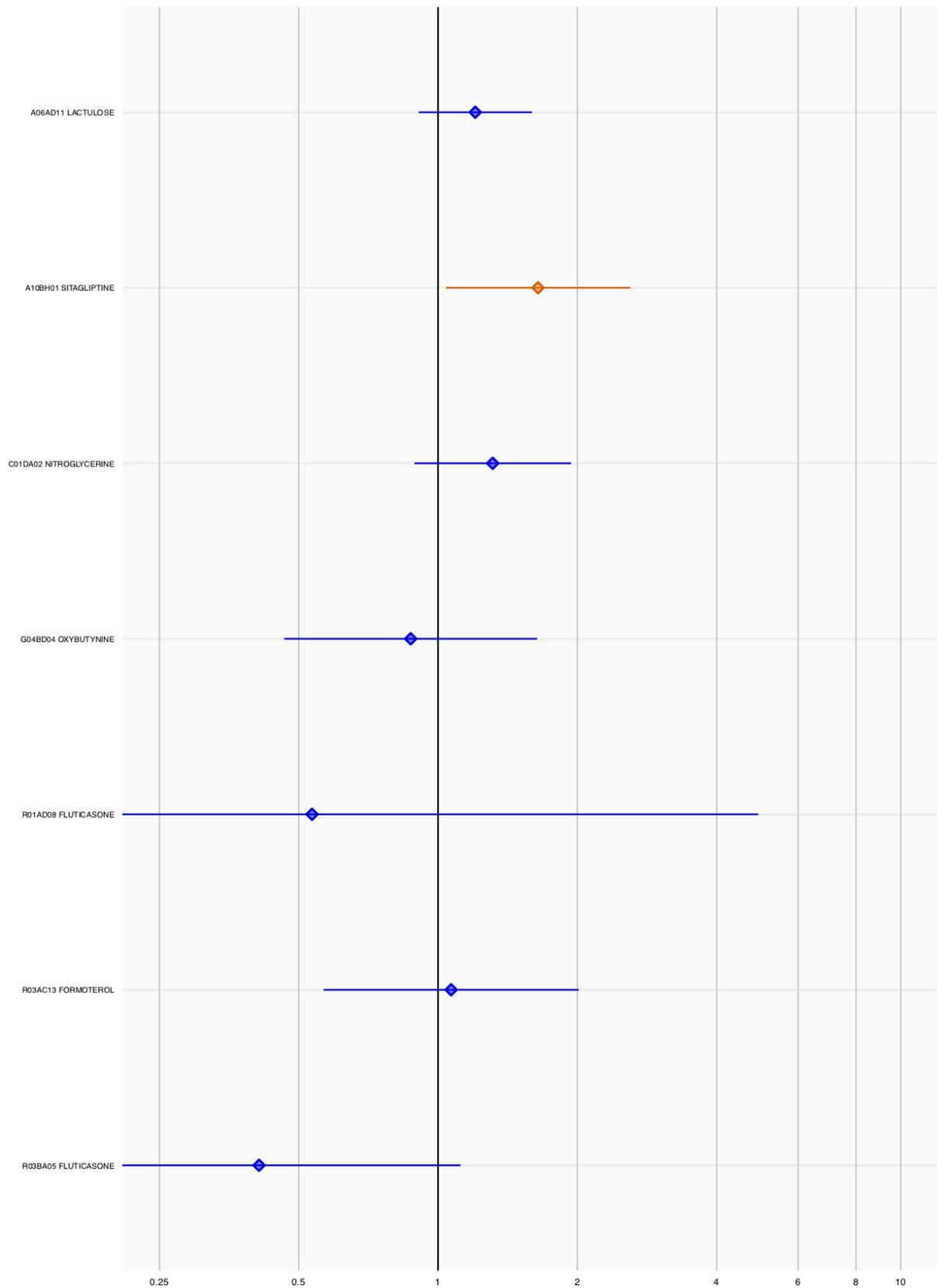


Figure 4. Point estimates of negative controls for ALLI specific definition (SCCS variant 2090) Estimates that are significantly different from 1 ($\alpha = 0.05$) are marked in orange, others are marked in blue.

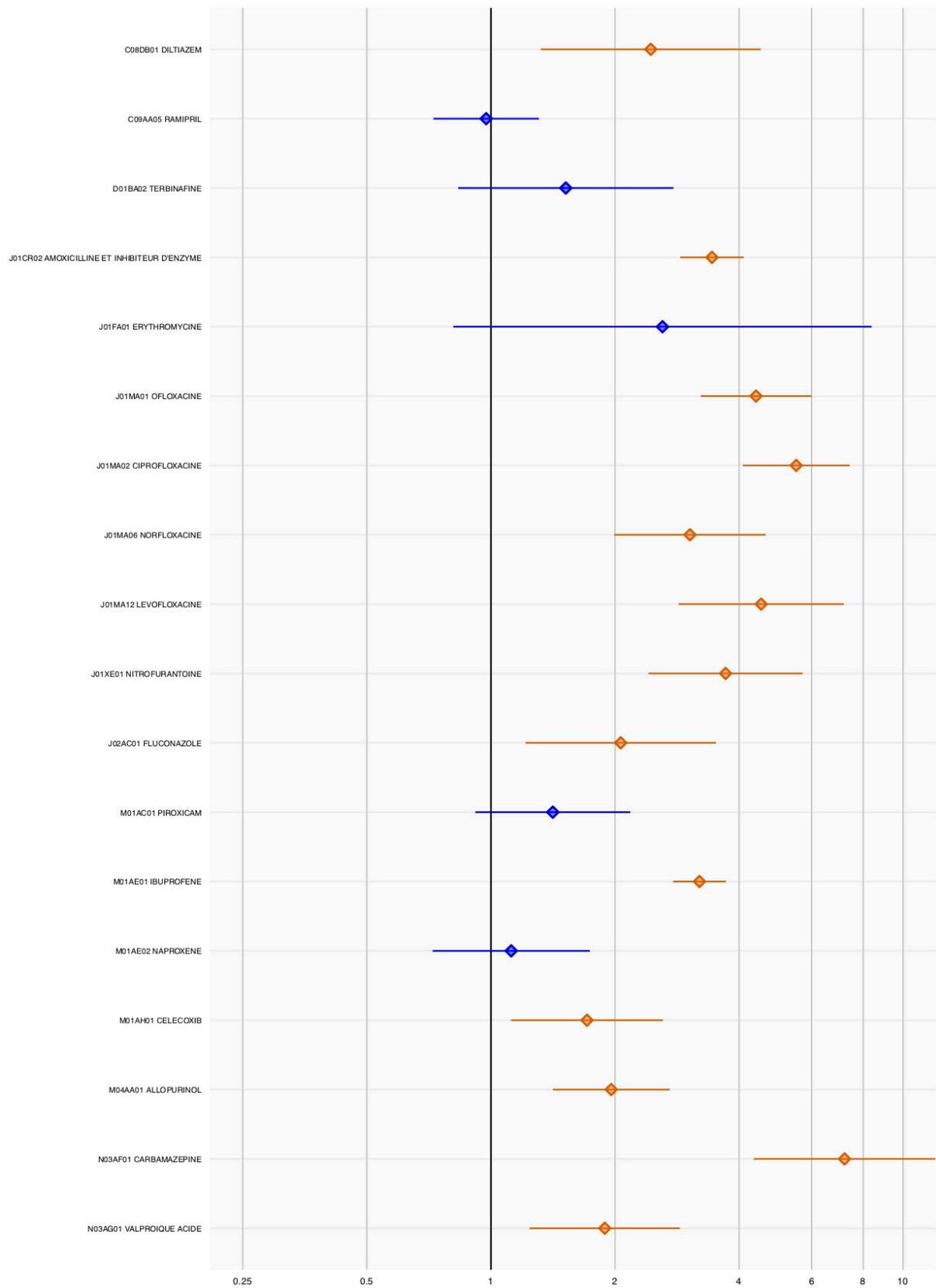


Figure 5. Point estimates of positive controls for ALLI specific definition (SCCS variant 2090). Estimates that are significantly different from 1 ($\alpha = 0.05$) are marked in orange, others are marked in blue.

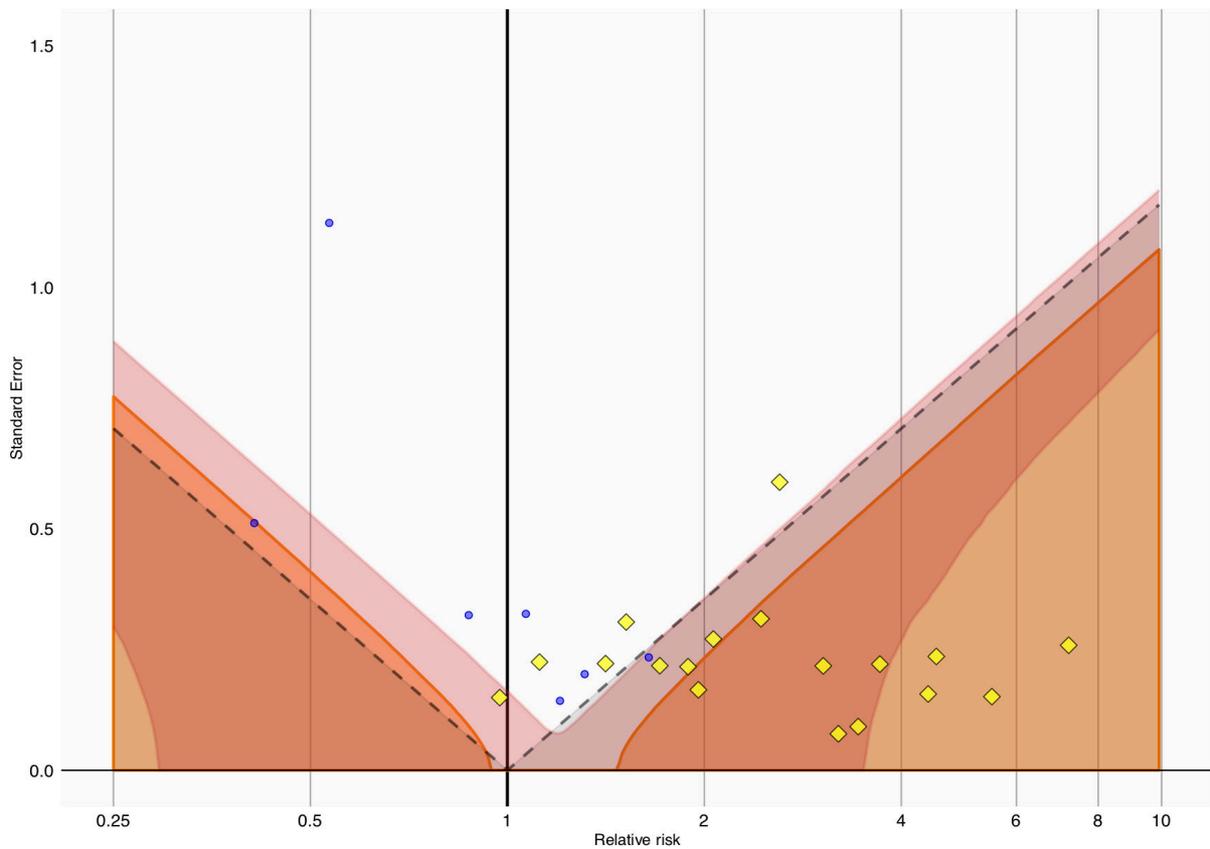


Figure 6. Point estimates from SCCS variant 2090, (ALI specific definition). Estimates below the dashed line have $p < 0.05$ using traditional p -value calculation. Estimates in the orange area have $p < 0.05$ using calibrated p -value calculation. Uncertainty in the p -value calibration is indicated by the red band. Blue dots indicate negative controls. Yellow diamonds indicate positive controls.

APPENDICES

APPENDIX 1 EXCLUSION CODES FOR ACUTE LIVER INJURY (ALI)

ICD10	Definition	In the 182 days preceding ALI	During ALI hospital stay
B18	Chronic viral hepatitis	To exclude	To exclude
C	Malignant neoplasms	To exclude	To exclude
F10	Mental and behavioural disorders due to use of alcohol	To exclude	To exclude
G31	Other degenerative diseases of nervous system, not elsewhere classified	To exclude	To exclude
I50	Heart failure	To exclude	To exclude
I81	Portal vein thrombosis	To exclude	To exclude
I85	Oesophageal varices	To exclude	To exclude
K70	Alcoholic liver disease	To exclude	To exclude
K74	Fibrosis and cirrhosis of liver	To exclude	To exclude
K76	Other diseases of liver	To exclude	To exclude
K80	Cholelithiasis	To exclude	To exclude
K83	Other diseases of biliary tract	To exclude	To exclude
R18	Ascites	To exclude	To exclude
Z95	Presence of cardiac and vascular implants and grafts	To exclude	To exclude
Z94.4	Liver transplant status	To exclude	
T36-T50	Poisoning by drugs, medicaments and biological substances		To exclude
T51-T65	Toxic effects of substances chiefly nonmedicinal as to source		To exclude
S36.1	Injury of liver or gallbladder		To exclude
B15	Acute hepatitis A		To exclude
B16	Acute hepatitis B		To exclude
B17	Other acute viral hepatitis		To exclude
B19	Unspecified viral hepatitis		To exclude
B24	Unspecified human immunodeficiency virus [HIV] disease	To exclude	To exclude
K77	Liver disorders in diseases classified elsewhere		To exclude
B25.1	Cytomegaloviral hepatitis		To exclude
K73	Chronic hepatitis, not elsewhere classified	To exclude	To exclude
K75.4	Autoimmune hepatitis	To exclude	To exclude
O26.6	Liver disorders in pregnancy, childbirth and the puerperium		To exclude

APPENDIX 2 DETECTABLE POSITIVE AND NEGATIVE CONTROLS WITH CORRESPONDING MINIMUM DETECTABLE RELATIVE RISK

ATC code	International nonproprietary name	Source reference set	Type of control	Specific definition		Sensitive definition	
				MDRR unsampled population	MDRR unsampled population		
A01AB09	MICONAZOLE	OMOP	Negative	1.77	1.70		
A02BX02	SUCRALFATE	OMOP	Negative	1.37	1.34		
A04AD01	SCOPOLAMINE	OMOP	Negative	3.27	3.27		
A06AD11	LACTULOSE	OMOP	Negative	1.11	1.10		
A07EC01	SULFASALAZINE	EU-ADR	Positive	1.45	1.43		
A10BG03	PIOGLITAZONE	OMOP	Positive	1.56	1.52		
A10BH01	SITAGLIPTINE	OMOP	Negative	1.25	1.24		
C01DA02	NITROGLYCERINE	EU-ADR	Negative	1.16	1.15		
C08CA05	NIFEDIPINE	OMOP	Positive	1.53	1.49		
C08DB01	DILTIAZEM	OMOP	Positive	1.29	1.27		
C09AA01	CAPTOPRIL	OMOP	Positive	1.72	1.65		
C09AA02	ENALAPRIL	OMOP	Positive	1.39	1.38		
C09AA03	LISINAPRIL	OMOP	Positive	1.43	1.42		
C09AA05	RAMIPRIL	OMOP	Positive	1.14	1.14		
C09AA06	QUINAPRIL	OMOP	Positive	3.96	3.96		
C09AA10	TRANDOLAPRIL	OMOP	Positive	1.77	1.70		
C09AA13	MOEXIPRIL	OMOP	Positive	5.76	5.76		
C10AD02	NICOTINIQUE ACIDE	OMOP	Positive	5.76	5.76		
D01BA01	GRISEOFULVINE	OMOP	Negative	1.80	1.77		
D01BA02	TERBINAFINE	OMOP	Positive	1.21	1.21		
G01AF04	MICONAZOLE	OMOP	Negative	1.65	1.58		
G04BD02	FLAVOXATE	OMOP	Negative	2.02	1.93		
G04BD04	OXYBUTYNINE	OMOP	Negative	1.27	1.26		
G04CA03	TERAZOSINE	EU-ADR	Negative	3.27	2.89		
H03BA02	PROPYLTHIOURACILE	OMOP	Positive	1.75	1.70		
H03BB02	THIAMAZOLE	OMOP	Positive	1.82	1.82		
J01CE02	PHENOXYMETHYLPENICILLINE	OMOP	Negative	1.45	1.44		
J01CR02	AMOXICILLINE ET INHIBITEUR DIENZYME	EU-ADR	Positive	1.06	1.05		
J01FA01	ERYTHROMYCINE	OMOP	Positive	1.28	1.27		

J01MAA01	OFLOXACINE	OMOP	Positive	1.10	1.10
J01MAA02	CIPROFLOXACINE	OMOP	Positive	1.11	1.10
J01MAA06	NORFLOXACINE	OMOP	Positive	1.11	1.10
J01MAA12	LEVOFLOXACINE	OMOP	Positive	1.14	1.14
J01XE01	NITROFURANTOINE	OMOP	Positive	1.16	1.16
J02AC01	FLUCONAZOLE	OMOP	Positive	1.15	1.14
J02AC02	ITRACONAZOLE	OMOP	Positive	2.08	2.08
J04AB02	RIFAMPICINE	OMOP	Positive	1.31	1.30
J04AC01	ISONIAZIDE	OMOP	Positive	1.49	1.45
J05AE09	TIPRANAVIR	OMOP	Positive	3.96	3.96
J05AE10	DARUNAVIR	OMOP	Positive	1.89	1.89
J05AF01	ZIDOVUDINE	OMOP	Positive	#N/A	#N/A
J05AF02	DIDANOSINE	OMOP	Positive	3.96	3.96
J05AF04	STAVUDINE	OMOP	Positive	5.76	5.76
J05AF05	LAMIVUDINE	OMOP	Positive	2.15	2.08
J05AF06	ABACAVIR	OMOP	Positive	3.96	3.96
J05AF07	TENOFOVIR DISOPROXIL	OMOP	Positive	2.02	1.97
J05AG01	NEVIRAPINE	OMOP	Positive	2.65	2.65
J05AG03	EFAVIRENZ	OMOP	Positive	2.65	2.65
L01AB01	BUSULFAN	OMOP	Positive	#N/A	#N/A
L01XE01	IMATINIB	OMOP	Positive	2.65	2.65
L02BA01	TAMOXIFENE	OMOP	Positive	1.97	1.97
L02BB01	FLUTAMIDE	OMOP	Positive	#N/A	#N/A
L03AB07	INTERFERON BETA-1A	OMOP	Positive	1.97	1.93
L04AD01	CYCLOSPORINE	OMOP	Positive	1.60	1.57
M01AB01	INDOMETACINE	OMOP	Positive	1.53	1.49
M01AB02	SULINDAC	OMOP	Positive	2.08	1.97
M01AB08	ETODOLAC	OMOP	Positive	1.31	1.30
M01AC01	PIROXICAM	OMOP	Positive	1.11	1.10
M01AE01	IBUPROFENE	OMOP	Positive	1.06	1.05
M01AE02	NAPROXENE	OMOP	Positive	1.10	1.10
M01AH01	CELECOXIB	OMOP	Positive	1.15	1.14
M01CC01	PENICILLAMINE	OMOP	Positive	2.47	2.24
M04AA01	ALLOPURINOL	OMOP	Positive	1.17	1.16
N02CA52	ERGOTAMINE EN ASSOCIATION SAUF AUX PSYCHOLEPTIQUES	OMOP	Negative	1.97	1.97
N02CC05	ALMOTRIPTAN	OMOP	Negative	1.38	1.37
N03AA03	PRIMIDONE	OMOP	Negative	2.34	2.34

N03AF01	CARBAMAZEPINE	OMOP	Positive	1.27	1.26
N03AG01	VALPROIQUE ACIDE	OMOP	Positive	1.21	1.20
N03AX09	LAMOTRIGINE	OMOP	Positive	1.33	1.32
N04BA02	LEVODOPA ET INHIBITEUR DE LA DECARBOXYLASE	EU-ADR	Negative	1.35	1.34
N04BX01	TOLCAPONE	OMOP	Positive	#N/A	#N/A
N05AH02	CLOZAPINE	OMOP	Positive	2.47	2.24
N05AN01	LITHIUM	OMOP	Negative	1.49	1.48
N07BB01	DISULFIRAM	OMOP	Positive	1.89	1.85
P01AB02	TINIDAZOLE	OMOP	Negative	1.43	1.42
R01AD08	FLUTICASONNE	OMOP	Negative	1.30	1.29
R03AC12	SALMETEROL	OMOP	Negative	1.51	1.50
R03AC13	FORMOTEROL	EU-ADR	Negative	1.25	1.24
R03BA05	FLUTICASONNE	OMOP	Negative	1.25	1.24
R06AX17	KETOTIFENE	OMOP	Negative	1.70	1.65
S01EC01	ACETAZOLAMIDE	OMOP	Positive	1.40	1.39

ATC = Anatomical Therapeutic and Chemical classification

MDRR = Minimum Detectable Relative Risk

Detectable positive controls 18 20

Detectable negative controls 7 7

APPENDIX 3 DESIGN VARIANT DESCRIPTION

Design approach	Analysis_ID	Controls per case	Required observation time prior to outcome	Outcomes to include	Risk windows	Case-control matching strategy	Remove unmatched case	Metric
CC	1001	up to 2	182d	All occurrences	7d and 0d	Age and gender	Yes	Odds ratio
CC	1002	up to 2	182d	All occurrences	30d and 0d	Age and gender	Yes	Odds ratio
CC	1003	up to 2	182d	All occurrences	60d and 0d	Age and gender	Yes	Odds ratio
CC	1004	up to 2	182d	All occurrences	60d and 7d	Age and gender	Yes	Odds ratio
CC	1005	up to 2	182d	All occurrences	60d and 15d	Age and gender	Yes	Odds ratio
CC	1006	up to 10	182d	All occurrences	7d and 0d	Age and gender	Yes	Odds ratio
CC	1007	up to 10	182d	All occurrences	30d and 0d	Age and gender	Yes	Odds ratio
CC	1008	up to 10	182d	All occurrences	60d and 0d	Age and gender	Yes	Odds ratio
CC	1009	up to 10	182d	All occurrences	60d and 7d	Age and gender	Yes	Odds ratio
CC	1010	up to 10	182d	All occurrences	60d and 15d	Age and gender	Yes	Odds ratio
CC	1011	up to 2	182d	First occurrence	7d and 0d	Age and gender	Yes	Odds ratio
CC	1012	up to 2	182d	First occurrence	30d and 0d	Age and gender	Yes	Odds ratio
CC	1013	up to 2	182d	First occurrence	60d and 0d	Age and gender	Yes	Odds ratio
CC	1014	up to 2	182d	First occurrence	60d and 7d	Age and gender	Yes	Odds ratio
CC	1015	up to 2	182d	First occurrence	60d and 15d	Age and gender	Yes	Odds ratio
CC	1016	up to 10	182d	First occurrence	7d and 0d	Age and gender	Yes	Odds ratio
CC	1017	up to 10	182d	First occurrence	30d and 0d	Age and gender	Yes	Odds ratio
CC	1018	up to 10	182d	First occurrence	60d and 0d	Age and gender	Yes	Odds ratio
CC	1019	up to 10	182d	First occurrence	60d and 7d	Age and gender	Yes	Odds ratio
CC	1020	up to 10	182d	First occurrence	60d and 15d	Age and gender	Yes	Odds ratio

Design approach	Analysis_ID	Outcomes to include	Naive period	Risk window	Pre-exposure window	Include age into the model	Include seasonality into the model	Adjusting for multiple drugs
SCCS	2001	All occurrences	182d	30d from dispensing first day	30d	No	No	No
SCCS	2002	All occurrences	182d	30d from dispensing first day	30d	No	No	Yes
SCCS	2003	All occurrences	182d	30d from dispensing first day	30d	No	Yes	No
SCCS	2004	All occurrences	182d	30d from dispensing first day	30d	No	Yes	Yes
SCCS	2005	All occurrences	182d	30d from dispensing first day	30d	Yes	No	No
SCCS	2006	All occurrences	182d	30d from dispensing first day	30d	Yes	No	Yes
SCCS	2007	All occurrences	182d	30d from dispensing first day	30d	Yes	Yes	No
SCCS	2008	All occurrences	182d	30d from dispensing first day	30d	Yes	Yes	Yes
SCCS	2009	All occurrences	182d	30d from dispensing first day	7d	No	No	No
SCCS	2010	All occurrences	182d	30d from dispensing first day	7d	No	No	Yes
SCCS	2011	All occurrences	182d	30d from dispensing first day	7d	No	Yes	No
SCCS	2012	All occurrences	182d	30d from dispensing first day	7d	No	Yes	Yes
SCCS	2013	All occurrences	182d	30d from dispensing first day	7d	Yes	No	No
SCCS	2014	All occurrences	182d	30d from dispensing first day	7d	Yes	No	Yes
SCCS	2015	All occurrences	182d	30d from dispensing first day	7d	Yes	Yes	No
SCCS	2016	All occurrences	182d	30d from dispensing first day	7d	Yes	Yes	Yes
SCCS	2017	All occurrences	182d	30d from dispensing first day	No	No	No	No
SCCS	2018	All occurrences	182d	30d from dispensing first day	No	No	No	Yes
SCCS	2019	All occurrences	182d	30d from dispensing first day	No	No	Yes	No
SCCS	2020	All occurrences	182d	30d from dispensing first day	No	No	Yes	Yes
SCCS	2021	All occurrences	182d	30d from dispensing first day	No	Yes	No	No
SCCS	2022	All occurrences	182d	30d from dispensing first day	No	Yes	No	Yes
SCCS	2023	All occurrences	182d	30d from dispensing first day	No	Yes	Yes	No
SCCS	2024	All occurrences	182d	30d from dispensing first day	No	Yes	Yes	Yes
SCCS	2025	All occurrences	182d	Period of dispensing	30d	No	No	No
SCCS	2026	All occurrences	182d	Period of dispensing	30d	No	No	Yes
SCCS	2027	All occurrences	182d	Period of dispensing	30d	No	Yes	No
SCCS	2028	All occurrences	182d	Period of dispensing	30d	No	Yes	Yes
SCCS	2029	All occurrences	182d	Period of dispensing	30d	Yes	No	No
SCCS	2030	All occurrences	182d	Period of dispensing	30d	Yes	No	Yes
SCCS	2031	All occurrences	182d	Period of dispensing	30d	Yes	Yes	No
SCCS	2032	All occurrences	182d	Period of dispensing	30d	Yes	Yes	Yes
SCCS	2033	All occurrences	182d	Period of dispensing	7d	No	No	No
SCCS	2034	All occurrences	182d	Period of dispensing	7d	No	No	Yes
SCCS	2035	All occurrences	182d	Period of dispensing	7d	No	Yes	No
SCCS	2036	All occurrences	182d	Period of dispensing	7d	No	Yes	Yes

SCCS	2037	All occurrences	182d	Period of dispensing	7d	Yes	No	No
SCCS	2038	All occurrences	182d	Period of dispensing	7d	Yes	No	Yes
SCCS	2039	All occurrences	182d	Period of dispensing	7d	Yes	Yes	No
SCCS	2040	All occurrences	182d	Period of dispensing	7d	Yes	Yes	Yes
SCCS	2041	All occurrences	182d	Period of dispensing	No	No	No	No
SCCS	2042	All occurrences	182d	Period of dispensing	No	No	No	Yes
SCCS	2043	All occurrences	182d	Period of dispensing	No	No	Yes	No
SCCS	2044	All occurrences	182d	Period of dispensing	No	No	Yes	Yes
SCCS	2045	All occurrences	182d	Period of dispensing	No	Yes	No	No
SCCS	2046	All occurrences	182d	Period of dispensing	No	Yes	No	Yes
SCCS	2047	All occurrences	182d	Period of dispensing	No	Yes	Yes	No
SCCS	2048	All occurrences	182d	Period of dispensing	No	Yes	Yes	Yes
SCCS	2049	First occurrence	182d	30d from dispensing first day	30d	No	No	No
SCCS	2050	First occurrence	182d	30d from dispensing first day	30d	No	No	Yes
SCCS	2051	First occurrence	182d	30d from dispensing first day	30d	No	Yes	No
SCCS	2052	First occurrence	182d	30d from dispensing first day	30d	No	Yes	Yes
SCCS	2053	First occurrence	182d	30d from dispensing first day	30d	Yes	No	No
SCCS	2054	First occurrence	182d	30d from dispensing first day	30d	Yes	No	Yes
SCCS	2055	First occurrence	182d	30d from dispensing first day	30d	Yes	Yes	No
SCCS	2056	First occurrence	182d	30d from dispensing first day	30d	Yes	Yes	Yes
SCCS	2057	First occurrence	182d	30d from dispensing first day	7d	No	No	No
SCCS	2058	First occurrence	182d	30d from dispensing first day	7d	No	No	Yes
SCCS	2059	First occurrence	182d	30d from dispensing first day	7d	No	Yes	No
SCCS	2060	First occurrence	182d	30d from dispensing first day	7d	No	Yes	Yes
SCCS	2061	First occurrence	182d	30d from dispensing first day	7d	Yes	No	No
SCCS	2062	First occurrence	182d	30d from dispensing first day	7d	Yes	No	Yes
SCCS	2063	First occurrence	182d	30d from dispensing first day	7d	Yes	Yes	No
SCCS	2064	First occurrence	182d	30d from dispensing first day	7d	Yes	Yes	Yes
SCCS	2065	First occurrence	182d	30d from dispensing first day	No	No	No	No
SCCS	2066	First occurrence	182d	30d from dispensing first day	No	No	No	Yes
SCCS	2067	First occurrence	182d	30d from dispensing first day	No	No	Yes	No
SCCS	2068	First occurrence	182d	30d from dispensing first day	No	No	Yes	Yes
SCCS	2069	First occurrence	182d	30d from dispensing first day	No	Yes	No	No
SCCS	2070	First occurrence	182d	30d from dispensing first day	No	Yes	No	Yes
SCCS	2071	First occurrence	182d	30d from dispensing first day	No	Yes	Yes	No
SCCS	2072	First occurrence	182d	30d from dispensing first day	No	Yes	Yes	Yes
SCCS	2073	First occurrence	182d	Period of dispensing	30d	No	No	No
SCCS	2074	First occurrence	182d	Period of dispensing	30d	No	No	Yes
SCCS	2075	First occurrence	182d	Period of dispensing	30d	No	Yes	No
SCCS	2076	First occurrence	182d	Period of dispensing	30d	No	Yes	No
SCCS	2077	First occurrence	182d	Period of dispensing	30d	Yes	No	Yes

SCCS	2078	First occurrence	182d	Period of dispensing	30d	Yes	No	Yes
SCCS	2079	First occurrence	182d	Period of dispensing	30d	Yes	Yes	No
SCCS	2080	First occurrence	182d	Period of dispensing	30d	Yes	Yes	Yes
SCCS	2081	First occurrence	182d	Period of dispensing	7d	No	No	No
SCCS	2082	First occurrence	182d	Period of dispensing	7d	No	No	Yes
SCCS	2083	First occurrence	182d	Period of dispensing	7d	No	Yes	No
SCCS	2084	First occurrence	182d	Period of dispensing	7d	No	Yes	Yes
SCCS	2085	First occurrence	182d	Period of dispensing	7d	Yes	No	No
SCCS	2086	First occurrence	182d	Period of dispensing	7d	Yes	No	Yes
SCCS	2087	First occurrence	182d	Period of dispensing	7d	Yes	Yes	No
SCCS	2088	First occurrence	182d	Period of dispensing	7d	Yes	Yes	Yes
SCCS	2089	First occurrence	182d	Period of dispensing	No	No	No	No
SCCS	2090	First occurrence	182d	Period of dispensing	No	No	No	Yes
SCCS	2091	First occurrence	182d	Period of dispensing	No	No	Yes	No
SCCS	2092	First occurrence	182d	Period of dispensing	No	No	Yes	Yes
SCCS	2093	First occurrence	182d	Period of dispensing	No	Yes	No	No
SCCS	2094	First occurrence	182d	Period of dispensing	No	Yes	No	Yes
SCCS	2095	First occurrence	182d	Period of dispensing	No	Yes	Yes	No
SCCS	2096	First occurrence	182d	Period of dispensing	No	Yes	Yes	Yes

Design Approach	Analysis_ID	Required observation time prior to outcome	Outcome to include	Exposure windows and exclusion period	Approach	Extrapolation of the aggregated control data	Metrics
CP	3001	182d	All occurrences	30d and 0d	Per-user approach	Raw (no stratification)	CPR
CP	3002	182d	All occurrences	30d and 0d	Per-user approach	Raw (no stratification)	RR predicted
CP	3003	182d	All occurrences	60d and 0d	Per-user approach	Raw (no stratification)	CPR
CP	3004	182d	All occurrences	60d and 0d	Per-user approach	Raw (no stratification)	RR predicted
CP	3005	182d	All occurrences	60d and 15d	Per-user approach	Raw (no stratification)	CPR
CP	3006	182d	All occurrences	60d and 15d	Per-user approach	Raw (no stratification)	RR predicted
CP	3007	182d	All occurrences	60d and 7d	Per-user approach	Raw (no stratification)	CPR
CP	3008	182d	All occurrences	60d and 7d	Per-user approach	Raw (no stratification)	RR predicted
CP	3009	182d	All occurrences	7d and 0d	Per-user approach	Raw (no stratification)	CPR
CP	3010	182d	All occurrences	7d and 0d	Per-user approach	Raw (no stratification)	RR predicted
CP	3011	182d	First occurrence	30d and 0d	Per-user approach	Raw (no stratification)	CPR
CP	3012	182d	First occurrence	30d and 0d	Per-user approach	Raw (no stratification)	RR predicted
CP	3013	182d	First occurrence	60d and 0d	Per-user approach	Raw (no stratification)	CPR
CP	3014	182d	First occurrence	60d and 0d	Per-user approach	Raw (no stratification)	RR predicted
CP	3015	182d	First occurrence	60d and 15d	Per-user approach	Raw (no stratification)	CPR
CP	3016	182d	First occurrence	60d and 15d	Per-user approach	Raw (no stratification)	RR predicted
CP	3017	182d	First occurrence	60d and 7d	Per-user approach	Raw (no stratification)	CPR
CP	3018	182d	First occurrence	60d and 7d	Per-user approach	Raw (no stratification)	RR predicted
CP	3019	182d	First occurrence	7d and 0d	Per-user approach	Raw (no stratification)	CPR
CP	3020	182d	First occurrence	7d and 0d	Per-user approach	Raw (no stratification)	RR predicted
CP	3021	182d	All occurrences	30d and 0d	Per-user approach	Stratified on age and gender	CPR
CP	3022	182d	All occurrences	30d and 0d	Per-user approach	Stratified on age and gender	RR predicted
CP	3023	182d	All occurrences	60d and 0d	Per-user approach	Stratified on age and gender	CPR
CP	3024	182d	All occurrences	60d and 0d	Per-user approach	Stratified on age and gender	RR predicted
CP	3025	182d	All occurrences	60d and 15d	Per-user approach	Stratified on age and gender	CPR
CP	3026	182d	All occurrences	60d and 15d	Per-user approach	Stratified on age and gender	RR predicted
CP	3027	182d	All occurrences	60d and 7d	Per-user approach	Stratified on age and gender	CPR
CP	3028	182d	All occurrences	60d and 7d	Per-user approach	Stratified on age and gender	RR predicted
CP	3029	182d	All occurrences	7d and 0d	Per-user approach	Stratified on age and gender	CPR
CP	3030	182d	All occurrences	7d and 0d	Per-user approach	Stratified on age and gender	RR predicted
CP	3031	182d	First occurrence	30d and 0d	Per-user approach	Stratified on age and gender	CPR
CP	3032	182d	First occurrence	30d and 0d	Per-user approach	Stratified on age and gender	RR predicted
CP	3033	182d	First occurrence	60d and 0d	Per-user approach	Stratified on age and gender	CPR
CP	3034	182d	First occurrence	60d and 0d	Per-user approach	Stratified on age and gender	RR predicted

CP	3035	182d	First occurrence	60d and 15d	Per-user approach	Stratified on age and gender	CPR
CP	3036	182d	First occurrence	60d and 15d	Per-user approach	Stratified on age and gender	RR predicted
CP	3037	182d	First occurrence	60d and 7d	Per-user approach	Stratified on age and gender	CPR
CP	3038	182d	First occurrence	60d and 7d	Per-user approach	Stratified on age and gender	RR predicted
CP	3039	182d	First occurrence	7d and 0d	Per-user approach	Stratified on age and gender	CPR
CP	3040	182d	First occurrence	7d and 0d	Per-user approach	Stratified on age and gender	RR predicted
CP	3041	182d	All occurrences	30d and 0d	Person-time approach	Raw (no stratification)	CPR
CP	3042	182d	All occurrences	30d and 0d	Person-time approach	Raw (no stratification)	RR predicted
CP	3043	182d	All occurrences	60d and 0d	Person-time approach	Raw (no stratification)	CPR
CP	3044	182d	All occurrences	60d and 0d	Person-time approach	Raw (no stratification)	RR predicted
CP	3045	182d	All occurrences	60d and 15d	Person-time approach	Raw (no stratification)	CPR
CP	3046	182d	All occurrences	60d and 15d	Person-time approach	Raw (no stratification)	RR predicted
CP	3047	182d	All occurrences	60d and 7d	Person-time approach	Raw (no stratification)	CPR
CP	3048	182d	All occurrences	60d and 7d	Person-time approach	Raw (no stratification)	RR predicted
CP	3049	182d	All occurrences	7d and 0d	Person-time approach	Raw (no stratification)	CPR
CP	3050	182d	All occurrences	7d and 0d	Person-time approach	Raw (no stratification)	RR predicted
CP	3051	182d	First occurrence	30d and 0d	Person-time approach	Raw (no stratification)	CPR
CP	3052	182d	First occurrence	30d and 0d	Person-time approach	Raw (no stratification)	RR predicted
CP	3053	182d	First occurrence	60d and 0d	Person-time approach	Raw (no stratification)	CPR
CP	3054	182d	First occurrence	60d and 0d	Person-time approach	Raw (no stratification)	RR predicted
CP	3055	182d	First occurrence	60d and 15d	Person-time approach	Raw (no stratification)	CPR
CP	3056	182d	First occurrence	60d and 15d	Person-time approach	Raw (no stratification)	RR predicted
CP	3057	182d	First occurrence	60d and 7d	Person-time approach	Raw (no stratification)	CPR
CP	3058	182d	First occurrence	60d and 7d	Person-time approach	Raw (no stratification)	RR predicted
CP	3059	182d	First occurrence	7d and 0d	Person-time approach	Raw (no stratification)	CPR
CP	3060	182d	First occurrence	7d and 0d	Person-time approach	Raw (no stratification)	RR predicted
CP	3061	182d	All occurrences	30d and 0d	Person-time approach	Stratified on age and gender	CPR
CP	3062	182d	All occurrences	30d and 0d	Person-time approach	Stratified on age and gender	RR predicted
CP	3063	182d	All occurrences	60d and 0d	Person-time approach	Stratified on age and gender	CPR
CP	3064	182d	All occurrences	60d and 0d	Person-time approach	Stratified on age and gender	RR predicted
CP	3065	182d	All occurrences	60d and 15d	Person-time approach	Stratified on age and gender	CPR
CP	3066	182d	All occurrences	60d and 15d	Person-time approach	Stratified on age and gender	RR predicted
CP	3067	182d	All occurrences	60d and 7d	Person-time approach	Stratified on age and gender	CPR
CP	3068	182d	All occurrences	60d and 7d	Person-time approach	Stratified on age and gender	RR predicted
CP	3069	182d	All occurrences	7d and 0d	Person-time approach	Stratified on age and gender	CPR
CP	3070	182d	All occurrences	7d and 0d	Person-time approach	Stratified on age and gender	RR predicted
CP	3071	182d	First occurrence	30d and 0d	Person-time approach	Stratified on age and gender	CPR

CP	3072	182d	First occurrence	30d and 0d	Person-time approach	Stratified on age and gender	RR predicted
CP	3073	182d	First occurrence	60d and 0d	Person-time approach	Stratified on age and gender	CPR
CP	3074	182d	First occurrence	60d and 0d	Person-time approach	Stratified on age and gender	RR predicted
CP	3075	182d	First occurrence	60d and 15d	Person-time approach	Stratified on age and gender	CPR
CP	3076	182d	First occurrence	60d and 15d	Person-time approach	Stratified on age and gender	RR predicted
CP	3077	182d	First occurrence	60d and 7d	Person-time approach	Stratified on age and gender	CPR
CP	3078	182d	First occurrence	60d and 7d	Person-time approach	Stratified on age and gender	RR predicted
CP	3079	182d	First occurrence	7d and 0d	Person-time approach	Stratified on age and gender	CPR
CP	3080	182d	First occurrence	7d and 0d	Person-time approach	Stratified on age and gender	RR predicted

**APPENDIX 4 PERFORMANCES DESIGN VARIANTS TESTED IN ALI SPECIFIC
AND SENSITIVE**

AlcaponeAli_Appendix4_v0.1.xlsx

**APPENDIX 5 CONTROL ESTIMATES FROM DESIGN VARIANTS TESTED IN
ALI SPECIFIC AND SENSITIVE DEFINITIONS**

AlcaponeAli_Appendix5_v0.1.xlsx

3. GENERAL DISCUSSION, CONCLUSION AND PERSPECTIVE

3.1. IT constraints

IT constraints was something we didn't really envisage when we were designing the project. However, we spent almost one year overcoming them. We faced our first IT issues at the implementation stage of ALCAPONE, at the beginning of the feasibility study, with the deployment of the SelfControlledCaseSeries package. As an open-source package in a collaborative framework, OHDSI SCCS package is regularly enhanced, with the integration of new options or bug correction. Through ALCAPONE we went very deep in the package options, testing some specific combinations of parameters in sometimes very large datasets. In a way, we were "the guinea pig" as M. Schuemie said. Considering coding error and the package debugging, we spent almost 6 months to be able to run a first complex SCCS design. We then tried to replicate it in a different machine, and a new surprise arose. The same script executed with the same R and package version gave different results. Thinking it was some sort of software issue we tried to boot the different machines, one after the other, on a same hard drive, but the problem was still present. Seeing no software solution to the problem, we tried to investigate more deeply the source of the discrepancies in the R package. It appeared that the discrepancy was observed for SCCS adjusting for event-dependant observation time as proposed by Farrington *et al.*¹⁸⁸ We went through some tests of the event-dependant observation period using simulated data, and it appeared that even if the optimum for the censor model computed by each platform was the same, slightly different estimates were generated. As no clue were found to the sensitivity to the choice of the hardware, and given that the discrepancies observed could have had significant impact on the final results, we lastly decided not to consider this option for the experiment. This example shows that even if this kind of complex model go through a large set of automated validations tests and simulation studies during its development, as is the case for packages developed by OHDSI,¹⁸⁹ there is always room for unexpected issues. Integrating comprehensive tests across different hardware environments to quality insurance routine could be a solution to control this type of issue. However, since not all possible configurations will ever be tested, the best way to ensure the consistency of a tool (and a potential result) is to test it across its own IT framework. Moreover, a descriptive of the platform and the hardware used to run analyses should always be provided in appendix of publications and reports.

In addition, ALCAPONE project taught us that computational power and calculation time are also to be carefully considered when designing a project involving large databases and complex calculations since they directly impact study timelines and deliverables.

3.2. Interest of health professional point of view

There is always a great temptation to consider results of such experiment just as numbers, going no further than applying bias-proofed supposed state-of-the-art methodology and taking the design variant apparently showing the best performances, whether generated results seem to be consistent or not. Here, however, in looking to the controls in more details, it clearly appeared that even the most advanced design does not address all the biases, and it is at this point that the first part of the word pharmacoepidemiology takes on its full meaning. Clinical pharmacology and physician point of view are required to correctly interpret the generated results: to judge if provided estimates likely result from a pharmacological effects or potential biases. Context matters when interpreting. It is the context which incites a physician to prescribe one drug rather than another, inducing potential channeling bias. In this case, context can take the form of patient symptoms, prescription guidelines, healthcare environment, drug reimbursement, or even physician class of age. Given that context is multifactorial by nature and therefore complex, who can help better understand its influence if not the actors themselves: the healthcare professionals? In this thesis work, a full comprehension of the SNDS variables was necessary to achieve a faithful ETL to the OMOP CDM, especially the outcome and drug related ones. Since drug information in primary care is almost fully captured by pharmacies, a pharmacist point of view was very valuable to correctly interpret SNIIRAM information. Dispensing rules and patients-pharmacist relationship are country specific. Even if some documents provide guidelines on how drugs should be dispensed, only a health professional involved in a day-to-day practice has a global vision of what is really happening on the field, how patients behave, how the rules are interpreted, how they are applied... or not, and especially how this is translated in the database. In the same way as a physician is not able to describe what happen in a pharmacy, a pharmacist is not able to depict what can occur in a doctor practice. A significant part of the confounding observed in pharmacoepidemiological studies could be understood appreciating how physicians

managed their patients. GPs are the backbone of patient management. They are responsible for patient follow-up in primary care, including dose adjustment, prescription renewal, as well as non-serious acute disease treatment. They know how to interpret global patient care pathways. However, since SNDS reports hospital discharge summaries, specialist physician inputs are of special importance. They are the ones who have a comprehensive view of the disease management from the symptoms to the treatment, including the diagnostic stage. As pharmacists help to understand SNIIRAM data, specialists are the key to interpret PMSI. However, although drugs in primary care settings are entered in the SNDS directly when they are dispensed, medical procedures and diagnostics code lead to a rather complex process in hospitals. Most of the time, the physician involved in patient treatment is not coding. According to the hospital or the department, coding process can be achieved through other physicians, residents, nurses or sometimes secretaries, with or without specific formation.¹⁹⁰ Since coding quality is correlated to the funds paid to the hospitals, most of them have also a dedicated “Medical information department” including physicians, in charge of the management and the quality of the data entered to the PMSI.¹⁹¹ Sometimes finding or interpreting a procedure or diagnostic code can be quite challenging, medical nomenclatures being complex. In this case, data-coding specialist input are very valuable. They help to make the link between the original medical condition and the way it is translated in the database. Through the definition of accurate HOI in the first stage of the project, the ETL of SNDS data to the OMOP CDM and the understanding of the generated estimates, ALCAPONE demonstrated that the best way to achieve a faithful decoding of the data, an appropriate data processing and a comprehensive interpretation of the results is to get a multidisciplinary team to work together.

3.3. Routine implementation and generalizability

SNDS is often considered as the world's largest continuous homogeneous claims database.¹⁴⁰ Information about outpatient drug exposure is recorded almost exhaustively and in detail, even if the fact that the dispensed drugs were actually taken by the patient, it is always subject to uncertainty. The quality of the outcome information available in hospital discharge summaries is ensured by the daily work of medical information departments.^{191, 192} Obviously, there may always be some discrepancies between the captured data and the actual patient conditions, but

ad-hoc validation studies seem to prove that SNDS provide specific diagnostic information.¹⁹³ Even if medical results, risk factors and outpatients conditions are missing, the quality and the richness of data available often allow to overcome this limitation. Proxies can serve in place of untracked diagnoses: dyslipidemia are not well tracked since the disease is mainly managed by GPs, but lipid lowering therapies dispensings are. Moreover, succession of cares are suggestive enough to confirm the doubtful health status of patients.^{194, 195} With 99% of the 66.6 million French inhabitants included, SNDS representativeness against French population is not a matter for discussion; nor is its power. These assets combined make of the SNDS a good support to perform signal detection and validation.

Comparing pharmacoepidemiological case-based approaches, this thesis aimed to take advantage of this potential. Obviously, even the best-performing designs identified in this work could be further improved, but the empirical assessment showed that performances of SCCS as it is, with carefully selected parameters, were decent enough and superior to those of CC and CP for the identification of drugs associated with ALI, AKI, MI and UGIB. Through the UGIB example, false positives revealed that some biases remained, especially protopathic bias and confounding by indication. However, most of the time, a clinical point of view allowed discrimination between true and false positives. Moreover, as we have seen in ALI example, a restrictive selection of cases could help to reduce this residual confounding. Besides, some positive controls were not detected during the experiment, sometimes showing no effect at all, or positive but non-significant effect. The positive but non-significant effect can be reasonably attributed to a lack of power resulting from a small number of outcomes or a weak exposure. One wonders about the real impact of such a drug in the overall population. Furthermore, since pharmacoepidemiology captures the actual effect of a drug in real life conditions, the absence of association can reasonably suggest that the event of interest is not a safety issue for the considered medicine in the day-to-day practice. This could be the consequence of the actual innocuity of the drug, or of confounding by (contra)indication, which would mean that the existing risk is correctly managed.

All in all, these works showed that SNDS is perfectly suitable to generate drug safety alerts in an accurate manner. Thus, a pertinent interpretation by health specialists of the estimates generated by the previously highlighted reference designs in the SNDS should provide valuable input for drug safety alert generations at a national level.

Such method can be used to validate a signal generated through another source and quantify the potential risk, or to screen routinely a large set of newly marketed drugs. To do so, reference methods could first be applied across all the drugs of a SNDS extraction. Risk already documented would be ruled out, and emerging alerts carefully studied to distinguish between biased, potential, and confirmed alerts. Second, newly marketed or suspected drugs could be screened on a yearly basis. This approach extend to other outcomes of interest for drug safety could consist in substantial progress in pharmacovigilance in France. However, the development of such tools should not overshadow a significant part of the adverse events observed in clinical practice are well-known adverse reactions of old drugs.¹⁹⁶ Patients, pharmacists and physicians must understand that drugs are not without risk. The more effective or powerful the drug is, the greater the risk of adverse reactions. In the end, there are only more or less safe (or dangerous) ways of prescribing and using them.⁵

REFERENCES

1. Strom BL, Kimmel SE, Hennessy S. Pharmacoepidemiology. 2012, Page 1 online resource.
2. Wettermark B. The intriguing future of pharmacoepidemiology. *Eur J Clin Pharmacol* 2013, 69 Suppl 1:43-51.
3. Cartwright N. What are randomised controlled trials good for? *Philosophical Studies* 2009, 147:59.
4. Evans SJ. Pharmacoepidemiology. *Br J Clin Pharmacol* 2012, 73:973-978.
5. Moore N. The past, present and perhaps future of pharmacovigilance: homage to Folke Sjoqvist. *Eur J Clin Pharmacol* 2013, 69 Suppl 1:33-41.
6. Bégaud B. *Dictionary of pharmacoepidemiology*. Chichester ; New York: Wiley; 2000.
7. Blin P, Dureau-Pournin C, Cottin Y, Benichou J, Mismetti P, Abouelfath A, Lassalle R, Droz C, Moore N. Comparative Effectiveness and Safety of Standard or Reduced Dose Dabigatran vs. Rivaroxaban in Nonvalvular Atrial Fibrillation. *Clinical Pharmacology & Therapeutics* 2019, 105:1439-1455.
8. World Health Organization. *The importance of pharmacovigilance. Safety monitoring of medicinal products*. Geneva, Switzerland: World Health Organization,; 2002.
9. International Society for Pharmacoepidemiology. About Pharmacoepidemiology. Available at: <https://www.pharmacoepi.org/about-ispe/about-pharmacoepidemiology/>. (Accessed 06/18)
10. the Uppsala Monitoring Centre. *Safety Monitoring of Medicinal Products: Guidelines for Setting Up and Running a Pharmacovigilance Centre*. Uppsala, Sweden: the Uppsala Monitoring Centre,; 2000.
11. Faillie JL, Montastruc F, Montastruc JL, Pariente A. Pharmacoepidemiology and its input to pharmacovigilance. *Therapie* 2016, 71:211-216.
12. Lund JL, Richardson DB, Sturmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. *Curr Epidemiol Rep* 2015, 2:221-228.
13. Yoshida K, Solomon DH, Kim SC. Active-comparator design and new-user design in observational studies. *Nat Rev Rheumatol* 2015, 11:437-441.
14. Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiology* 2009, 20:512-522.
15. Droz-Perroteau C, Blin P, Dureau-Pournin C, Thomas D, Danchin N, Tricoire J, Paillard F, Hercberg S, Guize L, Guiard E, et al. Six-year survival study after myocardial infarction: The EOLE prospective cohort study. Long-term survival after MI. *Therapie* 2019, 74:459-468.
16. Roujeau J-C, Kelly JP, Naldi L, Rzany B, Stern RS, Anderson T, Auquier A, Bastuji-Garin S, Correia O, Locati F, et al. Medication Use and the Risk of Stevens–Johnson Syndrome or Toxic Epidermal Necrolysis. *New England Journal of Medicine* 1995, 333:1600-1608.

17. Shapiro S. The design of a study of the drug etiology of agranulocytosis and aplastic anemia. *European Journal of Clinical Pharmacology* 1983, 24:833-836.
18. Dartigues JF, Gagnon M, Michel P, Letenneur L, Commenges D, Barberger-Gateau P, Auriaud S, Rigal B, Bedry R, Alperovitch A, et al. [The Paquid research program on the epidemiology of dementia. Methods and initial results]. *Rev Neurol (Paris)* 1991, 147:225-230.
19. Goldberg M, Leclerc A, Zins M. Cohort Profile Update: The GAZEL Cohort Study. *International Journal of Epidemiology* 2014, 44:77-77g.
20. The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). Guide on Methodological Standards in Pharmacoepidemiology (Revision 7). Available at: http://www.encepp.eu/standards_and_guidances.
21. Kleinbaum DG, Morgenstern H, Kupper LL. Selection bias in epidemiologic studies. *Am J Epidemiol* 1981, 113:452-463.
22. Heun R, Burkart M, Maier W. Selection biases during recruitment of patients and relatives for a family study in the elderly. *Journal of Psychiatric Research* 1995, 29:491-504.
23. Coughlin SS. Recall bias in epidemiologic studies. *Journal of Clinical Epidemiology* 1990, 43:87-91.
24. Parsons HM. What Happened at Hawthorne?: New evidence suggests the Hawthorne effect resulted from operant reinforcement contingencies. *Science* 1974, 183:922-932.
25. Wickstrom G, Bendix T. The "Hawthorne effect"--what did the original Hawthorne studies actually show? *Scand J Work Environ Health* 2000, 26:363-367.
26. Droz-Perroteau C. La persistance des traitements en vie réelle comme mesure d'impacte en pharmaco-épidémiologie. *Ecole Doctorale Sociétés, Politique, Santé Publique* 2012.
27. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, Smeeth L. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol* 2015, 44:827-836.
28. Ohno-Machado L. Realizing the full potential of electronic health records: the role of natural language processing. *J Am Med Inform Assoc* 2011, 18:539.
29. IBM. What IBM MarketScan Research Databases can do for you. Available at: <https://www.ibm.com/us-en/marketplace/marketscan-research-databases>. (Accessed 06/21)
30. LOI n° 2016-41 du 26 janvier 2016 de modernisation de notre système de santé.
31. Zins M, Goldberg M. The French CONSTANCES population-based cohort: design, inclusion and follow-up. *Eur J Epidemiol* 2015, 30:1317-1328.
32. Avorn J. Evaluating drug effects in the post-Vioxx world: there must be a better way. *Circulation* 2006, 113:2173-2176.

33. Wadman M. Experts call for active surveillance of drug safety. *Nature* 2007, 446:358-359.
34. Coloma PM, Schuemie MJ, Trifiro G, Gini R, Herings R, Hippisley-Cox J, Mazzaglia G, Giaquinto C, Corrao G, Pedersen L, et al. Combining electronic healthcare databases in Europe to allow for large-scale drug safety monitoring: the EU-ADR Project. *Pharmacoepidemiol Drug Saf* 2011, 20:1-11.
35. Coloma PM, Avillach P, Salvo F, Schuemie MJ, Ferrajolo C, Pariente A, Fourier-Réglat A, Molokhia M, Patadia V, van der Lei J, et al. A Reference Standard for Evaluation of Methods for Drug Safety Signal Detection Using Electronic Healthcare Record Databases. *Drug Safety* 2013, 36:13-23.
36. Schuemie MJ, Coloma PM, Straatman H, Herings RM, Trifiro G, Matthews JN, Prieto-Merino D, Molokhia M, Pedersen L, Gini R, et al. Using electronic health care records for drug safety signal detection: a comparative evaluation of statistical methods. *Med Care* 2012, 50:890-897.
37. Schuemie MJ, Gini R, Coloma PM, Straatman H, Herings RM, Pedersen L, Innocenti F, Mazzaglia G, Picelli G, van der Lei J, et al. Replication of the OMOP experiment in Europe: evaluating methods for risk identification in electronic health record databases. *Drug Saf* 2013, 36 Suppl 1:S159-169.
38. Valkhoff VE, Schade R, t Jong GW, Romio S, Schuemie MJ, Arfe A, Garbe E, Herings R, Lucchi S, Picelli G, et al. Population-based analysis of non-steroidal anti-inflammatory drug use among children in four European countries in the SOS project: what size of data platforms and which study designs do we need to assess safety issues? *BMC Pediatr* 2013, 13:192.
39. EU-ADR Alliance. About Us. Available at: <http://eu-adr-alliance.com/>. (Accessed 06/20)
40. Sentinel Coordinating Center. Available at: <https://www.sentinelinitiative.org/>. (Accessed 06/20)
41. 110th U.S. Congress. Food and Drug Administration Amendments Act of 2007. 2007.
42. Ryan PB, Stang PE, Overhage JM, Suchard MA, Hartzema AG, DuMouchel W, Reich CG, Schuemie MJ, Madigan D. A Comparison of the Empirical Performance of Methods for a Risk Identification System. *Drug Safety* 2013, 36:143-158.
43. DuMouchel W, Ryan PB, Schuemie MJ, Madigan D. Evaluation of Disproportionality Safety Signaling Applied to Healthcare Databases. *Drug Safety* 2013, 36:123-132.
44. Madigan D, Schuemie MJ, Ryan PB. Empirical Performance of the Case–Control Method: Lessons for Developing a Risk Identification and Analysis System. *Drug Safety* 2013, 36:73-82.
45. Norén 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 Safety* 2013, 36:107-121.

46. 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 Safety* 2013, 36:59-72.
47. Ryan PB, Schuemie MJ, Welebob E, Duke J, Valentine S, Hartzema AG. Defining a Reference Set to Support Methodological Research in Drug Safety. *Drug Safety* 2013, 36:33-47.
48. 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 Safety* 2013, 36:83-93.
49. Ryan PB, Schuemie MJ. Evaluating Performance of Risk Identification Methods Through a Large-Scale Simulation of Observational Data. *Drug Safety* 2013, 36:171-180.
50. Schuemie MJ, Madigan D, Ryan PB. Empirical Performance of LGPS and LEOPARD: Lessons for Developing a Risk Identification and Analysis System. *Drug Safety* 2013, 36:133-142.
51. Reich C, van Zandt M, Voss E. OMOP Common Data Model and Standardized Vocabularies. *2nd European OHDSI Symposium* 2019.
52. Overhage JM, Ryan PB, Reich CG, Hartzema AG, Stang PE. Validation of a common data model for active safety surveillance research. *J Am Med Inform Assoc* 2012, 19:54-60.
53. IMEDS. Available at: <http://imeds.reaganudall.org/>. (Accessed 06/20)
54. Hripcsak G, Duke JD, Shah NH, Reich CG, Huser V, Schuemie MJ, Suchard MA, Park RW, Wong IC, Rijnbeek PR, et al. Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers. *Stud Health Technol Inform* 2015, 216:574-578.
55. Platt R, Carnahan RM, Brown JS, Chrischilles E, Curtis LH, Hennessy S, Nelson JC, Racoosin JA, Robb M, Schneeweiss S, et al. The U.S. Food and Drug Administration's Mini-Sentinel program: status and direction. *Pharmacoepidemiol Drug Saf* 2012, 21 Suppl 1:1-8.
56. Curtis LH, Weiner MG, Boudreau DM, Cooper WO, Daniel GW, Nair VP, Raebel MA, Beaulieu NU, Rosofsky R, Woodworth TS, et al. Design considerations, architecture, and use of the Mini-Sentinel distributed data system. *Pharmacoepidemiology and Drug Safety* 2012, 21:23-31.
57. Rassen JA, Schneeweiss S. Using high-dimensional propensity scores to automate confounding control in a distributed medical product safety surveillance system. *Pharmacoepidemiology and Drug Safety* 2012, 21:41-49.
58. Myers JA, Rassen JA, Gagne JJ, Huybrechts KF, Schneeweiss S, Rothman KJ, Joffe MM, Glynn RJ. Effects of Adjusting for Instrumental Variables on Bias and Precision of Effect Estimates. *American Journal of Epidemiology* 2011, 174:1213-1222.
59. Toh S, Reichman ME, Houstoun M, Ross Southworth M, Ding X, Hernandez AF, Levenson M, Li L, McCloskey C, Shoabi A, et al. Comparative risk for angioedema associated with the use of drugs that target the renin-angiotensin-aldosterone system. *Arch Intern Med* 2012, 172:1582-1589.

60. Network AP. Available at: <http://aspennet.asia>. (Accessed 06/20)
61. As PENc, Andersen M, Bergman U, Choi NK, Gerhard T, Huang C, Jalbert J, Kimura M, Kimura T, Kubota K, et al. The Asian Pharmacoepidemiology Network (AsPEN): promoting multi-national collaboration for pharmacoepidemiologic research in Asia. *Pharmacoepidemiol Drug Saf* 2013, 22:700-704.
62. Lai EC-C, Man K, Pratt N, Kubota K. Introduction to Asian PharmacoEpidemiology Network. *10th Asian Conference on Pharmacoepidemiology* 2017. Available at: <http://aspennet.asia/pdf/AsPENSymposiumOutline.pdf>.
63. Pratt N, Andersen M, Bergman U, Choi N-K, Gerhard T, Huang C, Kimura M, Kimura T, Kubota K, Lai EC-C, et al. Multi-country rapid adverse drug event assessment: the Asian Pharmacoepidemiology Network (AsPEN) antipsychotic and acute hyperglycaemia study. *Pharmacoepidemiology and Drug Safety* 2013, 22:915-924.
64. Roughead EE, Chan EW, Choi N-K, Kimura M, Kimura T, Kubota K, Lai EC-C, Man KKC, Nguyen TA, Ooba N, et al. Variation in Association Between Thiazolidinediones and Heart Failure Across Ethnic Groups: Retrospective analysis of Large Healthcare Claims Databases in Six Countries. *Drug Safety* 2015, 38:823-831.
65. Asian Pharmacoepidemiology Network. Surveillance of Health Care in Asian Network (SCAN) Project: Infrastructure Development and analysis with the Asian Pharmacoepidemiology Network. Available at: <http://aspennet.asia/pdf/SCAN.pdf>. (Accessed 06/20)
66. Pratt N, Chan EW, Choi N-K, Kimura M, Kimura T, Kubota K, Lai EC-C, Man KKC, Ooba N, Park B-J, et al. Prescription sequence symmetry analysis: assessing risk, temporality, and consistency for adverse drug reactions across datasets in five countries. *Pharmacoepidemiology and Drug Safety* 2015, 24:858-864.
67. Lai EC-C, Shin J-Y, Kubota K, Man KKC, Park BJ, Pratt N, Roughead EE, Wong ICK, Kao Yang Y-H, Setoguchi S. Comparative safety of NSAIDs for gastrointestinal events in Asia-Pacific populations: A multi-database, international cohort study. *Pharmacoepidemiology and Drug Safety* 2018, 27:1223-1230.
68. Canadian Network for Observational Drug Effect Studies. Available at: <https://www.cnodes.ca>. (Accessed 06/20)
69. Government of Canada. Health Canada. Available at: <https://www.canada.ca/en/health-canada.html>. (Accessed 06/20)
70. Canadian Institutes of Health Research. Available at: <http://www.cihr-irsc.gc.ca>. (Accessed 06/20)
71. Platt RW, Dormuth CR, Chateau D, Filion K. Observational Studies of Drug Safety in Multi-Database Studies: Methodological Challenges and Opportunities. *EGEMS (Wash DC)* 2016, 4:1221.
72. Cameron C, Fireman B, Hutton B, Clifford T, Coyle D, Wells G, Dormuth CR, Platt R, Toh S. Network meta-analysis incorporating randomized controlled

- trials and non-randomized comparative cohort studies for assessing the safety and effectiveness of medical treatments: challenges and opportunities. *Syst Rev* 2015, 4:147.
73. Guertin JR, Rahme E, Dormuth CR, LeLorier J. Head to head comparison of the propensity score and the high-dimensional propensity score matching methods. *BMC Med Res Methodol* 2016, 16:22.
 74. Suissa S, Henry D, Caetano P, Dormuth CR, Ernst P, Hemmelgarn B, Leloirier J, Levy A, Martens PJ, Paterson JM, et al. CNODES: the Canadian Network for Observational Drug Effect Studies. *Open Med* 2012, 6:e134-140.
 75. Filion KB, Chateau D, Targownik LE, Gershon A, Durand M, Tamim H, Teare GF, Ravani P, Ernst P, Dormuth CR. Proton pump inhibitors and the risk of hospitalisation for community-acquired pneumonia: replicated cohort studies with meta-analysis. *Gut* 2014, 63:552-558.
 76. Azoulay L, Filion KB, Platt RW, Dahl M, Dormuth CR, Clemens KK, Durand M, Hu N, Juurlink DN, Paterson JM, et al. Association Between Incretin-Based Drugs and the Risk of Acute Pancreatitis. *JAMA Intern Med* 2016, 176:1464-1473.
 77. Filion KB, Azoulay L, Platt RW, Dahl M, Dormuth CR, Clemens KK, Hu N, Paterson JM, Targownik L, Turin TC, et al. A Multicenter Observational Study of Incretin-based Drugs and Heart Failure. *N Engl J Med* 2016, 374:1145-1154.
 78. Observational Health Data Sciences and Informatics. Available at: <https://www.ohdsi.org/>. (Accessed 06/20)
 79. Observational Health Data Sciences and Informatics. Frequently Asked Questions. Available at: <https://ohdsi.org/who-we-are/frequently-asked-questions/>. (Accessed 06/20)
 80. Observational Health Data Sciences and Informatics. OHDSI Methods Library. Available at: <https://ohdsi.github.io/MethodsLibrary/index.html>. (Accessed 06/20)
 81. Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. Empirical confidence interval calibration for population-level effect estimation studies in observational healthcare data. *Proceedings of the National Academy of Sciences* 2018, 115:2571-2577.
 82. Reps JM, Schuemie MJ, Suchard MA, Ryan PB, Rijnbeek PR. Design and implementation of a standardized framework to generate and evaluate patient-level prediction models using observational healthcare data. *Journal of the American Medical Informatics Association* 2018, 25:969-975.
 83. Schuemie MJ, Ryan PB, DuMouchel W, Suchard MA, Madigan D. Interpreting observational studies: why empirical calibration is needed to correct p-values. *Statistics in Medicine* 2014, 33:209-218.
 84. Observational Health Data Sciences and Informatics. ATLAS. Available at: <http://www.ohdsi.org/web/atlas/>. (Accessed 06/20)
 85. Observational Health Data Sciences and Informatics. OHDSI Forum. Available at: <http://forums.ohdsi.org/>. (Accessed 06/20)

86. Vashisht R, Jung K, Schuler A, Banda JM, Park RW, Jin S, Li L, Dudley JT, Johnson KW, Shervey MM, et al. Association of Hemoglobin A1c Levels With Use of Sulfonylureas, Dipeptidyl Peptidase 4 Inhibitors, and Thiazolidinediones in Patients With Type 2 Diabetes Treated With Metformin: Analysis From the Observational Health Data Sciences and Informatics Initiative. *HbA1c Levels and Second-Line Therapy for Type 2 Diabetes Treated With Metformin*. *JAMA Network Open* 2018, 1:e181755-e181755.
87. Hripcsak G, Ryan PB, Duke JD, Shah NH, Park RW, Huser V, Suchard MA, Schuemie MJ, DeFalco FJ, Perotte A, et al. Characterizing treatment pathways at scale using the OHDSI network. *Proc Natl Acad Sci U S A* 2016, 113:7329-7336.
88. European Health Data & Evidence Network. Available at: <https://www.ehden.eu>. (Accessed 06/20)
89. Innovative Medicines Initiative. Available at: <https://www.imi.europa.eu/>. (Accessed 06/20)
90. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol* 2005, 58:323-337.
91. Pariente A, Gregoire F, Fourrier-Reglat A, Haramburu F, Moore N. Impact of safety alerts on measures of disproportionality in spontaneous reporting databases: the notoriety bias. *Drug Saf* 2007, 30:891-898.
92. Moride Y, Haramburu F, Requejo AA, Begaud B. Under-reporting of adverse drug reactions in general practice. *Br J Clin Pharmacol* 1997, 43:177-181.
93. Hartnell NR, Wilson JP. Replication of the Weber effect using postmarketing adverse event reports voluntarily submitted to the United States Food and Drug Administration. *Pharmacotherapy* 2004, 24:743-749.
94. Zorych I, Madigan D, Ryan P, Bate A. Disproportionality methods for pharmacovigilance in longitudinal observational databases. *Stat Methods Med Res* 2013, 22:39-56.
95. Petri H, de Vet HC, Naus J, Urquhart J. Prescription sequence analysis: a new and fast method for assessing certain adverse reactions of prescription drugs in large populations. *Stat Med* 1988, 7:1171-1175.
96. Hallas J. Evidence of depression provoked by cardiovascular medication: a prescription sequence symmetry analysis. *Epidemiology* 1996, 7:478-484.
97. Wahab IA, Pratt NL, Wiese MD, Kalisch LM, Roughead EE. The validity of sequence symmetry analysis (SSA) for adverse drug reaction signal detection. *Pharmacoepidemiol Drug Saf* 2013, 22:496-502.
98. Brown JS, Kulldorff M, Chan KA, Davis RL, Graham D, Pettus PT, Andrade SE, Raebel MA, Herrinton L, Roblin D, et al. Early detection of adverse drug events within population-based health networks: application of sequential testing methods. *Pharmacoepidemiol Drug Saf* 2007, 16:1275-1284.
99. Kulldorff M, Davis RL, Kolczak† M, Lewis E, Lieu T, Platt R. A Maximized Sequential Probability Ratio Test for Drug and Vaccine Safety Surveillance. *Sequential Analysis* 2011, 30:58-78.

100. Li L. A conditional sequential sampling procedure for drug safety surveillance. *Stat Med* 2009, 28:3124-3138.
101. Murphy S, Castro V, Colecchi J, Dubey A, Gainer V, C H, Sordo M. Partners healthcare OMOP study report. 2011.
102. Jin H, Chen J, Kelman C, He H, McAullay D, O'Keefe CM. Mining Unexpected Associations for Signalling Potential Adverse Drug Reactions from Administrative Health Databases. In. Berlin, Heidelberg: Springer Berlin Heidelberg; 2006.
103. Norén GN, Hopstadius J, Bate A, Star K, Edwards IR. Temporal pattern discovery in longitudinal electronic patient records. *Data Mining and Knowledge Discovery* 2010, 20:361-387.
104. Jin H, Chen J, He H, Kelman C, McAullay D, Keefe CMO. Signaling Potential Adverse Drug Reactions from Administrative Health Databases. *IEEE Transactions on Knowledge and Data Engineering* 2010, 22:839-853.
105. Ji Y, Massanari RM, Ager J, Yen J, Miller RE, Ying H. A fuzzy logic-based computational recognition-primed decision model. *Information Sciences* 2007, 177:4338-4353.
106. Reps J, Garibaldi J, Aickelin U, Soria D, E. Gibson J, Hubbard R. *Comparison of algorithms that detect drug side effects using electronic healthcare database*. Vol. 17; 2013.
107. Reps JM, Garibaldi JM, Aickelin U, Soria D, Gibson JE, Hubbard RB. Signalling paediatric side effects using an ensemble of simple study designs. *Drug Saf* 2014, 37:163-170.
108. Reps JM, Garibaldi JM, Aickelin U, Gibson JE, Hubbard RB. A supervised adverse drug reaction signalling framework imitating Bradford Hill's causality considerations. *J Biomed Inform* 2015, 56:356-368.
109. Kuldorff M, Dashevsky I, Avery TR, Chan AK, Davis RL, Graham D, Platt R, Andrade SE, Boudreau D, Gunter MJ, et al. Drug safety data mining with a tree-based scan statistic. *Pharmacoepidemiol Drug Saf* 2013, 22:517-523.
110. Kuldorff M, Fang Z, Walsh SJ. A tree-based scan statistic for database disease surveillance. *Biometrics* 2003, 59:323-331.
111. Glynn RJ, Gagne JJ, Schneeweiss S. Role of disease risk scores in comparative effectiveness research with emerging therapies. *Pharmacoepidemiol Drug Saf* 2012, 21 Suppl 2:138-147.
112. Arbogast PG, Ray WA. Use of disease risk scores in pharmacoepidemiologic studies. *Stat Methods Med Res* 2009, 18:67-80.
113. Arbogast PG, Ray WA. Performance of disease risk scores, propensity scores, and traditional multivariable outcome regression in the presence of multiple confounders. *Am J Epidemiol* 2011, 174:613-620.
114. Etminan M. Pharmacoepidemiology II: the nested case-control study--a novel approach in pharmacoepidemiologic research. *Pharmacotherapy* 2004, 24:1105-1109.

115. Schneeweiss S, Suissa S. Discussion of Schuemie et al: "A plea to stop using the case-control design in retrospective database studies". *Stat Med* 2019, 38:4209-4212.
116. Schuemie MJ, Ryan PB, Man KKC, Wong ICK, Suchard MA, Hripcsak G. A plea to stop using the case-control design in retrospective database studies. *Stat Med* 2019, 38:4199-4208.
117. Moore N, Gulmez SE, Blin P, Lassalle R, Jove J, Theophile H, Begaud B, Larrey D, Benichou J. Relative risks from case-population data. *Epidemiology* 2013, 24:935-936.
118. Theophile H, Laporte JR, Moore N, Martin KL, Begaud B. The case-population study design: an analysis of its application in pharmacovigilance. *Drug safety : an international journal of medical toxicology and drug experience* 2011, 34:861-868.
119. Capella D, Pedros C, Vidal X, Laporte JR. Case-population studies in pharmacoepidemiology. *Drug safety : an international journal of medical toxicology and drug experience* 2002, 25:7-19.
120. van der Klauw MM, Stricker BH, Herings RM, Cost WS, Valkenburg HA, Wilson JH. A population based case-cohort study of drug-induced anaphylaxis. *Br J Clin Pharmacol* 1993, 35:400-408.
121. Moore N, Gulmez SE, Larrey D, Pageaux GP, Lignot S, Lassalle R, Jove J, Pariente A, Blin P, Benichou J, 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, 22:160-167.
122. 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. *Eur J Clin Pharmacol* 2000, 56:763-767.
123. Keisu M, Ekman E, Wiholm BE. Comparing risk estimates of sulphonamide-induced agranulocytosis from the Swedish Drug Monitoring System and a case-control study. *Eur J Clin Pharmacol* 1992, 43:211-214.
124. Laporte JR, Ibanez L, Ballarin E, Perez E, Vidal X. Fatal aplastic anaemia associated with nifedipine. *Lancet* 1998, 352:619-620.
125. 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, 36:1069-1078.
126. Schuemie MJ, Ryan P, Shaddox T, Suchard MA. SelfControlledCaseSeries: Self-Controlled Case Series. R package version 1.3.1. 2018. Available at: <https://github.com/OHDSI/SelfControlledCaseSeries>.
127. 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, 36 Suppl 1:S95-106.
128. Schuemie MJ. CaseCrossover: Case-Crossover. R package version 1.1.0. 2018. Available at: <https://github.com/OHDSI/CaseCrossover>.

129. Agence nationale de sécurité du médicament et des produits de santé. Available at: <https://ansm.sante.fr/>. (Accessed 06/23)
130. Moore N, Biour M, Paux G, Loupi E, Bégaud B, Boismare F, Royer RJ. Adverse drug reaction monitoring: doing it the French way. *Lancet* 1985, 2:1056-1058.
131. Vial T. Pharmacovigilance française : missions, organisation et perspectives. *Therapies* 2016, 71:135-142.
132. Arambourou S, Gardette C, Karvar A. Audit d'organisation de l'ANSM. 2015.
133. DRUGS-SAFE. Available at: <https://drugssafe.fr>. (Accessed 06/23)
134. Agence nationale de sécurité du médicament et des produits de santé. L'ANSM soutient deux plateformes en épidémiologie des produits de santé - Point d'information. 2014. Available at: <https://www.ansm.sante.fr/S-informer/Points-d-information-Points-d-information/L-ANSM-soutient-deux-plateformes-en-epidemiologie-des-produits-de-sante-Point-d-information>.
135. Agence nationale de sécurité du médicament et des produits de santé. L'ANSM et la Cnam s'associent pour créer Epi-Phare, une structure d'expertise publique en épidémiologie des produits de santé. 2019. Available at: <https://www.ansm.sante.fr/S-informer/Communiqués-Communiqués-Points-presse/L-ANSM-et-la-Cnam-s-associent-pour-créer-Epi-Phare-une-structure-d-expertise-publique-en-epidemiologie-des-produits-de-sante-Communiqué>.
136. Agence nationale de sécurité du médicament et des produits de santé. EPI-PHARE. Available at: [tps://www.ansm.sante.fr/L-ANSM/EPI-PHARE](https://www.ansm.sante.fr/L-ANSM/EPI-PHARE). (Accessed 06/23)
137. Agence nationale de sécurité du médicament et des produits de santé. EPI-PHARE s'appuiera sur un Centre partenaire en épidémiologie des produits de santé. 2019. Available at: <https://www.ansm.sante.fr/S-informer/Actualité/EPI-PHARE-s-appuiera-sur-un-Centre-partenaire-en-epidemiologie-des-produits-de-sante-Point-d-information>.
138. Agence nationale de sécurité du médicament et des produits de santé. Bonnes pratiques de pharmacovigilance. 2018.
139. Miremont-Salamé G, Théophile H, Haramburu F, Bégaud B. Causality assessment in pharmacovigilance: The French method and its successive updates. *Therapies* 2016, 71:179-186.
140. Bezin J, Duong M, Lassalle R, Droz C, Pariente A, Blin P, Moore N. The national healthcare system claims databases in France, SNIIRAM and EGB: Powerful tools for pharmacoepidemiology. *Pharmacoepidemiol Drug Saf* 2017, 26:954-962.
141. Santé SNdDd. Composantes du SNDS. Available at: <https://www.snds.gouv.fr/SNDS/Composantes-du-SNDS>. (Accessed 06/24)
142. 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, 58:286-290.

143. Système National des Données de Santé. Référentiel de Sécurité, Guide d'accompagnement. Available at: https://www.snds.gouv.fr/download/Guide_accompagnement.pdf.
144. Becher H, Kostev K, Schroder-Bernhardi D. Validity and representativeness of the "Disease Analyzer" patient database for use in pharmacoepidemiological and pharmaco-economic studies. *Int J Clin Pharmacol Ther* 2009, 47:617-626.
145. Jouaville SL, Miotti H, Coffin G, Sarfati B, Meilhoc A. Validité et limites de la base de données Longitudinal Patient Database France pour les études pharmacoépidémiologiques et pharmacoéconomiques. *Revue d'Épidémiologie et de Santé Publique* 2016, 64:S280-S281.
146. Bégaud B, Costagliola D. Rapport sur la surveillance et la promotion du bon usage du médicament en France. 2013.
147. Zapletal E, Rodon N, Grabar N, Degoulet P. Methodology of integration of a clinical data warehouse with a clinical information system: the HEGP case. *Stud Health Technol Inform* 2010, 160:193-197.
148. Murphy SN, Mendis M, Hackett K, Kuttan R, Pan W, Phillips LC, Gainer V, Berkowicz D, Glaser JP, Kohane I, et al. Architecture of the open-source clinical research chart from Informatics for Integrating Biology and the Bedside. *AMIA Annu Symp Proc* 2007:548-552.
149. Assistance Publique - Hopitaux de Paris. L'Entrepôt de Données de Santé. Available at: <https://recherche.aphp.fr/eds/>.
150. Salamanca E. AP-HP Data Analytics Platform Improving outcomes and use of routinely collected health data *HIMSS Europe 2018 & Health 2.0 Conference* 2018.
151. Malafaye N, Demoulin D, Mailhe P, Morell M, Pellecier D, Dunoyer C. Mise en place et exploitation d'un entrepôt de données au département d'information médicale du CHU de Montpellier, France. *Revue d'Épidémiologie et de Santé Publique* 2018, 66:S26.
152. DuMouchel W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. *The American Statistician* 1999, 53:177-190.
153. Agence nationale de sécurité du médicament et des produits de santé. Détection automatisée et enquête de pharmacovigilance. *vigilances* 2016, Page 1.
154. Dias P, Penedones A, Alves C, Ribeiro CF, Marques FB. The Role of Disproportionality Analysis of Pharmacovigilance Databases in Safety Regulatory Actions: a Systematic Review. *Curr Drug Saf* 2015, 10:234-250.
155. Montastruc JL, Sommet A, Bagheri H, Lapeyre-Mestre M. Benefits and strengths of the disproportionality analysis for identification of adverse drug reactions in a pharmacovigilance database. *Br J Clin Pharmacol* 2011, 72:905-908.
156. Arnaud M. Developing an automated system combining detection and prioritization safety signal detection from healthcare databases. 2017. Available at: <https://tel.archives-ouvertes.fr/tel-01661585>.

157. Arnaud M, Begaud B, Thiessard F, Jarrion Q, Bezin J, Pariente A, Salvo F. An Automated System Combining Safety Signal Detection and Prioritization from Healthcare Databases: A Pilot Study. *Drug Saf* 2018, 41:377-387.
158. Bacry E, Gaïffas S. Machine learning and big data in health: the partnership between Caisse Nationale d'Assurance Maladie and Ecole polytechnique. In: Editions C, ed. *Santé et intelligence artificielle*; 2018.
159. Morel M, Bacry E, Gaïffas S, Guilloux A, Leroy F. ConvSCCS: convolutional self-controlled case series model for lagged adverse event detection. *Biostatistics* 2019.
160. Bezin J, Mansiaux Y, Noize P, Salvo F, Bégaud B, Pariente A. Use of Lipid-Lowering Drugs and the Risk of Cataract: A Population-Based Nested Case-Control Study. *Clinical Pharmacology & Therapeutics* 2019, 105:458-465.
161. Maumus-Robert S, Bérard X, Mansiaux Y, Tubert-Bitter P, Debette S, Pariente A. Short-Term Risk of Aortoiliac Aneurysm or Dissection Associated With Fluoroquinolone Use. *Journal of the American College of Cardiology* 2019, 73:875-877.
162. Arnaud M, Pariente A, Bezin J, Bégaud B, Salvo F. Risk of Serious Trauma with Glucose-Lowering Drugs in Older Persons: A Nested Case–Control Study. *Journal of the American Geriatrics Society* 2018, 66:2086-2091.
163. European Medicines Agency. Guidance on the format of the risk management plan (RMP) in the EU 2018. Available at: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guidance-format-risk-management-plan-rmp-eu-integrated-format-rev-201_en.pdf.
164. Madigan D, Ryan PB, Schuemie M, Stang PE, Overhage JM, Hartzema AG, Suchard MA, DuMouchel W, Berlin JA. Evaluating the impact of database heterogeneity on observational study results. *Am J Epidemiol* 2013, 178:645-651.
165. Schuemie MJ. CaseControl: Case-Control. R package version 1.6.0. 2019. Available at: <https://github.com/OHDSI/CaseControl>.
166. European Medicines Agency. The European Union electronic Register of Post-Authorisation Studies (EU PAS Register). Available at: http://www.encepp.eu/encepp_studies/indexRegister.shtml. (Accessed 06/25)
167. European Medicines Agency. The ENCePP Seal. Available at: http://www.encepp.eu/encepp_studies/index.shtml. (Accessed 06/25)
168. European Medicines Agency. The ENCePP Code of Conduct for Scientific Independence and Transparency in the Conduct of Pharmacoepidemiological and Pharmacovigilance Studies. 2018. Available at: http://www.encepp.eu/code_of_conduct/documents/ENCePPCodeofConduct.pdf.
169. Armstrong B. A simple estimator of minimum detectable relative risk, sample size, or power in cohort studies. *Am J Epidemiol* 1987, 126:356-358.
170. Musonda P, Paddy Farrington C, Whitaker HJ. Sample sizes for self-controlled case series studies. *Statistics in Medicine* 2006, 25:2618-2631.

171. Edwardes MD. Sample size requirements for case-control study designs. *BMC Med Res Methodol* 2001, 1:11.
172. Whitaker H. Advances in self-controlled case series method. *32nd ICPE* 2016.
173. Sackett DL. Bias in analytic research. *Journal of Chronic Diseases* 1979, 32:51-63.
174. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, Müller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011, 12:77.
175. Schuemie M. MethodEvaluation: Package for evaluation of estimation methods. R package version 1.0.2. 2019. Available at: <https://github.com/OHDSI/MethodEvaluation>.
176. Schuemie M. Running the OHDSI Methods Benchmark. Available at: <https://ohdsi.github.io/MethodEvaluation/articles/OhdsiMethodsBenchmark.html>. (Accessed 06/27)
177. Inc.; AWS. Available at: <https://aws.amazon.com/>.
178. Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. Robust empirical calibration of p-values using observational data. *Stat Med* 2016, 35:3883-3888.
179. Gruber S, Tchetgen Tchetgen E. Limitations of empirical calibration of p-values using observational data. *Statistics in Medicine* 2016, 35:3869-3882.
180. Reich C, Ryan P, Belenkaya R, Natarajan K, Blacketer C. OMOP Common Data Model v5.1 Specifications. 2017. Available at: <https://github.com/OHDSI/CommonDataModel/releases/tag/v5.1.0>.
181. Article L5125-23. *Code de la santé publique* 2019.
182. l'Assurance Maladie. Medic'AM. Available at: <https://www.ameli.fr/l-assurance-maladie/statistiques-et-publications/donnees-statistiques/medicament/medic-am/medic-am-mensuel-2018.php>. (Accessed 06/26)
183. Thurin N, Lassalle R, Blin P, Pénichon M, Schuemie M, Gagne J, Rassen J, Benichou J, Weill A, Droz-Perroteau C, et al. Alert generation using the case-population approach in the French claims databases (ALCAPONE). *33rd International Conference on Pharmacoepidemiology and Therapeutic Risk Management* 2017. Available at: http://www.bordeauxpharmacoepi.eu/congres/icpe_2017/alcapone.pdf.
184. Direction de la Recherche des Etudes de l'Evaluation et des Statistiques. Visualisation de la structure du SNDS. Available at: <http://dico-snds.health-data-hub.fr/>. (Accessed 06/28)
185. Thurin N, Lassalle R, Jové J, Sylvestre R, Dupont J, Lacueille C, Blin P, Moore N, Droz-Perroteau C. Standardisation de l'utilisation des données du SNDS à des fins de recherche médicale : présentation d'un modèle de données optimisé centré sur le patient. *Journées EMOIS 2019* 2019. Available at: <http://emois.org/wp-content/uploads/2019/04/B2-session-MethodologieDuTraitementDeLinformation.zip>.
186. Thurin N, Lassalle R, Blin P, Pénichon M, Schuemie M, Gagne J, Rassen J, Benichou J, Weill A, Droz-Perroteau C, et al. Empirical assessment of OHDSI

- case-based methods and case-population for the identification of drug-related outcome in the French nationwide healthcare database (SNDS) 2019 *European OHDSI Symposium 2019*. Available at: <https://www.ohdsi-europe.org/images/symposium-2019/presentations/7-Lightning.pdf>.
187. Thurin N, Lassalle R, Blin P, Pénichon M, Schuemie M, Gagne J, Rassen J, Benichou J, Weill A, Droz-Perroteau C, et al. Empirical Assessment Of Case-based Methods For The Identification Of Drug-related Health Outcome Of Interest In The French Nationwide Healthcare Database (sn ds). *35th International Conference on Pharmacoepidemiology and Therapeutic Risk Management 2019*. Available at: http://www.bordeauxpharmacoepi.eu/congres/icpe_2019/empiricalAssessment.pdf.
 188. Farrington CP, Anaya-Izquierdo K, Whitaker HJ, Hocine MN, Douglas I, Smeeth L. Self-Controlled Case Series Analysis With Event-Dependent Observation Periods. *Journal of the American Statistical Association* 2011, 106:417-426.
 189. Observational Health Data Sciences and Informatics. *The Book of OHDSI*; 2019.
 190. Centre Hospitalier Haguenau. Projet d'organisation du codage PMSI. 2018. Available at: <http://www.departement-information-medicale.com/wp-content/uploads/2018/06/organisation-codage.pdf>.
 191. Marescaux C. Entre soin et contrôle de gestion : place du DIM dans l'organisation hospitalière. *L'information psychiatrique* 2011, 87:487-491.
 192. Caeyseele T, Bruandet A, Delaby F, Theis D. Création d'un outil de gestion des contrôles qualités du codage au DIM du CHRU de Lille. *Revue d'Épidémiologie et de Santé Publique* 2016, 64:S20.
 193. Bosco-Levy P, Duret S, Gilleron V, Chatellier G, Looten V, Moore N. Étude de validation des codes diagnostic de l'insuffisance cardiaque. *Revue d'Épidémiologie et de Santé Publique* 2018, 66:S22.
 194. Quantin C, Benzenine É, Hägi M, Auverlot B, Abrahamowicz M, Cottenet J, Fournier É, Biquet C, Compain D, Monnet É, et al. Évaluation du PMSI comme moyen d'identification des cas incidents de cancer colorectal. *Santé Publique* 2014, 26:55-63.
 195. Thurin N, Rouyer M, Gross-Goupil M, Soulié M, Roumiguié M, Le Moulec S, De Beaucoudrey L, Droz-Perroteau C, Blin P. Validation d'un algorithme complexe dans le SNDS. Exemple du cancer de la prostate résistant à la castration et métastatique. *Revue d'Épidémiologie et de Santé Publique* 2019, 67:S75.
 196. Pirmohamed M, James S, Meakin S, Green C, Scott AK, Walley TJ, Farrar K, Park BK, Breckenridge AM. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. *Bmj* 2004, 329:15-19.

APPENDICES

Appendix 1 Communications related to the thesis

Thurin N, Lassalle R, Blin P, Pénichon M, Schuemie M, Gagne J, Rassen J, Benichou J, Weill A, Droz-Perroteau C, et al. Empirical Assessment Of Case-based Methods For The Identification Of Drug-related Health Outcome Of Interest In The French Nationwide Healthcare Database (SNDS). 35th International Conference on Pharmacoepidemiology and Therapeutic Risk Management 2019. Available at: http://www.bordeauxpharmacoepi.eu/congres/2019/ICPE/BPE_ALCAPONE_ICPE_2019.pdf.

Thurin N, Lassalle R, Blin P, Pénichon M, Schuemie M, Gagne J, Rassen J, Benichou J, Weill A, Droz-Perroteau C, et al. Empirical assessment of OHDSI case-based methods and case-population for the identification of drug-related outcome in the French nationwide healthcare database (SNDS) 2ème Congrès « Le SNDS pour la recherche en santé » 2019. Available at: http://www.bordeauxpharmacoepi.eu/congres/2019/OHDSI/BPE_ALCAPONE_Poster_OHDSI_2019.pdf.

Thurin N, Lassalle R, Blin P, Pénichon M, Schuemie M, Gagne J, Rassen J, Benichou J, Weill A, Droz-Perroteau C, et al. Empirical assessment of case-based methods for the identification of drug-related Acute Liver Injury (ALI) in the French nationwide healthcare database (SNDS). 2019 ISPE Mid-Year Meeting 2019. Available at: http://www.bordeauxpharmacoepi.eu/congres/2019/ISPE-MIDYEAR/BPE_ALCAPONE_ISPE-MIDYEAR_2019.pdf.

Thurin N, Lassalle R, Blin P, Pénichon M, Schuemie M, Gagne J, Rassen J, Benichou J, Weill A, Droz-Perroteau C, et al. Empirical assessment of OHDSI case-based methods and case-population for the identification of drug-related outcome in the French nationwide healthcare database (SNDS) 2019 European OHDSI Symposium 2019. Available at: <https://www.ohdsi-europe.org/images/symposium-2019/presentations/7-Lightning.pdf>.

Thurin N, Lassalle R, Blin P, Pénichon M, Schuemie M, Gagne J, Rassen J, Benichou J, Weill A, Droz-Perroteau C, et al. Alert generation using the case-population approach in the French Nationwide Healthcare Database (SNDS): the ALCAPONE project. 2018 European OHDSI Symposium 2018. Available at: <https://www.ohdsi-europe.org/images/symposium-2018/posters/23-Nicolas-Thurin.pdf>.

Thurin N, Pénichon M, Lassalle R, Blin P, Droz-Perroteau C, Schuemie M, Moore N. Mapping the 66.6M-patient French Nationwide Healthcare Database (SNDS) to the OMOP CDM v5: a feasibility study. 2018 European OHDSI Symposium 2018. Available at: <https://www.ohdsi-europe.org/images/symposium-2018/posters/22-Nicolas-Thurin.pdf>.

Thurin N, Lassalle R, Blin P, Schuemie M, Droz-Perroteau C, Moore N. Évaluation de l'approche cas-population dans la base du SNDS en vue de la génération de signaux de pharmacovigilance. VIème Congrès National Conjoint ADELFI / EMOIS 2018. Available at: http://www.bordeauxpharmacoepi.eu/congres/2018/EMOIS/BPE_ALCAPONE_EMOIS_2018.pdf.

Thurin N, Lassalle R, Blin P, Pénichon M, Schuemie M, Gagne J, Rassen J, Benichou J, Weill A, Droz-Perroteau C, et al. Alert generation using the case-

population approach in the French claims databases (ALCAPONE). 33rd International Conference on Pharmacoepidemiology and Therapeutic Risk Management 2017. Available at: http://www.bordeauxpharmacoepi.eu/congres/2017/ICPE/BPE_ALCAPONE_ICPE_2017.pdf.

Thurin N, Lassalle R, Blin P, Pénichon M, Schuemie M, Gagne J, Rassen J, Benichou J, Weill A, Droz-Perroteau C, et al. Validation et calibration de l'approche cas-population dans le SNIIRAM en vue de la génération de signaux de pharmacovigilance : projet ALCAPONE. 9ème Colloque Données de Santé en vie réelle 2017. Available at : http://www.bordeauxpharmacoepi.eu/congres/2017/DSVR/BPE_ALCAPONE_DSVR_2017.pdf.

Appendix 2 Award letter for ENCePP Seal



EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH



European Network of Centres for
Pharmacoepidemiology and
Pharmacovigilance

Prof. Nicolas More
Bordeaux PharmacoEpi (BPE)
University of Bordeaux
146 Rue Leo Saignat
33076 Bordeaux
France

7 April 2016
EMA/253543/2016
ENCEPP Secretariat

RE: Your ENCePP Seal application for "Alert generation using the case-population approach in the French claims databases (ALCAPONE)"

Study Reference Number: ENCEPP/SDPP/13031

Dear Prof. Moore,

On behalf of the ENCePP Steering Group I am writing to you to acknowledge fulfilment of the conditions for your study to qualify for the ENCePP Seal. We have received all required documentation¹ and the study has been included in the EU PAS Register prior to its start.

I am, therefore, pleased to confirm that your study has been awarded *a priori* the ENCePP Seal which will be displayed with the study entry in the EU PAS Register.

I would like to remind you of the additional obligations the ENCePP Seal places on the conduct of your study, namely to adhere to the entirety of the provisions of the ENCePP Code of Conduct throughout the research process. In particular, you are obliged to:

- Update the entry of the study in the EU PAS Register, as milestones are reached or in case of changes to the protocol that may affect the interpretation of the study;
- Update the Checklist for Study Protocols and re-submit it via the EU PAS Register in case of changes to the protocol that may affect the interpretation of the study;
- After finalisation of the study report, provide the final version of the protocol through the EU PAS Register. Both the original and the final versions of the protocol will be made publicly available at that stage.
- Make publicly available a clear summary of the main results of the study, whether positive or negative and including results from prematurely terminated studies according to the timetable agreed in the research contract. In addition, an abstract (in English) of the study results shall be provided, via the EU PAS Register, for publication within 3 months following the final study report.

¹ Signed copy of Checklist of the ENCePP Code of Conduct, Declaration on compliance with the ENCePP Code of Conduct, ENCePP Checklist for Study Protocols and Declaration of Interests.



Appendix 3 Décision d'approbation n°187 de l'Institut des Données de Santé du 25 mai 2016



Décision d'approbation n°187 de l'Institut des Données de Santé du 25 mai 2016

- Vu la demande de l'INSERM concernant une étude dénommée « Génération d'alertes en bases de données avec la méthode du Cas-population - Projet ALCAPONE »
- Vu l'évaluation scientifique dont a fait l'objet cette étude
- Vu l'arrêté mentionné à l'article L. 161-28-1 du code de la sécurité sociale
- Vu le règlement intérieur de l'Institut des Données de Santé

Le Bureau de l'IDS constate :

- Que la demande de l'INSERM est recevable au titre de l'arrêté mentionné à l'article L.161-28-1 du code de la sécurité sociale ;
- Que l'INSERM s'engage à respecter la charte de déontologie de l'IDS.

Le Bureau de l'IDS approuve la demande de l'INSERM d'extraction d'un échantillon de données du SNIIRAM (ambulatoires et hospitalières), portant sur les années 2009 à 2014, dans le cadre de l'étude citée ci-dessus. Il demande que les résultats de cette étude, qui revêt un intérêt pour un ensemble d'acteurs de la santé, soient en outre communiqués à l'IDS.

Cette approbation, conformément à l'arrêté mentionné à l'article L.161-28-1 du code de la sécurité sociale, sera transmise par l'IDS, à la CNIL.

La Présidente de l'Institut des Données de Santé

Par délégué, le Directeur
Chantal Cases

Richard DECOTTIGNIES

Appendix 4 Notification d'autorisation CNIL pour le projet ALCAPONE



Le Vice-Président délégué

Monsieur Yves LEVY
Président Directeur Général
INSTITUT NATIONAL DE LA SANTÉ ET LA
RECHERCHE MÉDICALE
INSERM
8 RUE DE LA CROIX JARRY
75013 - PARIS

Paris, le 13 MARS 2017

N/Réf. : MMS/OTB/AE171045

Objet : NOTIFICATION D'AUTORISATION

Décision DE-2017-040 autorisant l'INSTITUT NATIONAL DE LA SANTÉ ET LA RECHERCHE MÉDICALE à mettre en œuvre un traitement de données de santé à caractère personnel portant sur l'étude dénommée « Génération d'alertes en bases des données avec la méthode Cas-population – Projet ALCAPONE » aux fins d'évaluer la pertinence de l'utilisation des bases de données de remboursement françaises (le SNIIRAM et l'EGB) dans le cadre de la génération de signaux de pharmacovigilance, à partir de la méthodologie OMOP et de l'approche cas population. (Demande d'autorisation n° 1970983)

Monsieur le Président Directeur Général,

Vous avez saisi notre Commission d'une demande d'autorisation relative à un traitement de données à caractère personnel ayant pour finalité :

ÉVALUER LA PERTINENCE DE L'UTILISATION DES BASES DE DONNÉES DE REMBOURSEMENT FRANÇAISES (LE SNIIRAM ET L'EGB) DANS LE CADRE DE LA GÉNÉRATION DE SIGNAUX DE PHARMACOVIGILANCE, À PARTIR DE LA MÉTHODOLOGIE OMOP ET DE L'APPROCHE CAS POPULATION.

Afin de réaliser cette étude, vous souhaitez obtenir l'extraction d'un échantillon de données ambulatoires et hospitalières à partir du Système National d'informations inter-régimes de l'Assurance Maladie (SNIIRAM).

Ce traitement relève de la procédure des articles 62 et suivants de la loi du 6 janvier 1978 dans sa rédaction antérieure à la loi n°2016-41 du 26 janvier 2016 de modernisation de notre système de santé, en vigueur au moment de la transmission de la demande.

Je prends acte de ce que l'Institut des données de santé (IDS) a approuvé votre demande d'extraction pour une durée totale de cinq ans (2009-2014) et que les résultats de cette étude, qui revêtent un intérêt pour un ensemble d'acteurs de la santé, soient en outre communiqués à l'IDS.

RÉPUBLIQUE FRANÇAISE

3 Place de Fontenoy, TSA 80715 - 75334 PARIS CEDEX 07 - 01 53 73 22 22 - www.cnil.fr

Les données à caractère personnel nécessaires à l'accomplissement des missions de la CNIL, définies par la loi sont conservées dans des fichiers réservés à son usage exclusif.
Les personnes concernées peuvent exercer leurs droits d'accès, de rectification et d'opposition en s'adressant au correspondant informatique et Libertés (CIL) de la CNIL via le formulaire prévu à cet effet sur cnil.fr ou par courrier postal.

En application de l'article 4-I de l'arrêté du 19 juillet 2013 modifié, relatif à la mise en œuvre du SNIIRAM « Pour tout traitement de données individuelles relatives aux bénéficiaires de l'assurance maladie, seuls sont autorisés à effectuer des recherches mettant en œuvre simultanément plus d'une des quatre variables sensibles (code commune, date des soins, mois et année de naissance, date de décès) avec d'autres données :

- les médecins-conseils de l'assurance maladie et les personnels habilités placés sous leur responsabilité, nommément désignés par les médecins responsables selon l'organisation des régimes d'assurance maladie
- à titre expérimental et pour une durée de trois ans, les médecins de l'Institut de veille sanitaire et les personnels habilités placés sous leur responsabilité, nommément désignés par le directeur de l'Institut de veille sanitaire, autorité médicale de l'Institut, dans la mesure où ces variables sont strictement nécessaires à l'accomplissement des missions de surveillance et d'alerte sanitaires.»

Vous indiquez que des mesures de sécurité physique et logique seront mises en place pour garantir la confidentialité des données et que le traitement informatique des données sera réalisé sous votre responsabilité et celle de vos collaborateurs.

J'attire votre attention sur les obligations qui incombent désormais à ces personnes qui doivent :

- n'utiliser les fichiers qu'à des fins d'analyse comparative de l'activité de soins,
- respecter et faire respecter le secret des informations cédées par toutes les personnes susceptibles de travailler sur ces données, ces personnes étant astreintes par écrit au secret professionnel,
- prendre toutes précautions utiles afin de préserver la sécurité des informations ainsi transmises et notamment empêcher qu'elles ne soient déformées, endommagées ou communiquées à des tiers non autorisés,
- ne pas rétrocéder ou divulguer à des tiers les informations fournies sous quelque forme que ce soit,
- ne pas procéder à des rapprochements, interconnexions, mises en relation, appariements avec tout fichier de données directement ou indirectement nominatives ou toute information susceptible de révéler l'identité d'une personne et/ou son état de santé,
- ne pas utiliser de façon détournée les informations transmises, notamment à des fins de recherche ou d'identification des personnes.

En outre, le responsable du projet devra s'engager à ce que les informations tirées des exploitations de fichiers et susceptibles d'être diffusées se présentent uniquement sous la forme de statistiques agrégées de telle sorte que les personnes concernées ne puissent pas être identifiées.

La durée de conservation relative aux catégories de données est fixée à cinq ans.

En application des articles 15 et 69 de la loi précitée et de la délibération n° 2009-674 du 26 novembre 2009 portant délégation d'attributions de la Commission nationale de l'informatique et des libertés à son président et à son vice-président délégué, j'autorise la mise en œuvre de ce traitement.

Je vous prie d'agréer, Monsieur le Président Directeur Général, l'expression de mes salutations distinguées.

Marie-France MAZARS



Mapping the 66.6M-patient French Nationwide Healthcare Database (SNDS) to the OMOP CDM v5: a feasibility study

Nicolas Thurin^{1,2}, Marine Pénichon¹, Régis Lassalle¹, Patrick Blin¹, Cécile Droz-Perroteau¹, Martijn Schuemie³, Nicholas Moore^{1,2}

¹Bordeaux PharmacoEpi, INSERM CIC1401, Université de Bordeaux, Bordeaux, France; ²INSERM U1219, France; ³Janssen R&D, Titusville, New Jersey, USA



Background

- SNDS (Système National des Données de Santé) is the French Nationwide Healthcare System Database
 - Covers about 99% of the French population (about 66.6 million persons) from birth (or immigration) to death (or emigration)
 - Merges information of reimbursed claims (SNIIRAM) linked to the national hospital-discharge summary (PMSI) and the national death registry (CépiDC)
 - Includes: demographics; physician or paramedical visits (without medical indication); dispensed drugs and medical devices; lab tests (without results); chronic medical conditions; hospitalization with ICD-10 codes for primary, linked and associated diagnoses, date and duration; reimbursed procedures
- EGB (Echantillon Généraliste de Bénéficiaires) is the 1/97th random permanent representative sample of SNDS.
 - 810 000 persons
 - Planned 20-year longitudinal data
 - Same structure than SNDS
- EGB is an ideal candidate for assessing the possibility to apply the OMOP CDM to the SNDS

Methods

- Extraction of an EGB sample (extraction of all the myocardial infarctions between 2009 and 2014)
- Data management of the 74 resulting datasets to 14 transitional ones
 - Socio demographic data
 - Medical visits
 - Chronic medical conditions
 - Hospitalization diagnoses
 - Drug dispensings
 - Lab tests
 - Outpatient procedures
 - Inpatient procedures
- ETL design
 - Theoretical mapping of the variables from EGB to OMOP CDM v5 (Figure 1)
 - Assessment of the mapping from source codes and thesauri to OMOP standardized vocabulary (Table 1.)
- Effective mapping of data
 - Variables mapping
 - Generation of the treatment duration for the drug_exposure table according to drug dispensings data
 - Codes mapping whenever present in OMOP vocabulary

Results

ETL design

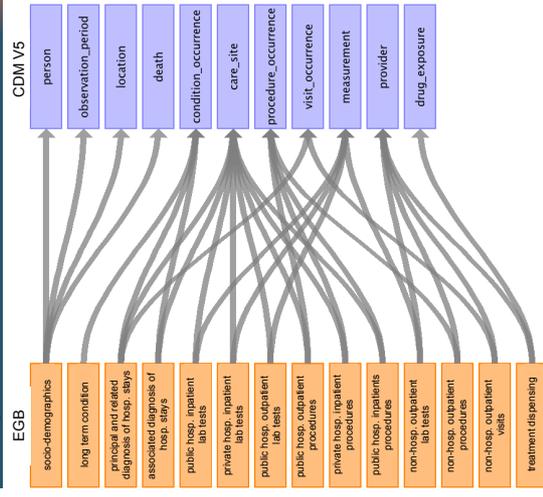


Figure 1. EGB data Mapping Approach to CDM v5 (clinical part only, without devices)

Effective mapping of data

- To date, 4 data tables have been generated
 - person
 - drug_exposure
 - location
 - observation_period
- Drug treatment duration for the drug_exposure table has been generated using medians of treatment duration and taking into account the 3-month drug boxes.

Conclusions

- Most of the standardized clinical data tables can be generated following OHDSI Extract Transform and Load (ETL) process. However the mapping of the French terminologies for procedures, lab tests, drugs and devices to OMOP Standard Vocabularies is required to complete fully the transformation of the clinical part of EGB and SNDS and enable the use of the whole set of OHDSI tools.

The theoretical mapping of the 14 transitional tables corresponding to the clinical part of the EGB has been realized through WhiteRabbit and RabbitFlat softwares (Figure 1.)

Table 1. Presence of EGB and SNDS vocabularies in the OMOP vocabularies

Domains	EGB / SNDS vocabularies	Presence in the OMOP vocabularies
Condition	ICD10	Yes
Drug	ATC	Yes
	CIP	No
	UCD	No
Device	LPP	No
	CCAM	No
Procedure	NABM	No
Measurement		No

ATC: Anatomical Therapeutic Chemical; CCAM: Classification Commune des Actes Médicaux; CIP: Code Identifiants de Présentations; ICD: International Classification of Diseases; LPP: Liste des Produits et Préparations; NABM: Nomenclature des Actes de Biologie Médicale; UCD: Unité Commune de Dispensation

Most of the national nomenclatures used in France are not included in the OMOP vocabularies

- CIP is used to identify drug boxes dispensed in community pharmacies
- UCD is used to identify expensive drugs dispensed in the hospitals
- LPP is mainly used to code reimbursed medical devices
- CCAM is used to code outpatient and inpatient procedures
- NABM is used to code inpatient and outpatient lab tests

- Other 5 data tables have not been created yet but do not seem to present technical difficulties
 - death
 - provider
 - condition_occurrence
 - visit_occurrence
 - care_site
- Generation of procedure_occurrence and measurements tables requires the mapping of CCAM and NABM vocabularies to standard concepts



Scan me to download this poster

Appendix 6 Poster displayed at the 33rd ICPE, August 2017, Montréal, QC, Canada



Pharmacologie médicale
Bordeaux PharmacoEpi
CIC Bordeaux CIC1401







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

N. Thurin^{1,2,3}, R. Lassalle¹, P. Blin¹, M. Pénichon¹, M. Schuemie⁴, J. J. Gagne⁵, J. Rassen⁶, J. Benichou^{7,2}, A. Weill⁸, C. Droz-Perroteau¹, N. Moore^{1,2,3}

¹Bordeaux PharmacoEpi, INSERM CIC1401, Université de Bordeaux, Bordeaux, France – ²INSERM U1219, Bordeaux, France – ³CHU de Bordeaux, Bordeaux, France – ⁴Janssen Research & Development, Observational Health Data Sciences and Informatics (OHDSI), New York, NY, USA – ⁵Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA – ⁶AstraZenca, New York, NY, USA – ⁷CHU de Rouen, INSERM U1219, Rouen, France – ⁸Caisse Nationale de l'Assurance Maladie des Travailleurs Salariés, Paris, France

Abstract

Background: France has a nationwide healthcare insurance system database – the SNIRAM (Système national d'information inter-régimes de l'Assurance maladie) – that covers about 99% of the French population. A 1/97th sample – the EGB (Echantillon généraliste de bénéficiaires) – is also available. SNIRAM has not been tested for drug safety alert generation. **Objectives:** To present the methodology and assess the feasibility of the ALCAPONE project. **Methods:** ALCAPONE is based on historical data from the SNIRAM, and the OMOP reference set which consists of 4 main outcomes – Acute Liver Injury (ALI), Myocardial Infarction (MI), Acute Kidney Injury (KI), and Upper Gastrointestinal Bleeding (UGIB) – and 165 positive and 235 negative drug controls. ALCAPONE consists of 3 main stages: (i) selection of detectable positive and negative controls (i.e. with a minimum detectable relative risk ≥ 1.25) through the realization of a feasibility study in the EGB, (ii) detection of the selected controls via 3 case-based designs: case-population approach (CP), case-control design (CC) and self-controlled case series (SCCS), including several variants (number of controls, risk window, adjustment strategy, etc.) and (iii) comparison of design performance using area under the ROC curve. Cases were identified between 01/01/2009 and 12/31/2014 according to hospitalization primary diagnoses. A narrow and a broad definition have been developed for each outcome. For each design and outcome, the accuracy of the measures of association will be used to calibrate the methods. **Results:** The feasibility study is currently ongoing. Based on the broad outcome definitions, 40 ALI, 6,334 MI, 758 KI and 1,774 UGIB have been identified in the EGB, versus 33 ALI, 3,202 MI, 94 KI and 1,390 UGIB for the narrow one. In respect of the reference set, 120 positive and 126 negative drug controls are present in the EGB. Power calculations are in process to determine which controls will have enough power to be investigated through the 80 CP, 40 CC and 336 SCCS variants. **Conclusions:** This project will identify and calibrate the best design to investigate ALI, MI, KI and UGIB in the SNIRAM, thus enabling the generation and validation of drug safety alerts.

Background

- The SNIRAM¹ is the French nationwide healthcare insurance system database covering 99% of the French population. It has not been tested for drug safety signal generation
- The EGB² is a 1/97th SNIRAM sample
- ALCAPONE (Alert generation using the case-population approach in the French claims databases) is a project aiming to:
 - Develop on SNIRAM the case-population approach for drug safety signal generation
 - Compare the performances of this approach with the case-control design and self-controlled case series ones, according to the Observational Medical Outcome Partnership (OMOP) methodology

¹ Système national d'information inter-régimes de l'Assurance maladie – ² Echantillon généraliste de bénéficiaires

Conflict of Interest Statement

This study is supported by an unconditional grant from the French Ministry of Health (PREPS, 2014, 14-0635), and supervised by an independent expert Scientific Committee. It was designed, conducted, and analyzed independently by the Bordeaux PharmacoEpi Platform, CIC Bordeaux CIC1401 of the Bordeaux University. This study was registered with the European Medicine agency's EUPAS registry (www.encepp.eu), under study number 13031, and carries the ENCePP Study seal.

Objectives

- To present the methodology of the ALCAPONE project.
- To assess the feasibility of the project through preliminary results from the EGB database.

Methods

- Study design**
 - OMOP reference set
 - 4 health outcomes of interest
 - Acute liver injury (ALI)
 - Acute kidney injury (KI)
 - Myocardial infarction (MI)
 - Upper gastrointestinal bleeding (UGIB)
 - Drug controls
 - Positive controls (CTR+) have been associated with the outcome of interest (RR>1)
 - Negative controls (CTR-) have not been associated with the outcome of interest (RR=1)
 - Historical data
 - From the EGB (feasibility study) and from the SNIRAM (final study)
 - Case-based extractions between 01/01/2009 and 12/31/2014
 - Project stages
 - Case-based patients extraction and selection of the detectable drug controls
 - Extraction (from EGB or SNIRAM) of 4 sub-populations: ALI, MI, KI, UGIB
 - According to a narrow definition
 - According to a broad definition.
 - Selection of the drugs available and reimbursed in the French community pharmacies among the ones of the OMOP Reference set.
 - Calculation of the minimum detectable relative risk (MDRR) with $\alpha = 0,05$ and $1-\beta = 0,80$ of each drug-outcome pair.

$$MDRR = \left[1 + \frac{z_{\alpha} - z_{1-\beta}}{2 \sqrt{\text{patients exposed}}} \right]^2$$
 - Elimination of the controls with MDRR > 1,25.
 - Generation of 4 sub-study databases composed of the cases extracted for a health outcome of interest and the corresponding reference containing the detectable drug controls.
 - Drug-outcome pairs detection
 - Generation of a measure of association for each drug-outcome pair
 - Via 3 study designs: (1) case-control, (2) self-controlled case series and (3) case-population
 - Each study design is repeated according to different settings e.g.:
 - Case-control: number of controls per case, matching strategy...
 - Self-controlled case series: adjustment strategy, pre-exposure window...
 - Case-population exposure window, exclusion period...
 - Each setting of a design is considered as a variant.
 - Generation of one measure of association by drug-outcome pair and design variant.
 - Comparison of design and design variants performances
 - Discriminating power
 - Detected CTR+ et CTR- → Specificity & Sensitivity → Area under the ROC curve
 - Accuracy of the measure of association (for CTR- only)
 - $MSE = \text{mean}[\log(RR_{est}) - \log(RR_{true})]^2$
 - $\text{Bias} = \text{mean}[\log(RR_{est}) - \log(RR_{true})]$
 - Coverage probability: frequency over replications that the confidence interval contains the true value.
 - Selection of the best design variant for each health outcome of interest.
 - Calibration of the selected design variant based on the CTR-.

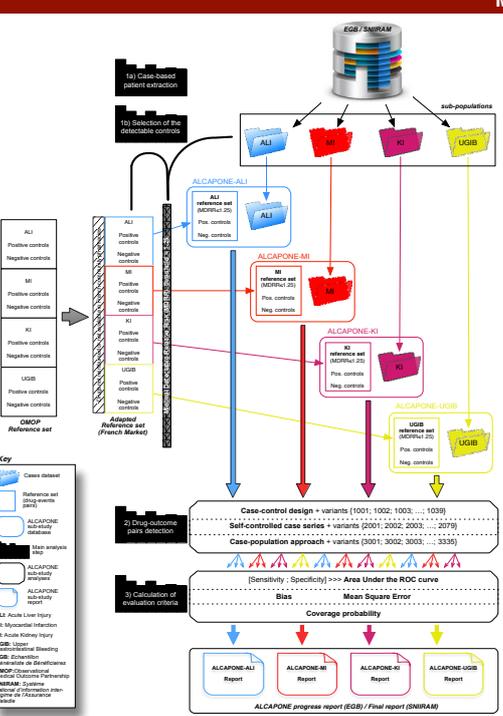


Figure 1: Overall ALCAPONE process

Results

Table 1: Number of positive and negative controls (CTR+ and CTR-) by health outcome of interest, present in the OMOP experiment, available in the French market, detectable in the EGB and expected in the SNIRAM [stage 1b of the Figure 1]

OMOP Reference set	OMOP Experiment ¹		French market Reference set		EGB (197th SNIRAM sample) Number of detectable controls ²		SNIRAM Expected number of detectable controls ²		
	Narrow definition	Broad definition	Narrow definition	Broad definition	Narrow definition	Broad definition	Narrow definition	Broad definition	
									Number of controls
ALI	CTR+	81	57	63	56	0	0	15	18
	CTR-	37	32	32	19	0	0	1	2
	MI	CTR+	36	26	33	26	3	5	23
KI	CTR+	24	19	-	19	0	3	11	18
	CTR-	64	34	-	32	0	0	5	16
UGIB	CTR+	24	24	22	19	5	7	18	19
	CTR-	66	53	49	38	1	1	30	31

¹ Results from the MarketScan Commercial Claims and Encounters database ² Drug controls with MDRR ≤ 1.25

Table 2: Outcomes included in the ALCAPONE project and corresponding number of patients by health outcome of interest definition in the EGB; Expected number for SNIRAM [stage 1a of the Figure 1]

		ALI		MI		KI		UGIB	
		Narrow Def.	Broad Def.						
EGB	n (outcomes)	33	40	3202	6334	94	93	1390	1771
	n (patients)	32	40	2757	4962	758	712	1213	1522
SNIRAM	n (outcomes)	3960	4800	384240	760080	11280	11160	168800	212520
	n (patients)	3840	4800	330840	595440	90960	85440	145560	182640

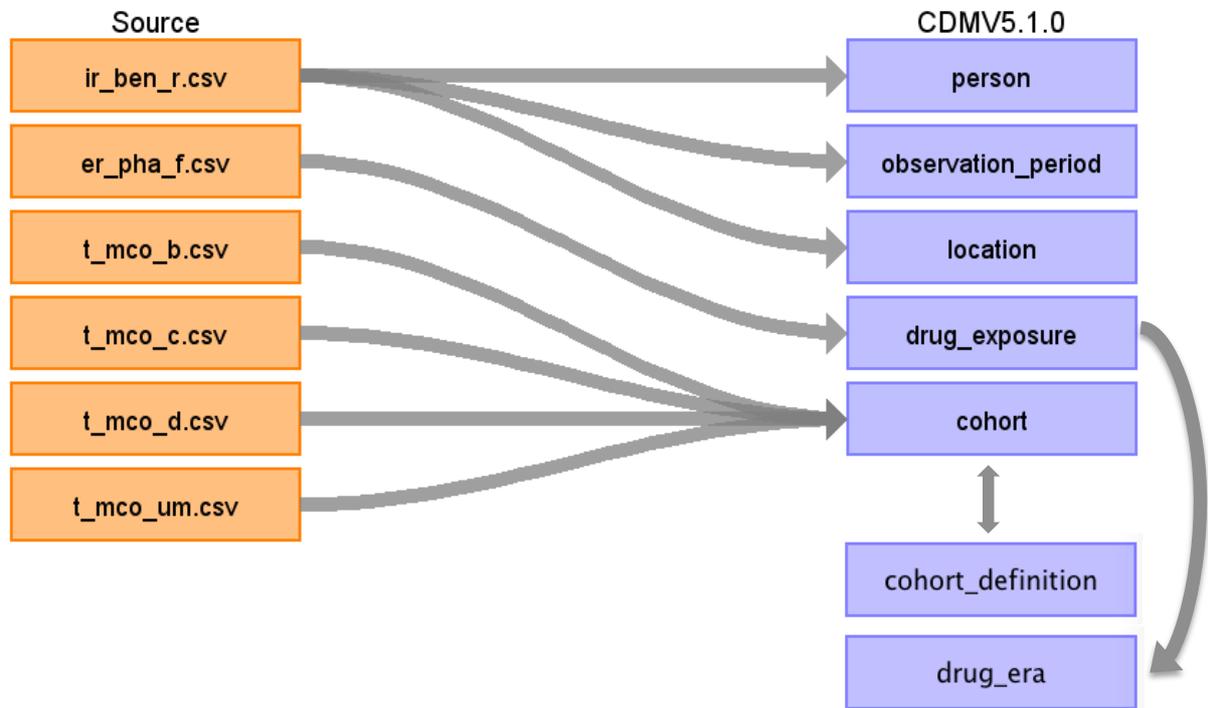
The left part of Table 1 shows the number of controls with MDRRs ≤ 1.25 in a database of the OMOP experiment. The right side displays the results of the feasibility study and its extrapolation to the SNIRAM. Table 2 presents the number of outcomes extracted from the EGB.

Example of Table 1, reading: Among the 81 positive controls of the ALI OMOP-Reference set, only 56 are available on the French market. The number of exposed cases in the EGB is not enough to detect an association ≤ 1.25 whatever the definition. By extrapolation, the SNIRAM would be powerful enough to detect an association ≤ 1.25 for 15 of the 56 positive controls and 1 of the 19 negative ones (narrow definition).

Conclusion

- The feasibility study shows that the EGB is not powerful enough, especially when the event and/or the exposure is rare. The SNIRAM seems to have a sufficient size to implement the ALCAPONE process.
- The step 1b "Selection of detectable controls" must be repeated after SNIRAM extraction to confirm the number of detectable drug controls. If necessary, additional ones could be added to enhance the French market Reference set.
- The identification of the optimal design for a health outcome of interest will enable the generation and the validation of drug safety alerts.

Appendix 7 Complete Extract, Transform and Load process of SNDS to OMOP CDM v5 for the ALCAPONE project



- **Table name: person**

Reading from IR_BEN_R.csv

Destination Field	Source Field	Logic	Comment
person_id			To be generated
gender_concept_id	ben_sex_cod		8507 = "Male" 8532 = "Female"
year_of_birth	ben_nai_ann		
month_of_birth			
day_of_birth			
birth_datetime			
race_concept_id			
ethnicity_concept_id			
location_id	ben_res_dpt bdi_dep		201 = department "02A" 202 = department "02B"
provider_id			
care_site_id			
person_source_value	num_enq_ano		
gender_source_value	ben_sex_cod		
gender_source_concept_id			
race_source_value			0
race_source_concept_id			
ethnicity_source_value			0
ethnicity_source_concept_id			

- **Table name: observation_period**

Reading from IR_BEN_R.csv

Destination Field	Source Field	Logic	Comment
observation_period_id			To be generated
person_id	num_enq_ano		
observation_period_start_date	year_of_birth		<ul style="list-style-type: none"> • If year of birth \geq 01/01/2009 then observation_period_start_date = 01/01/YEAR OF BIRTH • If year of birth $<$ 01/01/2009 then observation_period_start_date = 01/01/2009
observation_period_start_datetime			
observation_period_end_date	ben_dcd_ame		<ul style="list-style-type: none"> • If death date $<$ 31/12/2014 then observation_period_end_date = death date • If death date \geq 31/12/2014 then observation_period_end_date = 31/12/2014
observation_period_end_datetime			
period_type_concept_id			44814722 Period while enrolled in insurance

- **Table name: location**

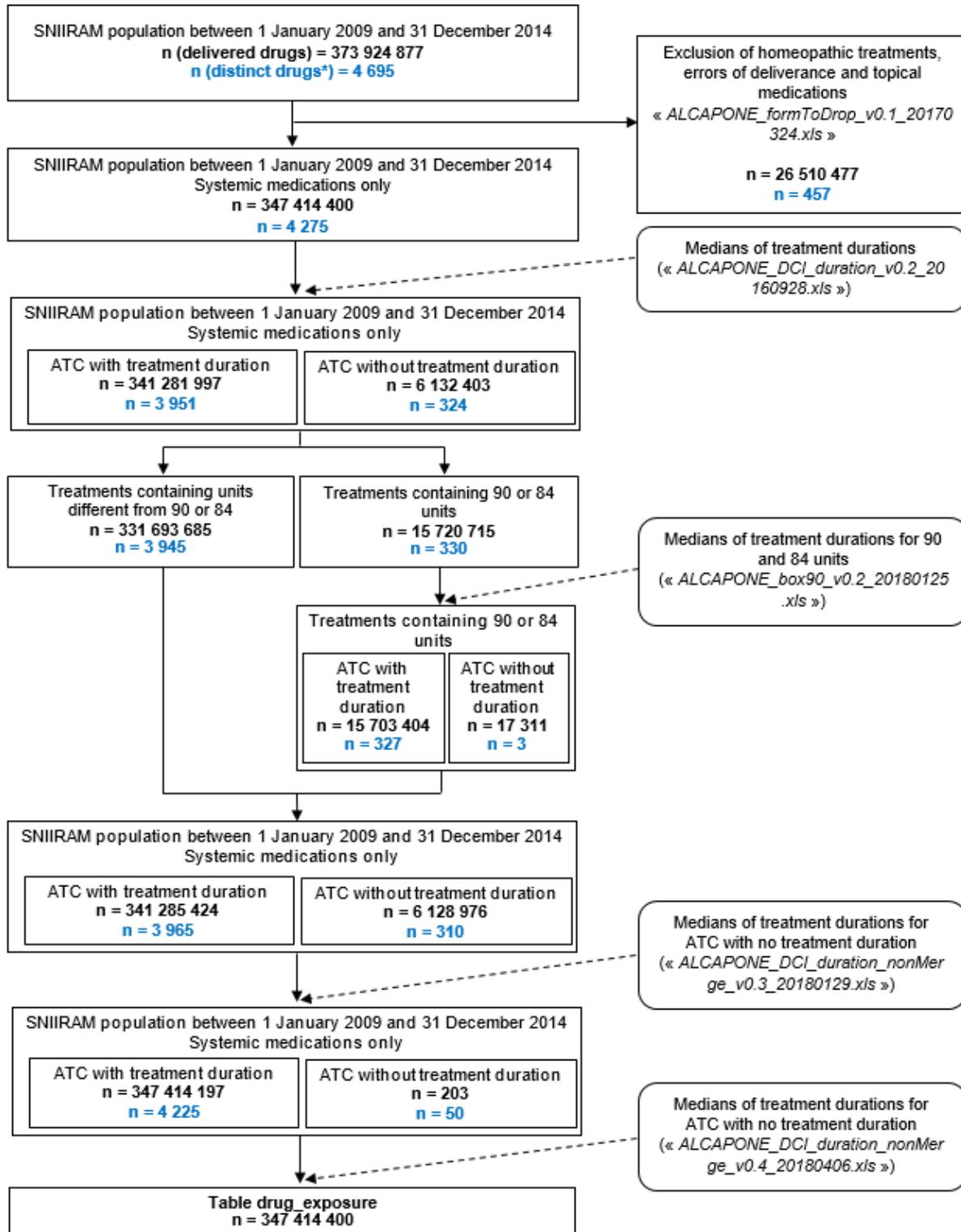
Reading from IR_BEN_R.csv

Destination Field	Source Field	Logic	Comment
county			
location_id			
address_1			
address_2			
location_source_value	ben_res_dpt		Department number in numeric format
city			
state			
zip			

- **Table name: drug_exposure**

Reading from ER_PHA_F.csv

Construction of drug treatment duration:



*distinct by ATC, number of units by box, dosage

Figure. ALCAPONE drug exposure selection.

- Homeopathic treatments, errors of deliverance and topical medications have been excluded;
- Treatment dispensations were merged with the medians of treatment durations (“ALCAPONE_DCI_duration_v0.2_20160928.xls”) by ATC code;
- Treatments containing 90 or 84 units were merged with a specific file of medians of treatment durations for 90 and 84 units (“ALCAPONE_box90_v0.2_20180125.xls”) by ATC code;
- Remaining ATC codes without duration were reviewed and merged by ATC code with *ad-hoc* files containing the missing medians.
 - “ALCAPONE_DCI_duration_nonMerge_v0.3_20180129.xls” corresponding to the unmapped ATC codes identified during the feasibility study
 - “ALCAPONE_DCI_duration_nonMerge_v0.4_20180406.xls” corresponding to the remaining unmapped ATC codes
 - ATC codes were finally translated into OMOP concept_id (vocabulary_id=ATC) using the OMOP vocabulary

Destination Field	Source Field	Logic	Comment
drug_exposure_id			To be generated
person_id	num_enq_an o		
drug_concept_id	pha_atc_c07		Vocabulary = ATC
drug_exposure_start_date	exe_soi_dtd		
drug_exposure_start_datetime			
drug_exposure_end_date	exe_soi_dtd pha_unt_nbr_ dses		drug_exposure_end_date=drug_exposure_start_date+days_supply
drug_exposure_end_datetime			
drug_type_concept_id			38000175 « Prescription dispensed in pharmacy »
stop_reason			
refills			
quantity			
days_supply			Median of treatment duration
sig			
route_concept_id			
effective_drug_dose			
dose_unit_concept_id			
lot_number			
provider_id			
visit_occurrence_id			
drug_source_value	pha_atc_c07		ATC code in character
drug_source_concept_id			44819117
route_source_value			
dose_unit_source_value			

- **Table name: drug_era**

Reading from drug_exposure

Destination Field	Source Field	Logic	Comment
drug_era_id	drug_exposure_id		To be generated
person_id	person_id		
drug_concept_id	drug_concept_id		Agregated at 7th ATC level
drug_era_start_date	drug_exposure_start_date		
drug_era_end_date	drug_exposure_end_date		Constructed from drug_exposure end_date : two exposures with a gap of 30 days or less form a single drug era
drug_exposure_count			The number of individual Drug exposure occurrences used to construct the Drug era
gap_days			The number of accumulated gap days in the whole drug era

- **Table name: cohort**

Reading from T_MCO_B.csv, T_MCO_C.csv, T_MCO_D.csv, T_MCO_UM.csv

Cohort table only contains patients responding to ALCAPONE inclusion criteria.

Destination Field	Source Field	Logic	Comment
cohort_definition_id			1001 1002 1003 1004 1005 1006 1007 1008 According to cohort_definition tables
subject_id	num_enq_and_o		
cohort_start_date	exe_soi_dtd		
cohort_end_date	exe_soi_dtf		

- **Table name: cohort_definition**

Destination Field	Source Field	Logic	Comment
cohort_definition_id	cohort_definition_id		1001
cohort_definition_name			ALI
cohort_definition_description			K71.1 K71.2 K71.6 K72.0 Exclusion of Z94.4 B18 B24 C F10 G31 I50 I81 I85 K70 K73 K74 K75.4 K76 K80 K83 R18 Z95 T86.4 B15 B16 B17 B19 B25.1 K77 026.6 S36.1 T36-T65
Definition_type_concept_id			
Cohort_definition_syntax			
Subject_concept_id			19
Cohort_instantiation_date			2014-12-31

Destination Field	Source Field	Logic	Comment
cohort_definition_id	cohort_definition_id		1002
cohort_definition_name			ALItot
cohort_definition_description			K71.1 K71.2 K71.6 K71.9 K72.0 Exclusion of Z94.4 B18 B24 C F10 G31 I50 I81 I85 K70 K73 K74 K75.4 K76 K80 K83 R18 Z95 T86.4 B15 B16 B17 B19 B25.1 K77 026.6 S36.1 T36-T65
Definition_type_concept_id			
Cohort_definition_syntax			
Subject_concept_id			19
Cohort_instantiation_date			2014-12-31

Destination Field	Source Field	Logic	Comment
cohort_definition_id	cohort_definition_id		1003
cohort_definition_name			KI
cohort_definition_description			N17.0 N17.1 N17.2 Exclusion of E14.2 N14.3 N18 N26 N28.0 N28.1 N99 090.4 P96.0 Q61 T86.1 Z90.5 Z94.0 T56 T57 T62.0
Definition_type_concept_id			
Cohort_definition_syntax			
Subject_concept_id			19
Cohort_instantiation_date			2014-12-31

Destination Field	Source Field	Logic	Comment
cohort_definition_id	cohort_definition_id		1004
cohort_definition_name			KItot
cohort_definition_description			N17.0 N17.1 N17.2 N17.8 N17.9 N19 Exclusion of E14.2 N14.3 N18 N26 N28.0

			N28.1 N99 090.4 P96.0 Q61 T86.1 Z90.5 Z94.0 T56 T57 T62.0
Definition_type_concept_id			
Cohort_definition_syntax			
Subject_concept_id			19
Cohort_instantiation_date			2014-12-31

Destination Field	Source Field	Logic	Comment
cohort_definition_id	cohort_definition_id		1005
cohort_definition_name			MI
cohort_definition_description			I21.0 I21.1 I21.2 I21.3 I21.4 I21.9
Definition_type_concept_id			
Cohort_definition_syntax			
Subject_concept_id			19
Cohort_instantiation_date			2014-12-31

Destination Field	Source Field	Logic	Comment
cohort_definition_id	cohort_definition_id		1006
cohort_definition_name			MItot
cohort_definition_description			I21.0 I21.1 I21.2 I21.3 I21.4 I21.9 I20.0
Definition_type_concept_id			
Cohort_definition_syntax			
Subject_concept_id			19
Cohort_instantiation_date			2014-12-31

Destination Field	Source Field	Logic	Comment
cohort_definition_id	cohort_definition_id		1007
cohort_definition_name			UGIB
cohort_definition_description			K25.0 K25.2 K25.4 K25.6 K26.0 K26.2 K26.4 K26.6 K27.0 K27.2 K27.4 K27.6 K28.0 K28.2 K28.4 K28.6 K29.0 K92.0 K92.1
Definition_type_concept_id			
Cohort_definition_syntax			
Subject_concept_id			19
Cohort_instantiation_date			2014-12-31

Destination Field	Source Field	Logic	Comment
cohort_definition_id	cohort_definition_id		1008
cohort_definition_name			UGIBtot
cohort_definition_description			K25.0 K25.2 K25.4 K25.6 K26.0 K26.2 K26.4 K26.6 K27.0 K27.2 K27.4 K27.6 K28.0 K28.2 K28.4 K28.6 K29.0 K92.0 K92.1 K92.2
Definition_type_concept_id			
Cohort_definition_syntax			
Subject_concept_id			19
Cohort_instantiation_date			2014-12-31