OHDSI Method Library Evaluation study protocol

**Version:** 0.2

**Authors:**

Martijn Schuemie, PhD, Janssen Research and Development

Marc Suchard, MD, PhD, University of California, Los Angeles

Patrick Ryan, PhD, Janssen Research and Development

George Hripcsak, MD, PhD, Columbia University

David Madigan, PhD, Columbia University

**Date:** 12 December 2018

**Acknowledgement:** The analysis is based in part on work from the Observational Health Sciences and Informatics collaborative. OHDSI (<http://ohdsi.org>) is a multi-stakeholder, interdisciplinary collaborative to create open-source solutions that bring out the value of observational health data through large-scale analytics.

The authors declare the following disclosures: Drs. Ryan, Schuemie are employees of Janssen Research & Development.

Table of contents

[2 List of abbreviations 3](#_Toc532392679)

[3 Abstract 3](#_Toc532392680)

[4 Amendments and Updates 3](#_Toc532392681)

[5 Rationale and Background 3](#_Toc532392682)

[6 Research Questions and Objectives 4](#_Toc532392683)

[6.1 Research Questions 4](#_Toc532392684)

[6.2 Objectives 4](#_Toc532392685)

[7 Research methods 4](#_Toc532392686)

[7.1 The OHDSI Methods Benchmark 4](#_Toc532392687)

[7.1.1 Gold standard 4](#_Toc532392688)

[7.1.2 Metrics 6](#_Toc532392689)

[7.2 Variables 7](#_Toc532392690)

[7.2.1 Exposures 7](#_Toc532392691)

[7.2.2 Outcomes 7](#_Toc532392692)

[7.2.3 Nesting cohorts 12](#_Toc532392693)

[7.3 Data Sources 12](#_Toc532392694)

[7.4 Methods to evaluate 14](#_Toc532392695)

[7.4.1 CohortMethod 14](#_Toc532392696)

[7.4.2 SelfControlledCohort 16](#_Toc532392697)

[7.4.3 CaseControl 17](#_Toc532392698)

[7.4.4 CaseCrossover 18](#_Toc532392699)

[7.4.5 SelfControlledCaseSeries 18](#_Toc532392700)

[7.5 Sample Size and Study Power 20](#_Toc532392701)

[7.6 Quality control 20](#_Toc532392702)

[7.7 Strengths and Limitations of the Research Methods 20](#_Toc532392703)

[8 Protection of Human Subjects 20](#_Toc532392704)

[9 Plans for Disseminating and Communicating Study Results 21](#_Toc532392705)

[10 References 21](#_Toc532392706)

[11 Appendix A: Negative controls 23](#_Toc532392707)

# List of abbreviations

CDM Common Data Model

OHDSI Observational Health Data Sciences and Informatics

OMOP Observational Medical Outcomes Partnership

PS Propensity Scores

SCCS Self-Controlled Case Series

SCC Self-Controlled Cohort

MSE Mean Squared Error

MDRR Minimum Detectable Relative Risk

CCAE Commercial Claims And Encounters

NLP Natural Language Processing

# Abstract

Here we evaluate a large set of population-level estimation methods in the OHDSI Methods Library using the OHDSI Method Evaluation Benchmark. The Benchmark consists of a set of 200 real negative and 600 synthetic positive controls. All methods are executed to produce estimates for the controls, and the estimated effect sizes are used to compute various performance metrics.

# Amendments and Updates

|  |  |  |  |
| --- | --- | --- | --- |
| Version | Date | Author(s) | Comments |
| 0.2 | 12 Dec 2018 | Martijn Schuemie | Added MDCR and JMDC databases |
| 0.1 | 25 Nov 2018 | Martijn Schuemie | Initial draft |

# Rationale and Background

When designing an observational study, there are many study designs to choose from, and many additional choices to make, and it is often unclear how these choices will affect the accuracy of the results. (e.g. If I match on propensity scores, will that lead to more or less bias than when I stratify? What about power?) The literature contains many papers evaluating one design choice at a time, but often with unsatisfactory scientific rigor; typically, a method is evaluated on one or two exemplar study from which we cannot generalize, or by using simulations which have an unclear relationship with the real world.

Here we apply a new benchmark for evaluating population-level estimation methods, one that can inform on how a particular study design and set of analysis choices perform in general. The benchmark consists of a gold standard of research hypothesis where the truth is known, and a set of metrics for characterizing a methods performance when applied to the gold standard. We distinguish between two types of tasks: (1) estimation of the average effect of an exposure on an outcome relative to no exposure (*effect estimation*), and (2) estimation of the average effect of an exposure on an outcome relative to another exposure (*comparative effect estimation*). The benchmark allows evaluation of a method on either or both tasks.

We apply the benchmark to establish the operating characteristics of several standard observational analysis methods. Specifically, we focus on those methods implemented in the OHDSI Methods Library, including the new-user cohort (optionally using propensity score adjustment), self-controlled case series (SCCS), case-control, self-controlled cohort, case-crossover, and case-time control designs. Besides the lessons learned from the OMOP experiment little is known about the performance of these designs in general, and even less of the performance of these specific implementations of the designs. We use the OHDSI Methods Benchmark to evaluate performance.

# Research Questions and Objectives

## Research Questions

What are the operating characteristics of the various population-level effect estimation methods in the OHDSI Methods Library?

## Objectives

To measure the operating characteristics of the various population-level effect estimation methods in the OHDSI Methods Library using the OHDSI Methods Benchmark.

# Research methods

## The OHDSI Methods Benchmark

The OHDSI Methods Benchmark consists of a gold standard (i.e. a set of research questions where the answer is known), and a set of metrics to characterize a method’s performance in estimating the answers.

### Gold standard

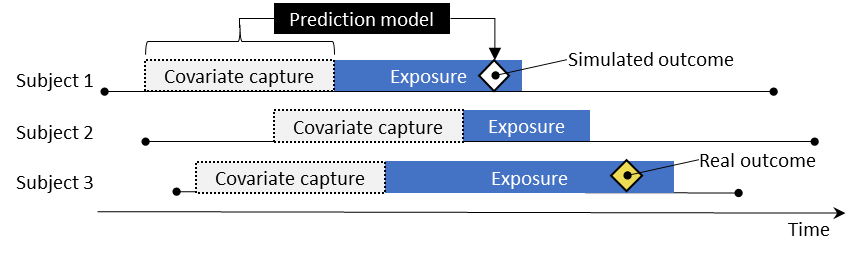
The gold standard comprises 800 entries, with each item specifying a target exposure, comparator exposure, outcome, nesting cohort, and true effect size. Four example entries are shown in Table 1. The true effect size refers to the absolute effect of the target on the outcome. Because the comparator is always believed to have no effect on the outcome, the true effect size also holds for the relative effect of the target compared to the comparator. Thus, each entry can be used for evaluating both effect estimation and comparative effect estimation. The nesting cohort identifies a more homogeneous subgroup of the population and can be used to evaluate methods such as the nested case-control design.

**Table 1**. Example entries in the gold standard.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Target | Comparator | Nesting cohort | Outcome | True effect size |
| Brinzolamide | Levobunolol | Glaucoma | Acute pancreatitis | 1.0 |
| Cevimeline | Pilocarpine | Sjogren's syndrome | Acute pancreatitis | 1.0 |
| Diclofenac | Celecoxib | Arthralgia | Acute stress disorder | 1.0 |
| Diclofenac | Celecoxib | Arthralgia | Ingrowing nail | 1.0 |

Of the total set, 200 entries are real negative controls with a presumed true relative risk of 1. We select these negative controls by first picking four diverse outcomes (acute pancreatitis, gastrointestinal bleeding, inflammatory bowel disease, and stroke) and four diverse exposures (ciprofloxacin, diclofenac, metformin, and sertraline). For each of the four outcomes, we create 25 entries with target and comparator exposures that are not believed to cause the outcome. For example, the top two rows of Table 1 are selected for the outcome of ‘acute pancreatitis’. For each of the four exposures, we select a single comparator exposure and create 25 entries with outcomes not believed to be caused by the target and comparator exposures. For example, the bottom two rows of Table 1 are selected for the exposure ‘diclofenac’, with celecoxib being used as the comparator for all diclofenac entries. To aid the creation of these entries, we generate candidate lists of negative controls for each of the four main outcomes and four main exposures using an automated procedure[1](#_ENREF_1), drawing on literature, product labels, and spontaneous reports. These candidates are used to construct target-comparator-outcome triplets where neither the target nor the exposure causes the outcome, and the target and comparator were either previously compared in a randomized trial per ClinicalTrials.gov, or both had the same four-digit ATC code (same indication) but not the same five-digit ATC code (different class). These candidates are ranked on prevalence of the exposures and outcome and manually reviewed until 25 were approved per initial outcome or exposure. Nesting cohorts were selected by manually reviewing the most prevalent conditions and procedures on the first day of the target or comparator treatment. The full set of negative controls is included as Appendix A.

The remaining 600 entries are positive controls. To avoid the aforementioned shortcomings of real positive controls, we choose to generate synthetic positive controls[2](#_ENREF_2). These are automatically derived from the 200 negative controls by adding simulated additional outcomes during the target exposure until a desired incidence rate ratio is achieved between before and after injection of the synthetic outcomes. For example, assume that, during exposure to diclofenac, n occurrences of ingrowing nail were observed. If we now add an additional n simulated occurrences during exposure, we have doubled the risk. Since this was a negative control, the relative risk compared to the counterfactual was one, but after injection, it becomes two. Using this process, we have generated positive controls based on each negative control with relative risks of 1.5, 2, and 4. To preserve (measured) confounding, predictive models are fitted for each outcome during target exposure and used to generate probabilities from which the synthetic outcomes are sampled. This process is depicted in Figure 1.



**Figure 1**. Synthesizing positive controls from negative controls. For each outcome a model is learned to predict the Poisson rate of the outcome during exposure using covariates captured prior to exposure. These covariates include demographics, as well as all recorded diagnoses, drug exposures, measurements, and medical procedures. A Poisson regression using L1 regularization is used to fit the prediction model. The predicted rates are subsequently used to sample simulated outcomes during exposure to increase the true effect size to the desired magnitude. The resulting positive control thus contains both real and simulated outcomes.

### Metrics

Once an effect estimation method has been selected for evaluation, it should be applied to estimate the effects for the negative and positive controls in the gold standard. We will assume such a method will produce an effect size estimate for each control, expressed as either relative risk, odds ratio, or hazard ratio, together with an indication of the uncertainty of the estimate expressed as a 95% confidence interval. We also assume that a two-sided p-value is computed for the null hypothesis of no effect.

Based on the estimates of a particular method for the 800 negative and positive controls, we can then compute the following metrics:

* **AUC**: the ability to discriminate between positive controls and negative controls.
* **Coverage**: how often the true effect size is within the 95% confidence interval.
* **Mean precision**: Computed as 1/(standard error)2, higher precision means narrower confidence intervals.
* **Mean squared error (MSE)**: Mean squared error between effect size point-estimate and the true effect size.
* **Type 1 error**: For negative controls, how often was the null rejected (at alpha = 0.05).
* **Type 2 error**: For positive controls, how often was the null not rejected (at alpha = 0.05).
* **Non-estimable**: For how many of the controls was the method unable to produce an estimate? There can be various reasons why an estimate cannot be produced, for example because there were no subjects left after propensity score matching, or because no subjects remained having the outcome.

These metrics are computed both overall, as well as stratified by true effect size, by each of the four initial outcomes and four initial exposures, and by amount of data as reflected by the minimum detectable relative risk (MDRR) computed using a standard approach[3](#_ENREF_3). When a method cannot compute an effect size, for example due to limited sample size, its estimate is assumed to be a relative risk of 1 with infinite confidence interval when computing these metrics.

In prior work we showed that the p-value and confidence interval can be empirically calibrated[2](#_ENREF_2)[4](#_ENREF_4) to restore their operating characteristics to near-nominal values using the observed distribution of estimates for negative and positive controls. We therefore also report the performance metrics listed above after calibration.

## Variables

### Exposures

Exposures are defined as exposure to any drug containing the ingredient specified in the gold standard. Consecutive exposures are merged if the gap between exposures is less than 30 days.

### Outcomes

The four main outcomes of interest (acute pancreatitis, gastrointestinal bleeding, inflammatory bowel disease, and stroke) are defined using manually crafted rule-based definitions including various diagnosis codes. The formal definitions are below. All other outcomes are defined as the first occurrence of the outcome concept or any of its descendants.

#### Acute pancreatitis

Initial Event Cohort

People having any of the following: 

* a condition occurrence of Acute pancreatitis1
  + condition type is any of Inpatient detail - primary, Inpatient detail - primary position, Inpatient header - primary, Inpatient header - primary position, Primary Condition, Inpatient detail - 1st position
  + visit occurrence is any of: Emergency Room Visit, Inpatient Visit

with continuous observation of at least 0 days prior and 0 days after event index date, and limit initial events to: **earliest event per person.**

Limit qualifying cohort to: **all events per person.**

End Date Strategy

No end date strategy selected. By default, the cohort end date will be the end of the observation period that contains the index event.

Cohort Collapse Strategy:

Collapse cohort by era with a gap size of 0 days.

Appendix 1: Concept Set Definitions

1.  Acute pancreatitis

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 199074 | Acute pancreatitis | Condition | SNOMED | NO | YES | NO |

#### GI bleeding

Initial Event Cohort

People having any of the following: 

* a condition occurrence of GI bleeding1
  + condition type is any of Inpatient detail - primary, Inpatient header - primary, Primary Condition, Inpatient detail - 1st position, Inpatient header - 1st position
  + visit occurrence is any of: Emergency Room Visit, Inpatient Visit

with continuous observation of at least 0 days prior and 0 days after event index date, and limit initial events to: **earliest event per person.**

Limit qualifying cohort to: **all events per person.**

End Date Strategy

No end date strategy selected. By default, the cohort end date will be the end of the observation period that contains the index event.

Cohort Collapse Strategy:

Collapse cohort by era with a gap size of 0 days.

Appendix 1: Concept Set Definitions

1. GI bleeding

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 4280942 | Acute gastrojejunal ulcer with perforation | Condition | SNOMED | NO | YES | NO |
| 28779 | Bleeding esophageal varices | Condition | SNOMED | NO | YES | NO |
| 198798 | Dieulafoy's vascular malformation | Condition | SNOMED | NO | YES | NO |
| 4112183 | Esophageal varices with bleeding, associated with another disorder | Condition | SNOMED | NO | YES | NO |
| 194382 | External hemorrhoids | Condition | SNOMED | NO | NO | NO |
| 192671 | Gastrointestinal hemorrhage | Condition | SNOMED | NO | YES | NO |
| 196436 | Internal hemorrhoids | Condition | SNOMED | NO | NO | NO |
| 4338225 | Peptic ulcer with perforation | Condition | SNOMED | NO | YES | NO |
| 194158 | Perinatal gastrointestinal hemorrhage | Condition | SNOMED | YES | YES | NO |
| 194690 | Nontraumatic hemoperitoneum | Condition | SNOMED | NO | NO | NO |
| 443530 | Hematochezia | Condition | SNOMED | NO | NO | NO |
| 194395 | Bleeding diverticulosis | Condition | SNOMED | NO | YES | NO |
| 4048064 | Hemorrhagic diverticulitis | Observation | SNOMED | NO | YES | NO |

#### Inflammatory Bowel Disease

Initial Event Cohort

People having any of the following: 

* a condition occurrence of Inflammatory Bowel Disease (IBD)1

with continuous observation of at least 0 days prior and 0 days after event index date, and limit initial events to: **earliest event per person.**

Limit qualifying cohort to: **all events per person.**

End Date Strategy

No end date strategy selected. By default, the cohort end date will be the end of the observation period that contains the index event.

Cohort Collapse Strategy:

Collapse cohort by era with a gap size of 0 days.

Appendix 1: Concept Set Definitions

1. Inflammatory Bowel Disease (IBD)

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 4074815 | Inflammatory bowel disease | Condition | SNOMED | NO | YES | NO |

#### Acute stroke (ischemic or hemorrhagic)

Initial Event Cohort

People having any of the following: 

* a condition occurrence of  Acute stroke (ischemic or hemorrhagic)1
  + condition type is any of Inpatient detail - primary, Inpatient detail - primary position, Inpatient header - primary, Inpatient header - primary position, Primary Condition, Inpatient detail - 1st position
  + visit occurrence is any of: Emergency Room Visit, Inpatient Visit

with continuous observation of at least 0 days prior and 0 days after event index date, and limit initial events to: **earliest event per person.**

Limit qualifying cohort to: **earliest event per person.**

End Date Strategy

No end date strategy selected. By default, the cohort end date will be the end of the observation period that contains the index event.

Cohort Collapse Strategy:

Collapse cohort by era with a gap size of 0 days.

Appendix 1: Concept Set Definitions

1.  Acute stroke (ischemic or hemorrhagic)

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 374060 | Acute ill-defined cerebrovascular disease | Condition | SNOMED | NO | YES | NO |
| 376713 | Cerebral hemorrhage | Condition | SNOMED | NO | YES | NO |
| 4108356 | Cerebral infarction due to embolism of cerebral arteries | Condition | SNOMED | NO | YES | NO |
| 4110192 | Cerebral infarction due to thrombosis of cerebral arteries | Condition | SNOMED | NO | YES | NO |
| 4043731 | Infarction - precerebral | Condition | SNOMED | NO | YES | NO |
| 4345688 | Intracerebral hemorrhage in fetus or newborn | Condition | SNOMED | YES | YES | NO |
| 42873157 | Intracranial hemorrhage following injury | Condition | SNOMED | YES | YES | NO |
| 260841 | Perinatal subarachnoid hemorrhage | Condition | SNOMED | YES | YES | NO |
| 432923 | Subarachnoid hemorrhage | Condition | SNOMED | NO | YES | NO |
| 441709 | Subarachnoid hemorrhage following injury with open intracranial wound | Condition | SNOMED | YES | YES | NO |
| 443752 | Ventricular hemorrhage | Condition | SNOMED | YES | YES | NO |

### Nesting cohorts

The nesting cohorts are defined as the group of people that have any occurrence of the diagnosis code or any of its descendants. The cohort start date is defined as the day of the first such diagnosis, and the cohort end date is defined as the end of observation.

## Data Sources

For our evaluation we use four databases:

**CCAE**

The IBM MarketScan® Commercial Database (CCAE) database represents data from 138 million individuals enrolled in United States employer-sponsored insurance health plans from January 2000 until March 2018. The data includes adjudicated health insurance claims (e.g. inpatient, outpatient, and outpatient pharmacy) as well as enrollment data from large employers and health plans who provide private healthcare coverage to employees, their spouses, and dependents. Additionally, it captures laboratory tests for a subset of the covered lives. This administrative claims-database includes a variety of fee-for-service, preferred provider organizations, and capitated health plans. The major data elements contained within this database are outpatient pharmacy dispensing claims (coded with National Drug Codes), and inpatient and outpatient medical claims, which provide procedure codes (coded in CPT-4, HCPCs, ICD-9-CM or ICD-10-PCS) and diagnosis codes (coded in ICD-9-CM or ICD-10-CM). The data also contain selected laboratory test results (those sent to a contracted thirds-party laboratory service provider) for a non-random sample of the population (coded with LOINC codes).

**PanTher**

The Optum© Pan-Therapeutic (PanTher) Electronic Health Records (EHR) dataset contains medical record data of 86 million subjects primarily from United States Integrated Delivery Networks. These include clinical information, inclusive of prescriptions as prescribed and administered, lab results, vital signs, body measurements, diagnoses, procedures, and information derived from clinical notes using Natural Language Processing (NLP). PanTher integrates provider data from different EHR platforms (i.e. Cerner, Epic, GE, McKesson, etc.) and different versions of the same EHR platform. The data spans January 2006 to December 2017.

**JMDC**

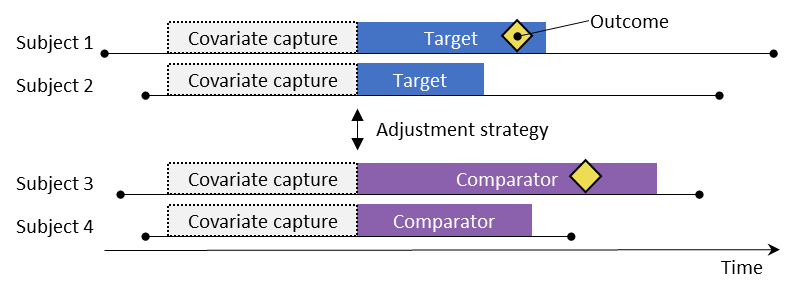
Japan Medical Data Center (JDMC) database consists of data from 60 society-managed health insurance plans covering workers aged 18 to 65 and their dependents (children younger than 18 years old and elderly people older than 65 years old). JMDC data includes membership status of the insured people and claims data provided by insurers under contract (e.g. patient-level demographic information, inpatient and outpatient data inclusive of diagnosis and procedures, and prescriptions as dispensed claims information). Claims data are derived from monthly claims issued by clinics, hospitals and community pharmacies. For claims only the month and year are provided; however, prescriptions, procedures, admission, discharge, and start of medical care are associated with a full date. The data covers 5 million subjects from January 2005 to December 2017.

**MDCR**

IBM MarketScan® Medicare Supplemental Database (MDCR) represents health services of retirees in the United States with primary or Medicare supplemental coverage through privately insured fee-for-service, point-of-service, or capitated health plans. These data include adjudicated health insurance claims (e.g. inpatient, outpatient, and outpatient pharmacy). Additionally, it captures laboratory tests for a subset of the covered lives. The data cover 10 million subjects from January 2000 to April 2018.

## Methods to evaluate

### CohortMethod

****

**Figure 2**. The new-user cohort design. Subjects observed to initiate the target treatment are compared to those initiating the comparator treatment. To adjust for differences between the two treatment groups several adjustment strategies can be used, such as stratification, matching, or weighting by the propensity score, or by adding baseline characateristcs to the outcome model. The chararacteristics included in the propensity model or outcome model are captured prior to treatment initiation.

The new-user cohort method attempts to emulate a randomized clinical trial[5](#_ENREF_5). Subjects that are observed to initiate one treatment (the target) are compared to subjects initiating another treatment (the comparator) and are followed for a specific amount of time following treatment initiation, for example the time they stay on the treatment. One crucial difference with a randomized trial is that there is no randomization, and therefore there might be systematic differences between the target and comparator populations. Without adjusting for these differences, estimates are likely to be confounded. A popular mechanism for adjusting for confounding is the use of Propensity Scores (PS). The PS is the probability of a subject receiving one treatment instead of the other, conditional on baseline characteristics[6](#_ENREF_6). In the past, PS were computed based on manually selected characteristics, and although the CohortMethod package can support such practices, we prefer the use of large-scale regularized regression using many generic characteristics[7](#_ENREF_7). These characteristics include demographics, as well as all diagnoses, drug exposures, measurement, and medical procedures observed prior to treatment initiation, typically ranging from 10,000 to 100,000 unique variables per model. The PS can be used in several ways, for example by stratifying the study population based on the PS, by matching target subjects to comparator subjects with similar PS, or by weighting subjects using Inverse Probability of Treatment Weighting (IPTW) derived from the PS.

Another strategy for adjusting for differences between the two groups is to include additional variables in the outcome model. One major limitation of this approach is that whereas there often is a wealth of data to fit a propensity model, with thousands of people in both treatment groups, the outcome tends to be rare, causing a paucity of data when trying to fit elaborate models with the outcome as dependent variable. One approach is to use both a PS and add the same variables that were used in the propensity model in the outcome model, thus adjusting for the same variables twice, but in different ways.

The new-user cohort method inherently is a method for comparative effect estimation, comparing one treatment to another. It is difficult to use this method to compare a treatment against no treatment, since it is hard to define a group of unexposed people that is comparable with the exposed group. If one wants to use this design for effect estimation, the preferred way is to select a comparator treatment for the same indication as the exposure of interest, where the comparator treatment is believed to have no effect on the outcome. Unfortunately, such a comparator might not always be available. In our gold standard, the comparators were specifically selected to have no effect, so we can also evaluate the cohort method’s performance for effect estimation.

**Evaluation settings**

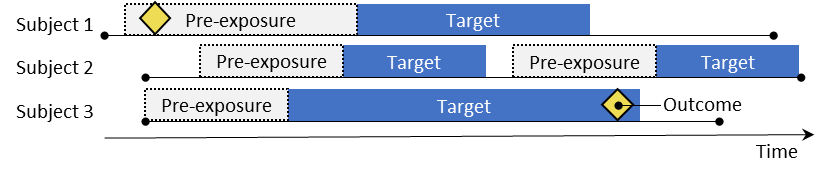
In our evaluation we focus on differences between the various adjustment strategies. All evaluations require 365 days of continuous observation prior to treatment initiation, capture a large set of covariates in the year prior to exposure, use a Cox proportional hazards model, and follow subjects from the day of treatment initiation until treatment discontinuation or end of observation, whichever is first. Subjects having both target and comparator exposures were removed. Propensity scores were computed using large-scale regularized regression[7](#_ENREF_7). All PS matching uses a caliper of 0.2 on the standardized logit scale. A full outcome model including all covariates that are also included in the PS is fitted using a large-scale Cox regression using regularization on all variables except the treatment variable.

**Table 2**. Analysis variants of the new-user cohort method included in the evaluation.

|  |  |  |  |
| --- | --- | --- | --- |
| Variant description | PS adjustment | Stratified outcome model | Add covariates to outcome model |
| No PS, simple outcome model | none | no | no |
| 1-on-1 matching, unstratified model | 1-on-1 matching | no | no |
| Variable ratio matching, stratified model | variable ratio matching | yes | no |
| Stratification | stratification | yes | no |
| IPTW | trimming + IPTW | no | no |
| Var. ratio matching + full model | variable ratio matching | yes | yes |

Table 2 lists the variants of the new-user cohort method included in the evaluation. A stratified outcome model is conditioned on the matched sets or PS strata and is required when using PS stratification or variable ratio matching. Variable ratio matching allows for more than one comparator subject to be selected for each target subject, as long as the matches stay within the predefined caliper[8](#_ENREF_8). Trimming is a common practice when performing IPTW to counter the effect of extreme weights[9](#_ENREF_9).

### SelfControlledCohort



**Figure 3**. The self-controlled cohort design. The rate of outcomes during exposure to the target is compared to the rate of outcomes in the time pre-exposure.

The self-controlled cohort (SCC) design compares the rate of outcomes during exposure to the rate of outcomes in the time just prior to the exposure. Because the same subject that make up the exposed group are also used as the control group, no adjustment for between-person differences need to be made.

**Evaluation settings**

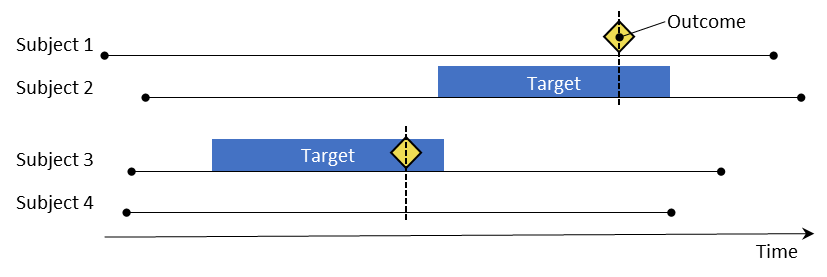
All evaluations of the SCC compare time while exposed to time prior to exposure, required 365 days of continuous observation prior to the exposure, and 183 days of continuous observation after the exposure start. Where possible, the pre-exposure window was set to the same length as the exposure window. All exposures were included, not just the first per person.

**Table 3**. Analysis variants of the self-controlled cohort design included in the evaluation.

|  |  |  |  |
| --- | --- | --- | --- |
| Description | Time at risk | Exposure start date | Require full observation |
| Time exposed, inc. exp. start date | Time exposed | In exposure window | no |
| 30 days, incl. exp. start date | 30 days | In exposure window | no |
| Time exposed, inc. exp. start date, require full obs. | Time exposed | In exposure window | yes |
| Time exposed, ex. exp. start date | Time exposed | Excluded | no |
| 30 days, ex. exp. start date | 30 days | Excluded | no |
| Time exposed, ex. exp. start date, require full obs. | Time exposed | Excluded | yes |

Table 3 shows the analysis variations include in our evaluation. We vary the definition of the time-at-risk to be either the entire time the subject was exposed, or just the first 30 days after exposure start. In all analyses the pre-exposure window was set to be the same length as the corresponding exposure window. In some variants the date when the exposure started was include in the time-at-risk, in others it was ignored. Sometimes the amount of observation time prior exposure is shorter than the time-at-risk window. By default, the pre-exposure window is then truncated to the available observation time, but in some analyses (marked ‘require full observation’) subjects were removed from the analyses if the pre-exposure observation time was too short.

### CaseControl



**Figure 4**. The case-control design. Subjects with the outcome (‘cases’) are compared to subjects without the outcome (‘controls’) in terms of their exposure status. Often, cases and controls are matched on various characteristics such as age and sex.

Case-control [10](#_ENREF_10) studies consider the question “are persons with a specific disease outcome exposed more frequently to a specific agent than those without the disease?” Thus, the central idea is to compare “cases”, i.e., subjects that experience the outcome of interest with “controls”, i.e., subjects that did not experience the outcome of interest. Often, one matches controls to cases based on characteristics such as age and sex to make them more comparable. Another widespread practice is to ‘nest’ the analysis within a specific subgroup of people, for example people that have all been diagnosed with one of the indications of the exposure of interest.

**Evaluation settings**

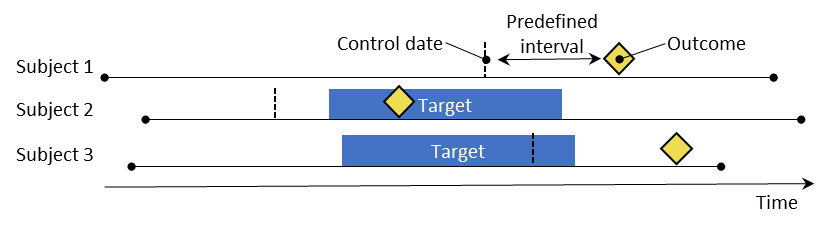
In all evaluation of the case-control design we match controls to cases on age and sex, set the index date of the cases to the date of the outcome, and use the same date as the index date the matched controls. We require 365 days of continuous observation prior to the index date and consider cases and controls to be exposed if they are exposed on the index date itself. The outcome model is logistic regression conditioned on the matched sets.

**Table 4**. Analysis variants of the case-control design included in the evaluation.

|  |  |  |
| --- | --- | --- |
| Description | Controls per case | Nesting in indication |
| 2 controls per case | 2 | no |
| 10 controls per case | 10 | no |
| Nesting in indication, 2 controls per case | 2 | yes |
| Nesting in indication, 10 controls per case | 10 | yes |

Table 4 enumerates the variants we evaluate. We either select up to two or ten matched controls per case, and optionally nest within the population within the indication.

### CaseCrossover



**Figure 5**. The case-crossover design. The time around the outcome is compared to a control date set at a predefined interval prior to the outcome date.

The case-crossover[11](#_ENREF_11) design evaluates whether the rate of exposure is different at the time of the outcome than at some predefined number of days prior to the outcome. It is trying to determine whether there is something special about the day the outcome occurred. Since cases serve as their own control, it is a self-controlled design, and should therefore be robust to confounding due to between-person differences. One concern is that, because the outcome date is always later than the control date, the method will be positively biased if the overall frequency of exposure increases over time (or negatively biased if there’s a decrease). To address this, the case-time-control design[12](#_ENREF_12) was developed, which adds matched controls to the case-crossover design to adjust for exposure trends.

**Evaluation settings**

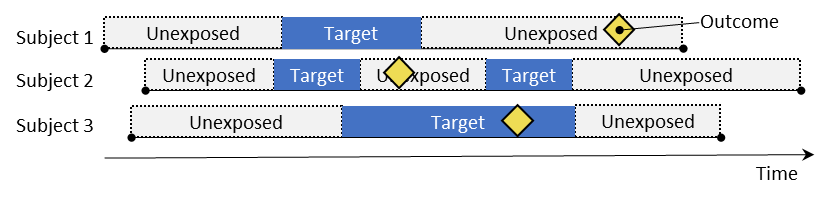
In all evaluation of the case-crossover design we require 365 days of continuous observation prior to the outcome date and consider subjects to be exposed if they are exposed on the outcome or control date itself. The outcome model is logistic regression conditioned on the persons.

**Table 5**. Analysis variants of the case-crossover design included in the evaluation.

|  |  |  |  |
| --- | --- | --- | --- |
| Description | Nesting in indication | Control window | Case-time-control |
| Simple case-crossover, -30 days | FALSE | -30 | no |
| Simple case-crossover, -180 days | FALSE | -180 | no |
| Nested case-crossover, -30 days | TRUE | -30 | no |
| Nested case-crossover, -180 days | TRUE | -180 | no |
| Nested case-time-control, -30 days | TRUE | -30 | yes |
| Nested case-time-control, -180 days | TRUE | -180 | yes |

Table 5 lists the variants of the case-crossover design included in the evaluation. When nested, the cases and, for the case-time-control extension the controls, are selected from the group of people having the indication.

### SelfControlledCaseSeries



**Figure 6**. The Self-Controlled Case Series design. The rate of outcomes during exposure is compared to the rate of outcomes when not exposed.

The Self-Controlled Case Series (SCCS) design[13](#_ENREF_13)[14](#_ENREF_14) compares the rate of outcomes during exposure to the rate of outcomes during all unexposed time, both before, between, and after exposures. It is a Poisson regression that is conditioned on the person. Thus, it seeks to answer the question: “Given that a patient has the outcome, is the outcome more likely during exposed time compared to non-exposed time?”. Like other self-controlled designs, the SCCS is robust to confounding due to between-person differences, but vulnerable to confounding due to time-varying effects. Several adjustments are possible to attempt to account for these.

**Evaluation settings**

In all evaluations we disregard the first 365 days in the analysis, except to determine the exposure status right after those initial 365 days. For example, if a 60-day prescription is started on day 340, the subject is considered exposed on days 366-400. Unless stated otherwise, the time-at-risk is assumed to start the day after exposure start, and end when exposure stops.

**Table 6**. Analysis variants of the case-control design included in the evaluation.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Description | Include start day | Pre-exposure window | Age and season | Event-dependent observation | All other exposures |
| Simple SCCS | no | no | no | no | no |
| Including exposure start day | yes | no | no | no | no |
| Using pre-exposure window | no | yes | no | no | no |
| Using age and season | no | no | yes | no | no |
| Using event-dependent observation | no | no | no | yes | no |
| Using all other exposures | no | no | no | no | yes |

Table 6 shows the variations of the SCCS included in the evaluation. We evaluate the effect of including the first day of exposure in the risk window, since this day could have many unrelated diagnoses being recorded while visiting the doctor. One frequent practice in SCCS designs is to set aside the time just prior to exposure to adjust for time-varying effects such as the contra-indications. We further evaluate adjusting for age and season by assuming a constant effect of age and season within each calendar month and using 5-knot cubic splines to model the effect across months. One important assumption underlying the SCCS is that the observation period end is independent of the date of the outcome. Because for some outcomes, especially ones that can be fatal such as stroke, this assumption can be violated an extension to the SCCS has been developed that corrects for any such dependency[15](#_ENREF_15). A final refinement of the SCCS is to include not just the exposure of interest, but all other exposures to drugs recorded in the database[16](#_ENREF_16), potentially adding thousands of additional variables to the model. L1-regularization is applied to the coefficients of all exposures except the exposure of interest.

## Sample Size and Study Power

The minimum detectable relative risk (MDRR) will be computed for each control, for each database. Results will be shown with and without filtering by MDRR.

## Quality control

The study R code and cohort definitions will undergo peer review.

The MethodEvaluation, EmpiricalCalibration, CohortMethod, SelfControlledCohort, CaseControl, CaseCrossover, and SelfControlledCaseSeries packages used in this study use unit tests for validation.

## Strengths and Limitations of the Research Methods

Strength

* Systematic empirical evaluation of a wide range of population-level estimation methods.
* Having a large set of negative and positive controls allows for accurate measurement of operating characteristics, and stratification to various types of controls.
* Negative controls capture both measured and unmeasured confounding.
* Positive controls are not known to physicians, and the evaluation is therefore not biased due to effects such as contra-indications and risk mitigations.

Limitations

* Operating characteristics will depend on the choice of database, and we therefore may not be able to generalize to databases not included in this study.
* Whereas negative controls reflect real confounding, both measured and unmeasured, positive controls do not retain all of that confounding. The injected outcomes only reflect measured bias.
* We do not vary the risk window definitions for the various methods and analyses settings. Instead, all designs use the equivalent of an ‘on-treatment’ design, since this is the risk window used for positive control synthesis. We believe the risk window definition is more part of the question we try to estimate an answer for instead of an analyses parameter for answering that question.

# Protection of Human Subjects

The study is using only de-identified data. Confidentiality of patient records will be maintained at all times. All study reports will contain aggregate data only and will not identify individual patients or physicians.

# Plans for Disseminating and Communicating Study Results

At least one paper describing the study and its results will be written and submitted for publication to a peer-reviewed scientific journal.

# References

1. Voss EA, Boyce RD, Ryan PB, et al. Accuracy of an automated knowledge base for identifying drug adverse reactions. *Journal of biomedical informatics* 2017;66:72-81. doi: 10.1016/j.jbi.2016.12.005 [published Online First: 2016/12/21]

2. Schuemie MJ, Hripcsak G, Ryan PB, et al. Empirical confidence interval calibration for population-level effect estimation studies in observational healthcare data. *Proceedings of the National Academy of Sciences of the United States of America* 2018;115(11):2571-77. doi: 10.1073/pnas.1708282114 [published Online First: 2018/03/14]

3. Armstrong B. A simple estimator of minimum detectable relative risk, sample size, or power in cohort studies. *American journal of epidemiology* 1987;126(2):356-8. [published Online First: 1987/08/01]

4. Schuemie MJ, Ryan PB, DuMouchel W, et al. Interpreting observational studies: why empirical calibration is needed to correct p-values. *Statistics in medicine* 2014;33(2):209-18. doi: 10.1002/sim.5925 [published Online First: 2013/08/01]

5. Hernan MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol* 2016;183(8):758-64. doi: 10.1093/aje/kwv254 [published Online First: 2016/03/20]

6. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70(1):41-55. doi: 10.1093/biomet/70.1.41

7. Tian Y, Schuemie MJ, Suchard MA. Evaluating large-scale propensity score performance through real-world and synthetic data experiments. *International journal of epidemiology* 2018 doi: 10.1093/ije/dyy120 [published Online First: 2018/06/26]

8. Rassen JA, Shelat AA, Myers J, et al. One-to-many propensity score matching in cohort studies. *Pharmacoepidemiology and drug safety* 2012;21 Suppl 2:69-80. doi: 10.1002/pds.3263 [published Online First: 2012/05/11]

9. Brookhart MA, Wyss R, Layton JB, et al. Propensity score methods for confounding control in nonexperimental research. *Circulation Cardiovascular quality and outcomes* 2013;6(5):604-11. doi: 10.1161/CIRCOUTCOMES.113.000359 [published Online First: 09/10]

10. Vandenbroucke JP, Pearce N. Case-control studies: basic concepts. *International journal of epidemiology* 2012;41(5):1480-9. doi: 10.1093/ije/dys147 [published Online First: 2012/10/10]

11. Maclure M. The Case-Crossover Design: A Method for Studying Transient Effects on the Risk of Acute Events. *American Journal of Epidemiology* 1991;133(2):144-53. doi: 10.1093/oxfordjournals.aje.a115853

12. Suissa S. The case-time-control design. *Epidemiology (Cambridge, Mass)* 1995;6(3):248-53. [published Online First: 1995/05/01]

13. Whitaker HJ, Farrington CP, Spiessens B, et al. Tutorial in biostatistics: the self-controlled case series method. *Statistics in medicine* 2006;25(10):1768-97. doi: 10.1002/sim.2302 [published Online First: 2005/10/13]

14. Farrington CP. Relative incidence estimation from case series for vaccine safety evaluation. *Biometrics* 1995;51(1):228-35. [published Online First: 1995/03/01]

15. Farrington CP, Anaya-Izquierdo K, Whitaker HJ, et al. Self-Controlled Case Series Analysis With Event-Dependent Observation Periods. *Journal of the American Statistical Association* 2011;106(494):417-26. doi: 10.1198/jasa.2011.ap10108

16. Simpson SE, Madigan D, Zorych I, et al. Multiple self-controlled case series for large-scale longitudinal observational databases. *Biometrics* 2013;69(4):893-902. doi: 10.1111/biom.12078 [published Online First: 2013/10/15]

# Appendix A: Negative controls

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| targetId | targetName | comparatorId | comparatorName | nestingId | nestingName | outcomeId | outcomeName | | type |
| 938044 | brinzolamide | 932815 | Levobunolol | 35606985 | Glaucoma | 1 | Acute pancreatitis | Exposure | |
| 954819 | cevimeline | 945286 | Pilocarpine | 35606954 | Sjogren's syndrome | 1 | Acute pancreatitis | Exposure | |
| 795113 | Chlorzoxazone | 704943 | Methocarbamol | 36516951 | Back pain | 1 | Acute pancreatitis | Exposure | |
| 1350310 | cilostazol | 1315865 | fondaparinux | 37622411 | Phlebosclerosis | 1 | Acute pancreatitis | Exposure | |
| 1517070 | desmopressin | 918906 | oxybutynin | 36718449 | Urinary incontinence | 1 | Acute pancreatitis | Exposure | |
| 1140088 | Dyphylline | 1110942 | omalizumab | 37203741 | Bronchospasm and obstruction | 1 | Acute pancreatitis | Exposure | |
| 943634 | epinastine | 1136422 | levocetirizine | 36009773 | Rhinitis allergic | 1 | Acute pancreatitis | Exposure | |
| 757352 | Eszopiclone | 704599 | Triazolam | 436962 | Insomnia | 1 | Acute pancreatitis | Exposure | |
| 19027958 | fesoterodine | 1517070 | desmopressin | 36718449 | Urinary incontinence | 1 | Acute pancreatitis | Exposure | |
| 1588712 | Follicle Stimulating Hormone | 1536743 | ganirelix | 37119655 | Infertility female | 1 | Acute pancreatitis | Exposure | |
| 1784444 | Ivermectin | 922868 | Permethrin | 36110310 | Arthropod infestation | 1 | Acute pancreatitis | Exposure | |
| 1589795 | Luteinizing Hormone | 1563600 | Chorionic Gonadotropin | 37119655 | Infertility female | 1 | Acute pancreatitis | Exposure | |
| 915855 | olopatadine | 943634 | epinastine | 35607032 | Conjunctivitis allergic | 1 | Acute pancreatitis | Exposure | |
| 1110942 | omalizumab | 1105775 | Aminophylline | 37203741 | Bronchospasm and obstruction | 1 | Acute pancreatitis | Exposure | |
| 987366 | lubiprostone | 19025115 | picosulfate sodium | 35702117 | Gastrointestinal motility and defaecation conditions | 1 | Acute pancreatitis | Exposure | |
| 19025115 | picosulfate sodium | 924939 | Bisacodyl | 35702117 | Gastrointestinal motility and defaecation conditions | 1 | Acute pancreatitis | Exposure | |
| 945286 | Pilocarpine | 938044 | brinzolamide | 35606985 | Glaucoma | 1 | Acute pancreatitis | Exposure | |
| 40163718 | prasugrel | 1315865 | fondaparinux | 37622411 | Phlebosclerosis | 1 | Acute pancreatitis | Exposure | |
| 951279 | Prilocaine | 732893 | Bupivacaine | 438112 | Neoplastic disease | 1 | Acute pancreatitis | Exposure | |
| 781182 | ramelteon | 965748 | Scopolamine | 436962 | Insomnia | 1 | Acute pancreatitis | Exposure | |
| 1136487 | ropivacaine | 732893 | Bupivacaine | 37522270 | Surgery | 1 | Acute pancreatitis | Exposure | |
| 19012925 | silodosin | 989482 | Dutasteride | 37119607 | Benign prostatic hyperplasia | 1 | Acute pancreatitis | Exposure | |
| 1336926 | tadalafil | 918906 | oxybutynin | 37119607 | Benign prostatic hyperplasia | 1 | Acute pancreatitis | Exposure | |
| 1341238 | Terazosin | 989482 | Dutasteride | 37119607 | Benign prostatic hyperplasia | 1 | Acute pancreatitis | Exposure | |
| 1311276 | vardenafil | 1336926 | tadalafil | 36919202 | Sexual dysfunction | 1 | Acute pancreatitis | Exposure | |
| 1317967 | aliskiren | 1309799 | eplerenone | 37622411 | Phlebosclerosis | 2 | GI bleeding | Exposure | |
| 43013884 | alogliptin | 1502826 | nateglinide | 35506622 | Type 2 diabetes mellitus | 2 | GI bleeding | Exposure | |
| 742185 | atomoxetine | 705944 | Methylphenidate | 36918919 | Atypical attention deficit syndrome | 2 | GI bleeding | Exposure | |
| 937439 | Bethanechol | 759740 | Pyridostigmine | 35702117 | Gastrointestinal motility and defaecation conditions | 2 | GI bleeding | Exposure | |
| 938044 | brinzolamide | 717136 | Neostigmine | 35606985 | Glaucoma | 2 | GI bleeding | Exposure | |
| 732893 | Bupivacaine | 1136487 | ropivacaine | 37522270 | Surgery | 2 | GI bleeding | Exposure | |
| 1598819 | Clomiphene | 1563600 | Chorionic Gonadotropin | 37119655 | Infertility female | 2 | GI bleeding | Exposure | |
| 1103006 | desloratadine | 1136422 | levocetirizine | 36009773 | Rhinitis allergic | 2 | GI bleeding | Exposure | |
| 943634 | epinastine | 1136422 | levocetirizine | 36009773 | Rhinitis allergic | 2 | GI bleeding | Exposure | |
| 1588712 | Follicle Stimulating Hormone | 1536743 | ganirelix | 37119655 | Infertility female | 2 | GI bleeding | Exposure | |
| 19087394 | lacosamide | 750119 | Ethosuximide | 36718530 | Epilepsy | 2 | GI bleeding | Exposure | |
| 1136422 | levocetirizine | 1105889 | Triprolidine | 36009773 | Rhinitis allergic | 2 | GI bleeding | Exposure | |
| 1589795 | Luteinizing Hormone | 1563600 | Chorionic Gonadotropin | 37119655 | Infertility female | 2 | GI bleeding | Exposure | |
| 993979 | Malathion | 922868 | Permethrin | 36110310 | Arthropod infestation | 2 | GI bleeding | Exposure | |
| 734275 | Phenobarbital | 19087394 | lacosamide | 36718530 | Epilepsy | 2 | GI bleeding | Exposure | |
| 945286 | Pilocarpine | 938044 | brinzolamide | 35606985 | Glaucoma | 2 | GI bleeding | Exposure | |
| 951279 | Prilocaine | 732893 | Bupivacaine | 37522270 | Surgery | 2 | GI bleeding | Exposure | |
| 751347 | Primidone | 19087394 | lacosamide | 36702242 | Encephalopathies | 2 | GI bleeding | Exposure | |
| 781182 | ramelteon | 731188 | quazepam | 436962 | Insomnia | 2 | GI bleeding | Exposure | |
| 1136487 | ropivacaine | 1036884 | Tetracaine | 73553 | Arthropathy | 2 | GI bleeding | Exposure | |
| 965748 | Scopolamine | 781182 | ramelteon | 436962 | Insomnia | 2 | GI bleeding | Exposure | |
| 19012925 | silodosin | 989482 | Dutasteride | 37119607 | Benign prostatic hyperplasia | 2 | GI bleeding | Exposure | |
| 1236744 | Terbutaline | 1143374 | Ephedrine | 36009700 | Asthma | 2 | GI bleeding | Exposure | |
| 1036884 | Tetracaine | 702774 | Mepivacaine | 37522270 | Surgery | 2 | GI bleeding | Exposure | |
| 1111706 | zafirlukast | 1110942 | omalizumab | 36009700 | Asthma | 2 | GI bleeding | Exposure | |
| 1781733 | Atovaquone | 1702559 | Tinidazole | 36110538 | Arthropod-borne disease | 3 | Stroke | Exposure | |
| 1768849 | Cefaclor | 1796435 | Cefixime | 372328 | Otitis media | 3 | Stroke | Exposure | |
| 1769535 | Cefadroxil | 1778162 | Cefuroxime | 36102976 | Upper respiratory tract infections | 3 | Stroke | Exposure | |
| 1771162 | Cefazolin | 1778162 | Cefuroxime | 37203741 | Bronchospasm and obstruction | 3 | Stroke | Exposure | |
| 1796458 | cefdinir | 1778162 | Cefuroxime | 35406429 | Otitis media | 3 | Stroke | Exposure | |
| 1796435 | Cefixime | 1769535 | Cefadroxil | 35708603 | Pharyngitis | 3 | Stroke | Exposure | |
| 1774470 | Cefotaxime | 1778162 | Cefuroxime | 36102976 | Upper respiratory tract infections | 3 | Stroke | Exposure | |
| 1749008 | cefpodoxime | 1768849 | Cefaclor | 35406429 | Otitis media | 3 | Stroke | Exposure | |
| 1778162 | Cefuroxime | 1776684 | Ceftazidime | 36102971 | Lower respiratory tract and lung infections | 3 | Stroke | Exposure | |
| 954819 | cevimeline | 945286 | Pilocarpine | 35606954 | Sjogren's syndrome | 3 | Stroke | Exposure | |
| 1103006 | desloratadine | 1136422 | levocetirizine | 36009773 | Rhinitis allergic | 3 | Stroke | Exposure | |
| 1711246 | entecavir | 1703603 | famciclovir | 36102154 | Viral infectious disorders | 3 | Stroke | Exposure | |
| 19027958 | fesoterodine | 916005 | Solifenacin | 36718449 | Urinary incontinence | 3 | Stroke | Exposure | |
| 1763779 | Griseofulvin | 1741309 | terbinafine | 36102955 | Tinea infections | 3 | Stroke | Exposure | |
| 1784444 | Ivermectin | 922868 | Permethrin | 4085103 | Arthropod dermatosis | 3 | Stroke | Exposure | |
| 19087394 | lacosamide | 750119 | Ethosuximide | 36718530 | Epilepsy | 3 | Stroke | Exposure | |
| 1136422 | levocetirizine | 943634 | epinastine | 36009773 | Rhinitis allergic | 3 | Stroke | Exposure | |
| 1517740 | paricalcitol | 1512446 | Doxercalciferol | 35506745 | Hyperparathyroidism | 3 | Stroke | Exposure | |
| 19033354 | Petrolatum | 966956 | Silver Sulfadiazine | 37320089 | Skin dystrophy | 3 | Stroke | Exposure | |
| 19025115 | picosulfate sodium | 987366 | lubiprostone | 35702117 | Gastrointestinal motility and defaecation conditions | 3 | Stroke | Exposure | |
| 965748 | Scopolamine | 781182 | ramelteon | 436962 | Insomnia | 3 | Stroke | Exposure | |
| 1776684 | Ceftazidime | 1715117 | Aztreonam | 36102971 | Lower respiratory tract and lung infections | 3 | Stroke | Exposure | |
| 1715117 | Aztreonam | 902722 | Tobramycin | 35306246 | Cystic fibrosis | 3 | Stroke | Exposure | |
| 1782521 | isoniazid | 1749301 | Ethambutol | 36110782 | Tuberculosis | 3 | Stroke | Exposure | |
| 1790868 | Amikacin | 902722 | Tobramycin | 35306246 | Cystic fibrosis | 3 | Stroke | Exposure | |
| 742185 | atomoxetine | 705944 | Methylphenidate | 36918919 | Atypical attention deficit syndrome | 4 | Inflammatory Bowel Disease | Exposure | |
| 1338005 | Bisoprolol | 1386957 | Labetalol | 37622529 | Hypertension | 4 | Inflammatory Bowel Disease | Exposure | |
| 915542 | brimonidine | 717136 | Neostigmine | 35606985 | Glaucoma | 4 | Inflammatory Bowel Disease | Exposure | |
| 1598819 | Clomiphene | 1563600 | Chorionic Gonadotropin | 37119655 | Infertility female | 4 | Inflammatory Bowel Disease | Exposure | |
| 981709 | dorzolamide | 915542 | brimonidine | 35606985 | Glaucoma | 4 | Inflammatory Bowel Disease | Exposure | |
| 1363053 | Doxazosin | 913782 | tolterodine | 37622529 | Hypertension | 4 | Inflammatory Bowel Disease | Exposure | |
| 989482 | Dutasteride | 1341238 | Terazosin | 37119607 | Benign prostatic hyperplasia | 4 | Inflammatory Bowel Disease | Exposure | |
| 996416 | Finasteride | 989482 | Dutasteride | 37119607 | Benign prostatic hyperplasia | 4 | Inflammatory Bowel Disease | Exposure | |
| 1550023 | Insulin Lispro | 1502905 | Insulin Glargine | 35506622 | Type 2 diabetes mellitus | 4 | Inflammatory Bowel Disease | Exposure | |
| 1347384 | irbesartan | 1334456 | Ramipril | 37622529 | Hypertension | 4 | Inflammatory Bowel Disease | Exposure | |
| 1386957 | Labetalol | 1313200 | Nadolol | 37622529 | Hypertension | 4 | Inflammatory Bowel Disease | Exposure | |
| 989878 | Lidocaine | 732893 | Bupivacaine | 37522270 | Surgery | 4 | Inflammatory Bowel Disease | Exposure | |
| 1102527 | Meperidine | 965748 | Scopolamine | 36803580 | Delivery | 4 | Inflammatory Bowel Disease | Exposure | |
| 1114122 | Nalbuphine | 989878 | Lidocaine | 37522270 | Surgery | 4 | Inflammatory Bowel Disease | Exposure | |
| 753626 | Propofol | 708298 | Midazolam | 37522270 | Surgery | 4 | Inflammatory Bowel Disease | Exposure | |
| 1513103 | Raloxifene | 1598819 | Clomiphene | 35506596 | Polycystic ovaries | 4 | Inflammatory Bowel Disease | Exposure | |
| 965748 | Scopolamine | 704599 | Triazolam | 436962 | Insomnia | 4 | Inflammatory Bowel Disease | Exposure | |
| 916005 | Solifenacin | 913782 | tolterodine | 36718449 | Urinary incontinence | 4 | Inflammatory Bowel Disease | Exposure | |
| 924566 | tamsulosin | 916005 | Solifenacin | 4025202 | Elimination pattern | 4 | Inflammatory Bowel Disease | Exposure | |
| 1341238 | Terazosin | 924566 | tamsulosin | 37119607 | Benign prostatic hyperplasia | 4 | Inflammatory Bowel Disease | Exposure | |
| 902427 | Timolol | 717136 | Neostigmine | 35606985 | Glaucoma | 4 | Inflammatory Bowel Disease | Exposure | |
| 1524674 | zoledronic acid | 1512480 | Ibandronate | 36416551 | Osteoporosis | 4 | Inflammatory Bowel Disease | Exposure | |
| 708298 | Midazolam | 965748 | Scopolamine | 436962 | Insomnia | 4 | Inflammatory Bowel Disease | Exposure | |
| 1103640 | Methadone | 780442 | varenicline | 36919128 | Drug dependence | 4 | Inflammatory Bowel Disease | Exposure | |
| 781182 | ramelteon | 965748 | Scopolamine | 436962 | Insomnia | 4 | Inflammatory Bowel Disease | Exposure | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 440083 | Acute stress disorder | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 436677 | Adjustment disorder | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 433753 | Alcohol abuse | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 261880 | Atelectasis | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 4084966 | Candida infection of genital region | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 81250 | Carcinoma in situ of breast | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 381581 | Chalazion | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 198075 | Condyloma acuminatum | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 260134 | Croup | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 436373 | Developmental speech disorder | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 443767 | Diabetic oculopathy | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 72715 | Exostosis | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 78804 | Fibrocystic disease of breast | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 441788 | Human papilloma virus infection | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 139099 | Ingrowing nail | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 141216 | Molluscum contagiosum infection | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 134898 | Non-toxic uninodular goiter | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 195501 | Polycystic ovaries | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 373478 | Presbyopia | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 440690 | Social phobia | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 133141 | Tinea pedis | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 140641 | Verruca vulgaris | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 436070 | Vitamin D deficiency | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 440638 | Lyme disease | Outcomel | |
| 1124300 | Diclofenac | 1118084 | Celecoxib | 36516812 | Arthralgia | 433527 | Endometriosis | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 433753 | Alcohol abuse | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 81250 | Carcinoma in situ of breast | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 436373 | Developmental speech disorder | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 192367 | Dysplasia of cervix | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 72715 | Exostosis | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 78804 | Fibrocystic disease of breast | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 441788 | Human papilloma virus infection | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 381549 | Migraine with aura | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 135777 | Neoplasm of uncertain behavior of skin | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 134898 | Non-toxic uninodular goiter | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 433736 | Obesity | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 440374 | Obsessive-compulsive disorder | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 321596 | Peripheral venous insufficiency | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 373478 | Presbyopia | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 372409 | Sciatica | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 141932 | Seborrheic keratosis | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 313459 | Sleep apnea | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 440690 | Social phobia | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 140641 | Verruca vulgaris | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 378424 | Astigmatism | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 136661 | Non-toxic nodular goiter | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 437409 | Intracranial injury | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 73302 | Curvature of spine | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 4242416 | Cutis laxa | Outcomel | |
| 1797513 | Ciprofloxacin | 1734104 | Azithromycin | 35406429 | Otitis media | 375519 | Alcohol withdrawal syndrome | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 260139 | Acute bronchitis | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 256439 | Allergic rhinitis due to pollen | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 133834 | Atopic dermatitis | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 438409 | Attention deficit hyperactivity disorder | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 138102 | Benign neoplasm of skin | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 436665 | Bipolar disorder | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 257012 | Chronic sinusitis | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 134438 | Contact dermatitis | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 201606 | Crohn's disease | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 197610 | Cyst of ovary | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 194696 | Dysmenorrhea | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 78804 | Fibrocystic disease of breast | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 441788 | Human papilloma virus infection | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 139099 | Ingrowing nail | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 379805 | Myopia | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 135777 | Neoplasm of uncertain behavior of skin | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 440374 | Obsessive-compulsive disorder | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 380733 | Otalgia | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 372328 | Otitis media | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 4311499 | Primary malignant neoplasm of respiratory tract | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 141932 | Seborrheic keratosis | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 140641 | Verruca vulgaris | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 433753 | Alcohol abuse | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 201061 | Diaphragmatic hernia | Outcomel | |
| 1503297 | Metformin | 1560171 | Glipizide | 35506622 | Type 2 diabetes mellitus | 260134 | Croup | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 260139 | Acute bronchitis | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 257007 | Allergic rhinitis | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 4324765 | Arthropathy of knee joint | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 138102 | Benign neoplasm of skin | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 4084966 | Candida infection of genital region | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 257012 | Chronic sinusitis | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 260134 | Croup | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 201061 | Diaphragmatic hernia | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 73008 | Enthesopathy | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 133141 | Tinea pedis | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 439727 | Human immunodeficiency virus infection | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 441788 | Human papilloma virus infection | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 433811 | Hydronephrosis | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 201093 | Infection of skin AND/OR subcutaneous tissue | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 139099 | Ingrowing nail | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 198678 | Intestinal infectious disease | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 436659 | Iron deficiency anemia | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 374919 | Multiple sclerosis | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 135777 | Neoplasm of uncertain behavior of skin | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 321596 | Peripheral venous insufficiency | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 4311499 | Primary malignant neoplasm of respiratory tract | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 75650 | Proteinuria | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 137053 | Seborrheic dermatitis | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 141932 | Seborrheic keratosis | Outcomel | |
| 739138 | Sertraline | 743670 | Venlafaxine | 36918942 | Depression | 140641 | Verruca vulgaris | Outcomel | |