OHDSI Large-Scale Evidence Generation and Evaluation in a Network of Databases (LEGEND): Study of the Effects of Treatments for Depression

**Version:** 0.3

Martijn Schuemie, PhD, Janssen Research and Development

Patrick Ryan, PhD, Janssen Research and Development

Nicole Pratt, PhD, University of South Australia

Seng Chan You, MD, Ajou University, Korea

George Hripcsak, MD, Columbia University

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

**Date:** 11 September 2018

**Acknowledgment:** 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: Dr. Schuemie and Dr. Ryan are employees of Janssen Research & Development.

# Table of contents

[1 Table of contents 2](#_Toc524437264)

[2 List of abbreviations 3](#_Toc524437265)

[3 Abstract 4](#_Toc524437266)

[4 Amendments and Updates 4](#_Toc524437267)

[5 Milestones 4](#_Toc524437268)

[6 Rationale and Background 4](#_Toc524437269)

[7 Research Questions and Objectives 5](#_Toc524437270)

[7.1 Research Questions 5](#_Toc524437271)

[7.2 Objectives 6](#_Toc524437272)

[8 Research methods 6](#_Toc524437273)

[8.1 Study Design 6](#_Toc524437274)

[8.2 Data Source(s) 6](#_Toc524437275)

[8.2.1 Truven MarketScan Commercial Claims and Encounters (CCAE) 7](#_Toc524437276)

[8.2.2 Truven MarketScan Medicare Supplemental Beneficiaries (MDCR) 7](#_Toc524437277)

[8.2.3 Truven MarketScan Multi-state Medicaid (MDCD) 7](#_Toc524437278)

[8.2.4 Optum ClinFormatics (Optum) 7](#_Toc524437279)

[8.2.5 Clinical Practice Research Datalink (CPRD) 8](#_Toc524437280)

[8.2.6 QuintilesIMS Disease Analyzer (DA) Germany 8](#_Toc524437281)

[8.2.7 Japan Medical Data Center (JMDC) 8](#_Toc524437282)

[8.2.8 Optum® de-identified Electronic Health Record Dataset 9](#_Toc524437283)

[8.2.9 Korea National Health Insurance Service (NHIS) National Sample Cohort (NSC) 9](#_Toc524437284)

[8.3 Study population 9](#_Toc524437285)

[8.3.1 Subgroups 9](#_Toc524437286)

[8.4 Exposures 10](#_Toc524437287)

[8.4.1 All drugs 10](#_Toc524437288)

[8.4.2 Psychotherapy 11](#_Toc524437289)

[8.4.3 Electroconvulsive therapy 16](#_Toc524437290)

[8.5 Outcomes 18](#_Toc524437291)

[8.5.1 Acute liver injury 18](#_Toc524437292)

[8.5.2 Acute myocardial infarction 19](#_Toc524437293)

[8.5.3 Alopecia 20](#_Toc524437294)

[8.5.4 Constipation 20](#_Toc524437295)

[8.5.5 Decreased libido 21](#_Toc524437296)

[8.5.6 Delirium 21](#_Toc524437297)

[8.5.7 Diarrhea 22](#_Toc524437298)

[8.5.8 Fracture 22](#_Toc524437299)

[8.5.9 Gastrointestinal hemorrhage 23](#_Toc524437300)

[8.5.10 Hyperprolactinemia 26](#_Toc524437301)

[8.5.11 Hyponatremia 26](#_Toc524437302)

[8.5.12 Hypotension 27](#_Toc524437303)

[8.5.13 Hypothyroidism 27](#_Toc524437304)

[8.5.14 Insomnia 28](#_Toc524437305)

[8.5.15 Nausea 28](#_Toc524437306)

[8.5.16 Open-angle glaucoma 29](#_Toc524437307)

[8.5.17 Unconfirmed open-angle glaucoma 29](#_Toc524437308)

[8.5.18 Seizure 30](#_Toc524437309)

[8.5.19 Stroke 31](#_Toc524437310)

[8.5.20 Suicide and suicidal ideation 32](#_Toc524437311)

[8.5.21 Tinnitus 32](#_Toc524437312)

[8.5.22 Ventricular arrhythmia and sudden cardiac death 33](#_Toc524437313)

[8.5.23 Vertigo 33](#_Toc524437314)

[8.5.24 Negative control outcomes 35](#_Toc524437315)

[8.5.25 Positive control outcomes 36](#_Toc524437316)

[8.6 Covariates 36](#_Toc524437317)

[8.6.1 Propensity score covariates 36](#_Toc524437318)

[9 Data Analysis Plan 37](#_Toc524437319)

[9.1 Calculation of time-at risk 37](#_Toc524437320)

[9.2 Model Specification 37](#_Toc524437321)

[9.2.1 Pooling effect estimates across databases 38](#_Toc524437322)

[9.3 Analyses to perform 38](#_Toc524437323)

[9.3.1 Comparative analyses 38](#_Toc524437324)

[9.3.2 Descriptive analyses 38](#_Toc524437325)

[9.4 Output 39](#_Toc524437326)

[9.5 Evidence Evaluation 39](#_Toc524437327)

[10 Study Diagnostics 40](#_Toc524437328)

[10.1 Sample Size and Study Power 40](#_Toc524437329)

[10.2 Cohort Comparability 40](#_Toc524437330)

[10.3 Systematic Error Assessment 40](#_Toc524437331)

[11 Strengths and Limitations of the Research Methods 40](#_Toc524437332)

[12 Protection of Human Subjects 40](#_Toc524437333)

[13 Management and Reporting of Adverse Events and Adverse Reactions 40](#_Toc524437334)

[14 Plans for Disseminating and Communicating Study Results 41](#_Toc524437335)

[15 References 41](#_Toc524437336)

[Appendix A. Differences with previous depression study 42](#_Toc524437337)

# List of abbreviations

ATC Anatomic Therapeutic Chemical

CYCLOPS Cyclic coordinate descent for logistic, Poisson and survival analysis

SNOMED Systematized Nomenclature of Medicine

OHDSI Observational Health Data Sciences and Informatics

OMOP Observational Medical Outcomes Partnership

T Target cohort

C Comparator cohort

O Outcome cohort

PS Propensity Scores

LASSO Least absolute shrinkage and selection operator

CI Confidence Interval

ECT Electroconvulsive therapy

MedDRA Medical Dictionary for Regulatory Activities

NDRI Norepinephrine-dopamine reuptake inhibitors

SARI Serotonin antagonist and reuptake inhibitor

SNRI Serotonin–norepinephrine reuptake inhibitor

SSRI Selective serotonin reuptake inhibitor

TCA Tricyclic antidepressant

# Abstract

In this study we will generate population-level estimates at scale for one disease: depression. We perform every possible pairwise comparison between depression treatments for a large set of outcomes of interest. Most of these outcomes are generic safety outcomes, but some outcomes are related more specifically to the effectiveness of antidepressant treatment.

# Amendments and Updates

|  |  |  |  |
| --- | --- | --- | --- |
| 0.1 | 17 August 2018 | M.Schuemie | First draft |
| 0.2 | 5 September 2018 | M. Schuemie, N. Pratt, S.C. You, P. Ryan | Updated indication and outcome definitions. Added section on changes with previous study. |
| 0.3 | 11 September 2018 | M. Schuemie, P. Ryan | Updated several additional outcome definitions to match those used in hypertension study. |
|  |  |  |  |

# Milestones

|  |  |
| --- | --- |
| Milestone | Planned / Estimated Date |
| Start of analysis | 9 September 2018 |
| End of analysis | 30 September 2018 |
| Presentation of results | 11 October 2018 |

# Rationale and Background

The Large-scale Evidence Generation and Evaluation in a Network of Databases (LEGEND) project aims to generate reliable evidence on the effects of medical interventions using observational healthcare data to support clinical decision making. LEGEND follows ten guiding principles (see Supplementary Material); chief among these stand that we generate evidence at large-scale to achieve completeness and facilitate analysis of the overall distribution of effect size estimates across treatments and outcomes. We also generate evidence consistently by applying a systematic approach across all research questions and disseminate evidence regardless on the estimates effects to avoid publication bias. These aims help overcome the questionable reliable of observational research [[1](#_ENREF_1)].

In this study we will generate population-level estimates at scale for one disease: depression. We perform every possible pairwise comparison between depression treatments for a large set of outcomes of interest. Most of these outcomes are generic safety outcomes, but some outcomes are related more specifically to the effectiveness of antidepressant treatment. This study is similar to our previous study into the effects of depression treatments [[1](#_ENREF_1)]. Changes since the previous study and their rationales are described in Appendix A.

# Research Questions and Objectives

## Research Questions

In this study, we are interested in every pairwise comparison between any two treatments in table 1. Treatments will be compared at the treatment level (e.g. comparing bupropion to mirtazapine), but also at the class level (e.g. SSRIs versus SNRIs).

|  |  |  |
| --- | --- | --- |
| **Type** | **Class** | **Treatment** |
| Drug | Norepinephrine-dopamine reuptake inhibitor (NDRI) | Bupropion |
| Drug | Tetracyclic antidepressant | Mirtazapine |
| Procedure | ECT | Electroconvulsive therapy |
| Procedure | Psychotherapy | Psychotherapy |
| Drug | Serotonin antagonist and reuptake inhibitor (SARI) | Trazodone |
| Drug | Serotonin-norepinephrine reuptake inhibitor (SNRI) | Desvenlafaxine |
| Drug | Serotonin-norepinephrine reuptake inhibitor (SNRI) | Duloxetine |
| Drug | Serotonin-norepinephrine reuptake inhibitor (SNRI) | Venlafaxine |
| Drug | Selective serotonin reuptake inhibitor (SSRI) | Citalopram |
| Drug | Selective serotonin reuptake inhibitor (SSRI) | Escitalopram |
| Drug | Selective serotonin reuptake inhibitor (SSRI) | Fluoxetine |
| Drug | Selective serotonin reuptake inhibitor (SSRI) | Paroxetine |
| Drug | Selective serotonin reuptake inhibitor (SSRI) | Sertraline |
| Drug | Selective serotonin reuptake inhibitor (SSRI) | vilazodone |
| Drug | Tricyclic antidepressant (TCA) | Amitriptyline |
| Drug | Tricyclic antidepressant (TCA) | Doxepin |
| Drug | Tricyclic antidepressant (TCA) | Nortriptyline |

**Table 1**. List of depression treatments considered in this study

For each comparison of two treatments, we are interested in the comparative effect on each of the outcomes listed in table 2.

|  |  |
| --- | --- |
| Acute liver injury | Hypotension |
| Acute myocardial infarction | Hypothyroidism |
| Alopecia | Insomnia |
| Constipation | Nausea |
| Decreased libido | Open-angle glaucoma |
| Delirium | Unconfirmed open-angle glaucoma |
| Diarrhea | Seizure |
| Fracture | Stroke |
| Gastrointestinal hemorrhage | Suicide and suicidal ideation |
| Hyperprolactinemia | Tinnitus |
| Hyponatremia | Ventricular arrhythmia and sudden cardiac death |
|  | Vertigo |

**Table 2.** Outcomes of interest considered in this study

Primary research question

* For each comparison between two depression treatments, for each of the outcomes of interest, what is the hazard ratio?

We further consider the six following subgroups of interest:

* Renal impairment
* Hepatic impairment
* Pregnant women
* Children (age < 18)
* Elderly (age >=65)
* Gender = female

Secondary research questions

* For each comparison between two depression treatments, for each of the outcomes of interest, how does the hazard ratio change within 6 subgroups of interest?
* What is the incidence rate of each outcome of interest in each exposure group?

## Objectives

Primary objective

* Generate evidence for the comparative effectiveness for each pairwise comparison of depression treatments for the outcomes of interest.

Secondary objectives

* Asses the bias inherent in each analysis by including negative and positive control outcomes.

# Research methods

## Study Design

This study will be a set of retrospective, observational, new-user cohort studies. By ‘retrospective’ we mean the study will use data already collected at the start of the study. By ‘observational’ we mean no intervention will take place in the course of this study. By ‘new-user’ we mean we will only analyze the first exposure of a subject to the treatment of interest. By ‘cohort study’ we mean two cohorts, a target and comparator cohort, will be followed from index date (start of first exposure) to some end date, and assessed for the occurrence of the outcomes of interest.

## Data Source(s)

The analyses will be performed across a network of observational healthcare databases. All databases have been transformed into the OMOP Common Data Model, version 4 or OMOP Common Data Model, version 5. The complete specification for OMOP Common Data Model, version 4 is available at: <http://omop.org/cdm>. The complete specification for OMOP Common Data Model, version 5 is available at: <https://github.com/OHDSI/CommonDataModel>. The following databases will be included in this analysis:

* Truven MarketScan Commercial Claims and Encounters (CCAE)
* Truven MarketScan Medicare Supplemental Beneficiaries (MDCR)
* Truven MarketScan Multi-state Medicaid (MDCD)
* Optum ClinFormatics (Optum)
* Clinical Practice Research Datalink (CPRD)
* QuintilesIMS Disease Analyzer (DA) Germany
* Japan Medical Data Center (JMDC)
* Optum® de-identified Electronic Health Record Dataset
* Korea National Health Insurance Service (NHIS) National Sample Cohort (NSC)
* <<add others who agree to participate>>

### Truven MarketScan Commercial Claims and Encounters (CCAE)

Truven Health MarketScan® Commercial Claims and Encounters Database (CCAE) represent data from individuals enrolled in United States employer-sponsored insurance health plans. 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.

### Truven MarketScan Medicare Supplemental Beneficiaries (MDCR)

Truven Health MarketScan® Medicare Supplemental and Coordination of Benefits 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.

### Truven MarketScan Multi-state Medicaid (MDCD)

Truven Health MarketScan® Multi-State Medicaid Database (MDCD) adjudicated US health insurance claims for Medicaid enrollees from multiple states and includes hospital discharge diagnoses, outpatient diagnoses and procedures, and outpatient pharmacy claims as well as ethnicity and Medicare eligibility. Members maintain their same identifier even if they leave the system for a brief period however the dataset lacks lab data. [For further information link to RWE site for Truven MDCD.

### Optum ClinFormatics (Optum)

Optum Clinformatics Extended DataMart is an adjudicated US administrative health claims database for members of private health insurance, who are fully insured in commercial plans or in administrative services only (ASOs), Legacy Medicare Choice Lives (prior to January 2006), and Medicare Advantage (Medicare Advantage Prescription Drug coverage starting January 2006). The population is primarily representative of commercial claims patients (0-65 years old) with some Medicare (65+ years old) however ages are capped at 90 years. It includes data captured from administrative claims processed from inpatient and outpatient medical services and prescriptions as dispensed, as well as results for outpatient lab tests processed by large national lab vendors who participate in data exchange with Optum. This dataset also provides date of death (month and year only) for members with both medical and pharmacy coverage from the Social Security Death Master File (however after 2011 reporting frequency changed due to changes in reporting requirements) and location information for patients is at the US state level.

### Clinical Practice Research Datalink (CPRD)

The Clinical Practice Research Datalink (CPRD) is a governmental, not-for-profit research service, jointly funded by the NHS National Institute for Health Research (NIHR) and the Medicines and Healthcare products Regulatory Agency (MHRA), a part of the Department of Health, United Kingdom (UK). CPRD consists of data collected from UK primary care for all ages. This includes conditions, observations, measurements, and procedures that the general practitioner is made aware of in additional to any prescriptions as prescribed by the general practitioner. In addition to primary care, there are also linked secondary care records for a small number of people.\nThe major data elements contained within this database are outpatient prescriptions given by the general practitioner (coded with Multilex codes) and outpatient clinical, referral, immunization or test events that the general practitioner knows about (coded in Read or ICD10 or LOINC codes). The database also contains the patients’ year of births and any date of deaths.

### QuintilesIMS Disease Analyzer (DA) Germany

The QuintilesIMS Disease Analyzer (DA) Germany database consists of data collected from physician practices and medical centers for all ages. Mostly primary care physician data however some data from specialty practices (where practices are electronically connected to each other) and some lab data is included. Key attributes include demographics, prescriptions as prescribed at brand level, diagnosis, lab measurements, actions (e.g. referrals, sick notes).

### Japan Medical Data Center (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 as associated with a full date.\nAll diagnoses are coded using ICD-10. All prescriptions refer to national Japanese drug codes, which have been linked to ATC. Procedures are encoded using local procedure codes, which the vendor has mapped to ICD-9 procedure codes. The annual health checkups report a standard battery of measurements (e.g. BMI), which are not coded but clearly described.

### Optum® de-identified Electronic Health Record Dataset

Optum© de-identified Electronic Health Record Dataset represents Humedica’s Electronic Health Record data a medical records database. The medical record data includes 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).

### Korea National Health Insurance Service (NHIS) National Sample Cohort (NSC)

Todo: add description

## Study population

All subjects in the database will be included who meet the following criteria: (note: the index date is the start of the first treatment for depression)

* Exposure to one of the treatments of interest
* At least 365 days of observation time prior to the index date
* No exposure to the two treatments in each pairwise comparison before the index date
* A diagnose of depressive disorder on or preceding the index date
* No diagnosis of bipolar disorder or schizophrenia on or preceding the index date

No diagnose of the outcome of interest preceding the index date

### Subgroups

Interaction effects will be estimates with the following subgroups:

* Renal impairment
* Hepatic impairment
* Pregnant women
* Children (age < 18)
* Elderly (age >=65)
* Gender = female

**Renal impairment**

Having any of the following concepts observed in the 365 days prior to the index date:

|  |  |  |  |
| --- | --- | --- | --- |
| **Concept ID** | **Concept Name** | **Domain** | **Vocabulary** |
| 4090651 | Dialysis finding | Observation | SNOMED |
| 4032243 | Dialysis procedure | Procedure | SNOMED |
| 45889365 | Dialysis Services and Procedures | Procedure | CPT4 |
| 2786488 | Irrigation of Peritoneal Cavity using Dialysate, Percutaneous Approach | Procedure | ICD10PCS |
| 2788041 | Performance of Urinary Filtration, Single | Procedure | ICD10PCS |
| 4030518 | Renal impairment | Condition | SNOMED |

**Hepatic impairment**

Having any of the following concepts observed in the 365 days prior to the index date:

|  |  |  |  |
| --- | --- | --- | --- |
| **Concept ID** | **Concept Name** | **Domain** | **Vocabulary** |
| 200763 | Chronic hepatitis | Condition | SNOMED |
| 4064161 | Cirrhosis of liver | Condition | SNOMED |
| 4245975 | Hepatic failure | Condition | SNOMED |
| 4337543 | Hepatic necrosis | Condition | SNOMED |
| 4055224 | Toxic liver disease | Condition | SNOMED |

**Pregnant women**

Todo: add description

**Children (age < 18)**

Defined as having index year – year of birth < 18.

**Elderly (age >=65)**

Defined as having index year – year of birth >= 65.

**Gender = female**

Defined as having gender = female (concept ID 8532).

## Exposures

In this study, we are interested in every pairwise comparison between any two treatments in table 1. Treatments will be compared at the treatment level (e.g. comparing bupropion to mirtazapine), but also at the class level (e.g. SSRIs versus SNRIs).

### All drugs

Index rule defining the index date:

* First exposure to any drug containing the RxNorm ingredient of interest.

Inclusion rules based on the index date:

* At least 365 days of observation time prior to the index date
* No exposure to the target or comparator ingredient or procedure before the index date
* A diagnose of depressive disorder on or preceding the index date
* No diagnose of bipolar disorder or schizophrenia on or preceding the index date
* No diagnose of the outcome of interest preceding the index date

Note that prior exposure to other depression treatments prior to the index date is allowed. For example, when comparing sertraline to duloxetine, sertraline users are not allowed to have prior exposure so sertraline or duloxetine but are allowed to have prior exposure to fluoxetine.

The end of the exposure cohort is defined as the end of the first exposure, allowing for 30-day gaps between consecutive prescriptions.

### Psychotherapy

Initial Event Cohort

People having any of the following:

* a procedure of Psychotherapy2

with continuous observation of at least 365 days prior and 0 days after event index date, and limit initial events to: **earliest event per person.**

Inclusion Criteria #1: Prior depressive disorder

People having all of the following criteria:

* at least 1 occurrences of a condition occurrence of depressive disorder1 starting between all days Before and 0 days After event index date

Inclusion Criteria #2: No prior bipolar disorder or schizophrenia

People having all of the following criteria:

* at most 0 occurrences of a condition occurrence of Schizophrenia and bipolar disorder3 starting between all days Before and 0 days After event index date

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

The cohort end date is defined as 30 days following the cohort start date. If new psychotherapy is observed before those 30 days are over, the exposure cohort is extended to 30 days after the new psychotherapy.

Appendix 1: Concept Set Definitions

1.  depressive disorder

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 440383 | depressive disorder | Condition | SNOMED | NO | YES | NO |

2. Psychotherapy

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 4119335 | Analytical psychology | Procedure | SNOMED | NO | NO | NO |
| 4084202 | Anti-criminal psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4079608 | Anti-suicide psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4048385 | Brief group psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4295027 | Brief solution focused psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4299728 | Client-centered psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4164790 | Conjoint psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4208314 | Couple psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4083706 | Crisis intervention | Procedure | SNOMED | NO | NO | NO |
| 4083131 | Daily life psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4121662 | Developmental psychodynamic psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4226276 | Eclectic psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4258834 | Educational psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4148765 | Encounter group therapy | Procedure | SNOMED | NO | NO | NO |
| 2007747 | Exploratory verbal psychotherapy | Procedure | ICD9Proc | NO | NO | NO |
| 4137086 | Expressed emotion family therapy | Procedure | SNOMED | NO | NO | NO |
| 4048387 | Expressive psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4173581 | Extended family therapy | Procedure | SNOMED | NO | NO | NO |
| 46286403 | Family intervention for psychosis | Procedure | SNOMED | NO | NO | NO |
| 2213546 | Family psychotherapy (conjoint psychotherapy) (with patient present) | Procedure | CPT4 | NO | NO | NO |
| 4028920 | Family psychotherapy procedure | Procedure | SNOMED | NO | NO | NO |
| 46286330 | Focal psychodynamic therapy | Procedure | SNOMED | NO | NO | NO |
| 4226275 | Formal psychological therapy | Procedure | SNOMED | NO | NO | NO |
| 45765516 | Functional family therapy | Procedure | SNOMED | NO | NO | NO |
| 4079939 | Functional psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4079500 | General psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4117915 | Generic Jungian-based therapy | Procedure | SNOMED | NO | NO | NO |
| 4100341 | Group analytical psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 44808677 | Group cognitive behavioural therapy | Procedure | SNOMED | NO | NO | NO |
| 4136352 | Group marathon therapy | Procedure | SNOMED | NO | NO | NO |
| 4268909 | Group primal therapy | Procedure | SNOMED | NO | NO | NO |
| 4296166 | Group psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 2213548 | Group psychotherapy (other than of a multiple-family group) | Procedure | CPT4 | NO | NO | NO |
| 2617477 | Group psychotherapy other than of a multiple-family group, in a partial hospitalization setting, approximately 45 to 50 minutes | Observation | HCPCS | NO | NO | NO |
| 4196062 | Group reassurance | Procedure | SNOMED | NO | NO | NO |
| 2213554 | Individual psychophysiological therapy incorporating biofeedback training by any modality (face-to-face with the patient), with psychotherapy (eg, insight oriented, behavior modifying or supportive psychotherapy); 30 minutes | Procedure | CPT4 | NO | NO | NO |
| 2213555 | Individual psychophysiological therapy incorporating biofeedback training by any modality (face-to-face with the patient), with psychotherapy (eg, insight oriented, behavior modifying or supportive psychotherapy); 45 minutes | Procedure | CPT4 | NO | NO | NO |
| 2007730 | Individual psychotherapy | Procedure | ICD9Proc | NO | NO | NO |
| 4088889 | Individual psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4103512 | Interactive group medical psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 2617478 | Interactive group psychotherapy, in a partial hospitalization setting, approximately 45 to 50 minutes | Observation | HCPCS | NO | NO | NO |
| 4221997 | Interactive individual medical psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 40482841 | Interpersonal psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4119334 | Jungian-based therapy | Procedure | SNOMED | NO | NO | NO |
| 4118797 | Long-term exploratory psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4118798 | Long-term psychodynamic psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 44792695 | Marital psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 2213547 | Multiple-family group psychotherapy | Procedure | CPT4 | NO | NO | NO |
| 4118800 | Narrative family psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4242119 | Occupational social therapy | Observation | SNOMED | NO | NO | NO |
| 2007749 | Other individual psychotherapy | Procedure | ICD9Proc | NO | NO | NO |
| 2007750 | Other psychotherapy and counselling | Procedure | ICD9Proc | NO | NO | NO |
| 45887728 | Other Psychotherapy Procedures | Procedure | CPT4 | NO | NO | NO |
| 45763911 | Parent-infant psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 2007746 | Play psychotherapy | Procedure | ICD9Proc | NO | NO | NO |
| 4083133 | Potential suicide care | Procedure | SNOMED | NO | NO | NO |
| 4084195 | Provocative therapy | Procedure | SNOMED | NO | NO | NO |
| 2213544 | Psychoanalysis | Procedure | CPT4 | NO | NO | NO |
| 2007731 | Psychoanalysis | Procedure | ICD9Proc | NO | NO | NO |
| 4114491 | Psychoanalytic and psychodynamic therapy | Procedure | SNOMED | NO | NO | NO |
| 2007763 | Psychodrama | Procedure | ICD9Proc | NO | NO | NO |
| 4202234 | Psychodrama | Procedure | SNOMED | NO | NO | NO |
| 4199042 | Psychodynamic psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4128268 | Psychodynamic-interpersonal psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4118801 | Psychotherapeutic approaches using specific settings | Procedure | SNOMED | NO | NO | NO |
| 4327941 | Psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4083129 | Psychotherapy - behavioral | Procedure | SNOMED | NO | NO | NO |
| 4079938 | Psychotherapy - cognitive | Procedure | SNOMED | NO | NO | NO |
| 45889353 | Psychotherapy for crisis | Procedure | CPT4 | NO | NO | NO |
| 45888237 | Psychotherapy for Crisis Services and Procedures | Procedure | CPT4 | NO | NO | NO |
| 43527991 | Psychotherapy for crisis; each additional 30 minutes (List separately in addition to code for primary service) | Procedure | CPT4 | NO | NO | NO |
| 43527990 | Psychotherapy for crisis; first 60 minutes | Procedure | CPT4 | NO | NO | NO |
| 45887951 | Psychotherapy Services and Procedures | Procedure | CPT4 | NO | NO | NO |
| 2108571 | Psychotherapy services provided (MDD, MDD ADOL) | Observation | CPT4 | NO | NO | NO |
| 43527986 | Psychotherapy, 30 minutes with patient and/or family member | Procedure | CPT4 | NO | NO | NO |
| 43527987 | Psychotherapy, 30 minutes with patient and/or family member when performed with an evaluation and management service (List separately in addition to the code for primary procedure) | Procedure | CPT4 | NO | NO | NO |
| 43527904 | Psychotherapy, 45 minutes with patient and/or family member | Procedure | CPT4 | NO | NO | NO |
| 43527988 | Psychotherapy, 45 minutes with patient and/or family member when performed with an evaluation and management service (List separately in addition to the code for primary procedure) | Procedure | CPT4 | NO | NO | NO |
| 43527905 | Psychotherapy, 60 minutes with patient and/or family member | Procedure | CPT4 | NO | NO | NO |
| 43527989 | Psychotherapy, 60 minutes with patient and/or family member when performed with an evaluation and management service (List separately in addition to the code for primary procedure) | Procedure | CPT4 | NO | NO | NO |
| 4148398 | Psychotherapy/sociotherapy | Procedure | SNOMED | NO | NO | NO |
| 4083130 | Rehabilitation for disabling psychiatric problem | Procedure | SNOMED | NO | NO | NO |
| 44791916 | Relationship psychosexual therapy | Procedure | SNOMED | NO | NO | NO |
| 4265313 | Relationship psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4084201 | Samaritans advisory service | Procedure | SNOMED | NO | NO | NO |
| 4233181 | Sensate focus technique | Procedure | SNOMED | NO | NO | NO |
| 4272803 | Sexual psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4035812 | Sexual psychotherapy, female therapist - female patient | Procedure | SNOMED | NO | NO | NO |
| 4012488 | Sexual psychotherapy, female therapist - male patient | Procedure | SNOMED | NO | NO | NO |
| 4132436 | Sexual psychotherapy, group | Procedure | SNOMED | NO | NO | NO |
| 4143316 | Sexual psychotherapy, group, all female | Procedure | SNOMED | NO | NO | NO |
| 4219683 | Sexual psychotherapy, group, all male | Procedure | SNOMED | NO | NO | NO |
| 4151904 | Sexual psychotherapy, group, male and female | Procedure | SNOMED | NO | NO | NO |
| 4278094 | Sexual psychotherapy, male therapist - female patient | Procedure | SNOMED | NO | NO | NO |
| 4249602 | Sexual psychotherapy, male therapist - male patient | Procedure | SNOMED | NO | NO | NO |
| 4234476 | Sexual surrogate therapy | Procedure | SNOMED | NO | NO | NO |
| 4179241 | Short-term psychodynamic therapy | Procedure | SNOMED | NO | NO | NO |
| 4234402 | Social psychotherapy | Observation | SNOMED | NO | NO | NO |
| 4128406 | Specific task orientated psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4080044 | Stimulative psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4262582 | Structural family psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4263758 | Structural psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4126653 | Supportive expressive psychodynamic psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 2007748 | Supportive verbal psychotherapy | Procedure | ICD9Proc | NO | NO | NO |
| 4311943 | Supportive verbal psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4225728 | Suppressive psychotherapy | Procedure | SNOMED | NO | NO | NO |
| 4080048 | Therapeutic psychology | Procedure | SNOMED | NO | NO | NO |
| 44808259 | Therapeutic role play | Procedure | SNOMED | NO | NO | NO |

3. Schizophrenia and bipolar disorder

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 436665 | Bipolar disorder | Condition | SNOMED | NO | YES | NO |
| 435783 | Schizophrenia | Condition | SNOMED | NO | YES | NO |

### Electroconvulsive therapy

Initial Event Cohort

People having any of the following:

* a procedure of Electroconvulsive therapy1

with continuous observation of at least 365 days prior and 0 days after event index date, and limit initial events to: **earliest event per person.**

Inclusion Criteria #1: Prior depressive disorder

People having all of the following criteria:

* at least 1 occurrences of a condition occurrence of depressive disorder2 starting between all days Before and 0 days After event index date

Inclusion Criteria #2: No prior bipolar disorder or schizophrenia

People having all of the following criteria:

* at most 0 occurrences of a condition occurrence of Schizophrenia and bipolar disorderstarting between all days Before and 0 days After event index date

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

The cohort end date is defined as 30 days following the cohort start date. If new psychotherapy is observed before those 30 days are over, the exposure cohort is extended to 30 days after the new psychotherapy.

Appendix 1: Concept Set Definitions

1. Electroconvulsive therapy

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 4111663 | Bilateral electroconvulsive therapy | Procedure | SNOMED | NO | NO | NO |
| 4030840 | Electroconvulsive therapy | Procedure | SNOMED | NO | NO | NO |
| 2108578 | Electroconvulsive therapy (ECT) provided (MDD) | Observation | CPT4 | NO | NO | NO |
| 2213552 | Electroconvulsive therapy (includes necessary monitoring) | Procedure | CPT4 | NO | NO | NO |
| 4020981 | Electronarcosis | Procedure | SNOMED | NO | NO | NO |
| 4210144 | First treatment in a course of electroconvulsive therapy | Procedure | SNOMED | NO | NO | NO |
| 4336318 | Multiple electroconvulsive therapy | Procedure | SNOMED | NO | NO | NO |
| 4332436 | Multiple monitored electroconvulsive therapy | Procedure | SNOMED | NO | NO | NO |
| 2007728 | Other electroshock therapy | Procedure | ICD9Proc | NO | NO | NO |
| 44508134 | Other specified electroconvulsive therapy | Procedure | OPCS4 | NO | NO | NO |
| 2108579 | Patient referral for electroconvulsive therapy (ECT) documented (MDD) | Observation | CPT4 | NO | NO | NO |
| 2007727 | Subconvulsive electroshock therapy | Procedure | ICD9Proc | NO | NO | NO |
| 4004830 | Subconvulsive electroshock therapy | Procedure | SNOMED | NO | NO | NO |
| 4210145 | Subsequent treatment in a course of electroconvulsive therapy | Procedure | SNOMED | NO | NO | NO |

2.  depressive disorder

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 440383 | depressive disorder | Condition | SNOMED | NO | YES | NO |

3. Schizophrenia and bipolar disorder

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 436665 | Bipolar disorder | Condition | SNOMED | NO | YES | NO |
| 435783 | Schizophrenia | Condition | SNOMED | NO | YES | NO |

## Outcomes

### Acute liver injury

Note: This algorithm uses the set of codes identified by Udo et al. [[2](#_ENREF_2)]

Initial Event Cohort

People having any of the following:

* a condition occurrence of acute liver injury1
  + for the first time in the person's history
  + 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.**

For people matching the Primary Events, include:

Having all of the following criteria:

* exactly 0 occurrences of a condition occurrence of acute liver injury exclusion concepts2

starting between 365 days Before and 60 days After event index date

Limit cohort of initial events to: **earliest event per person.**

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

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

Appendix 1: Concept Set Definitions

1. acute liver injury

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 200763 | Chronic hepatitis | Condition | SNOMED | YES | YES | NO |
| 377604 | Hepatic coma | Condition | SNOMED | NO | YES | NO |
| 196029 | Hepatic coma due to viral hepatitis | Condition | SNOMED | YES | YES | NO |
| 4337543 | Hepatic necrosis | Condition | SNOMED | NO | YES | NO |
| 194087 | Hepatitis due to infection | Condition | SNOMED | YES | YES | NO |
| 196455 | Hepatorenal syndrome | Condition | SNOMED | NO | YES | NO |
| 194990 | Inflammatory disease of liver | Condition | SNOMED | NO | YES | NO |
| 4291005 | Viral hepatitis | Condition | SNOMED | YES | YES | NO |

2. acute liver injury exclusion concepts

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 192956 | Cholecystitis | Condition | SNOMED | NO | YES | NO |
| 200763 | Chronic hepatitis | Condition | SNOMED | NO | YES | NO |
| 4212540 | Chronic liver disease | Condition | SNOMED | NO | YES | NO |
| 197917 | Disorder of biliary tract | Condition | SNOMED | NO | YES | NO |
| 192353 | Disorder of gallbladder | Condition | SNOMED | NO | YES | NO |
| 192963 | Disorder of pancreas | Condition | SNOMED | NO | YES | NO |
| 196456 | Gallstone | Condition | SNOMED | NO | YES | NO |
| 4130518 | Neoplasm of liver | Condition | SNOMED | NO | YES | NO |
| 4291005 | Viral hepatitis | Condition | SNOMED | NO | YES | NO |

### Acute **myocardial** infarction

Initial Event Cohort

People having any of the following:

* a condition occurrence of Acute myocardial Infarction2

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

For people matching the Primary Events, include:

Having any of the following criteria:

* at least 1 occurrences of a visit occurrence of Inpatient or ER visit1

where event starts between all days Before and 0 days After index start date and event ends between 0 days Before and all days After index start date

Limit cohort of initial events to: **all events per person.**

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

End Date Strategy

Date Offset Exit Criteria

This cohort definition end date will be the index event's start date plus 7 days

Cohort Collapse Strategy:

Collapse cohort by era with a gap size of 180 days.

Appendix 1: Concept Set Definitions

1. Inpatient or ER visit

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| 262 | Emergency Room and Inpatient Visit | Visit | Visit | NO | YES | NO |
| 9203 | Emergency Room Visit | Visit | Visit | NO | YES | NO |
| 9201 | Inpatient Visit | Visit | Visit | NO | YES | NO |

2. Acute myocardial Infarction

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| 4329847 | Myocardial infarction | Condition | SNOMED | NO | YES | NO |
| 314666 | Old myocardial infarction | Condition | SNOMED | YES | YES | NO |

### Alopecia

Initial Event Cohort

People having any of the following: 

* a condition occurrence of Alopecia1
  + for the first time in the person's history

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.**

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

Appendix 1: Concept Set Definitions

1. Alopecia

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 133280 | Alopecia | Condition | SNOMED | NO | YES | NO |
| 133959 | Syphilitic alopecia | Condition | SNOMED | YES | YES | NO |

### Constipation

Note: This algorithm requires the occurrence of 2 or more diagnoses, as recommended by Mody et al. [[3](#_ENREF_3)]

Initial Event Cohort

People having any of the following: 

* a condition occurrence of Constipation1
  + for the first time in the person's history

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.**

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

Appendix 1: Concept Set Definitions

1. Constipation

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 75860 | Constipation | Condition | SNOMED | NO | YES | NO |

### Decreased libido

Initial Event Cohort

People having any of the following: 

* a condition occurrence of Decreased libido1
  + for the first time in the person's history

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.**

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

Appendix 1: Concept Set Definitions

1. Decreased libido

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 436246 | Reduced libido | Condition | SNOMED | NO | YES | NO |
| 443262 | Lack or loss of sexual desire | Condition | SNOMED | NO | YES | NO |
| 4087317 | Lack of libido | Condition | SNOMED | NO | YES | NO |

### Delirium

Note: This algorithm relies on diagnosis codes associated with hospitalization. This approach may lead to underreporting, as described by McCoy et al. [[4](#_ENREF_4)]

Initial Event Cohort

People having any of the following: 

* a condition occurrence of Delirium1
  + for the first time in the person's history
  + 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.**

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

Appendix 1: Concept Set Definitions

1. Delirium

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 377830 | Alcohol withdrawal delirium | Condition | SNOMED | YES | YES | NO |
| 373995 | Delirium | Condition | SNOMED | NO | YES | NO |

### Diarrhea

Initial Event Cohort

People having any of the following:

* a condition occurrence of Diarrhea1

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

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

End Date Strategy

Date Offset Exit Criteria

This cohort definition end date will be the index event's start date plus 1 days

Cohort Collapse Strategy:

Collapse cohort by era with a gap size of 30 days.

Appendix 1: Concept Set Definitions

1. Diarrhea

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| 196523 | Diarrhea | Condition | SNOMED | NO | YES | NO |
| 4134607 | Diarrheal disorder | Condition | SNOMED | NO | YES | NO |
| 201773 | Enteritis of small intestine | Condition | SNOMED | NO | NO | NO |
| 80141 | Functional diarrhea | Condition | SNOMED | NO | YES | NO |
| 4207688 | Infectious enteritis | Condition | SNOMED | NO | NO | NO |
| 4324838 | Noninfectious enteritis | Condition | SNOMED | NO | NO | NO |
| 197596 | Toxic gastroenteritis | Condition | SNOMED | NO | YES | NO |
| 196620 | Viral enteritis | Condition | SNOMED | NO | YES | NO |

### Fracture

Note: This algorithm follows Lanteigne et al. [[5](#_ENREF_5)]

Initial Event Cohort

People having any of the following: 

* a condition occurrence of Fracture1
  + for the first time in the person's history

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.**

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

Appendix 1: Concept Set Definitions

1. Fracture

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 435093 | Closed fracture of femur | Condition | SNOMED | NO | YES | NO |
| 441974 | Closed fracture of forearm | Condition | SNOMED | NO | YES | NO |
| 4230399 | Closed fracture of hip | Condition | SNOMED | NO | YES | NO |
| 441422 | Closed fracture of humerus | Condition | SNOMED | NO | YES | NO |
| 439166 | Closed fracture of radius | Condition | SNOMED | NO | YES | NO |
| 4278672 | Fracture of forearm | Condition | SNOMED | NO | YES | NO |
| 442619 | Fracture of humerus | Condition | SNOMED | NO | YES | NO |
| 433856 | Fracture of neck of femur | Condition | SNOMED | NO | YES | NO |
| 4131595 | Fracture of radius | Condition | SNOMED | NO | YES | NO |
| 73571 | Pathological fracture | Condition | SNOMED | NO | YES | NO |

### Gastrointestinal hemorrhage

Initial Event Cohort

People having any of the following:

* a condition occurrence of Gastrointestinal hemorrhage GI bleeding2

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

For people matching the Primary Events, include:

Having any of the following criteria:

* at least 1 occurrences of a visit occurrence of Inpatient or ER visit1

where event starts between all days Before and 0 days After index start date and event ends between 0 days Before and all days After index start date

Limit cohort of initial events to: **all events per person.**

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

End Date Strategy

Date Offset Exit Criteria

This cohort defintion end date will be the index event's start date plus 7 days

Cohort Collapse Strategy:

Collapse cohort by era with a gap size of 30 days.

Appendix 1: Concept Set Definitions

1. Inpatient or ER visit

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| 262 | Emergency Room and Inpatient Visit | Visit | Visit | NO | YES | NO |
| 9203 | Emergency Room Visit | Visit | Visit | NO | YES | NO |
| 9201 | Inpatient Visit | Visit | Visit | NO | YES | NO |

2. Gastrointestinal hemorrhage GI bleeding

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| 4138962 | Acute duodenal ulcer without hemorrhage AND without perforation | Condition | SNOMED | YES | YES | NO |
| 4195231 | Acute gastric ulcer without hemorrhage AND without perforation | Condition | SNOMED | YES | YES | NO |
| 4147683 | Acute gastrojejunal ulcer without hemorrhage AND without perforation | Condition | SNOMED | NO | NO | NO |
| 4163865 | Acute peptic ulcer without hemorrhage AND without perforation | Condition | SNOMED | YES | YES | NO |
| 195584 | Acute peptic ulcer without hemorrhage AND without perforation but with obstruction | Condition | SNOMED | YES | YES | NO |
| 40482685 | Angiodysplasia of duodenum | Condition | SNOMED | NO | YES | NO |
| 28779 | Bleeding esophageal varices | Condition | SNOMED | NO | YES | NO |
| 4222896 | Chronic duodenal ulcer without hemorrhage AND without perforation | Condition | SNOMED | YES | YES | NO |
| 4296611 | Chronic gastric ulcer without hemorrhage AND without perforation | Condition | SNOMED | YES | YES | NO |
| 200769 | Chronic gastric ulcer without hemorrhage, without perforation AND without obstruction | Condition | SNOMED | YES | YES | NO |
| 4177387 | Chronic gastrojejunal ulcer without hemorrhage AND without perforation | Condition | SNOMED | YES | YES | NO |
| 434400 | Chronic gastrojejunal ulcer without hemorrhage AND without perforation but with obstruction | Condition | SNOMED | YES | YES | NO |
| 438795 | Chronic gastrojejunal ulcer without hemorrhage, without perforation AND without obstruction | Condition | SNOMED | YES | YES | NO |
| 4204555 | Chronic peptic ulcer without hemorrhage AND without perforation | Condition | SNOMED | YES | YES | NO |
| 24973 | Chronic peptic ulcer without hemorrhage AND without perforation but with obstruction | Condition | SNOMED | YES | YES | NO |
| 23808 | Chronic peptic ulcer without hemorrhage, without perforation AND without obstruction | Condition | SNOMED | YES | YES | NO |
| 2002608 | Control of hemorrhage and suture of ulcer of stomach or duodenum | Procedure | ICD9Proc | NO | YES | NO |
| 198798 | Dieulafoy's vascular malformation | Condition | SNOMED | NO | YES | NO |
| 4198381 | Duodenal ulcer disease | Condition | SNOMED | NO | YES | NO |
| 4209746 | Duodenal ulcer without hemorrhage AND without perforation | Condition | SNOMED | YES | YES | NO |
| 4112183 | Esophageal varices with bleeding, associated with another disorder | Condition | SNOMED | NO | YES | NO |
| 2108900 | Esophagogastroduodenoscopy, flexible, transoral; with control of bleeding, any method | Procedure | CPT4 | NO | YES | NO |
| 2108878 | Esophagoscopy, flexible, transoral; with control of bleeding, any method | Procedure | CPT4 | NO | YES | NO |
| 4265600 | Gastric ulcer | Condition | SNOMED | NO | YES | NO |
| 4248429 | Gastric ulcer without hemorrhage AND without perforation | Condition | SNOMED | YES | YES | NO |
| 192671 | Gastrointestinal hemorrhage | Condition | SNOMED | NO | YES | NO |
| 4101104 | Gastrojejunal ulcer without hemorrhage AND without perforation | Condition | SNOMED | YES | YES | NO |
| 443530 | Hematochezia | Condition | SNOMED | YES | YES | NO |
| 197925 | Hemorrhage of rectum and anus | Condition | SNOMED | YES | YES | NO |
| 4027663 | Peptic ulcer | Condition | SNOMED | NO | YES | NO |
| 4291028 | Peptic ulcer without hemorrhage AND without perforation | Condition | SNOMED | YES | YES | NO |

### Hyperprolactinemia

Initial Event Cohort

People having any of the following: 

* a condition occurrence of Hyperprolactinemia1
  + for the first time in the person's history

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.**

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

Appendix 1: Concept Set Definitions

1. Hyperprolactinemia

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 4030186 | Hyperprolactinemia | Condition | SNOMED | NO | YES | NO |

### Hyponatremia

Initial Event Cohort

People having any of the following:

* a condition occurrence of Hyponatremia1

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

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

End Date Strategy

Date Offset Exit Criteria

This cohort defintion end date will be the index event's start date plus 1 days

Cohort Collapse Strategy:

Collapse cohort by era with a gap size of 90 days.

Appendix 1: Concept Set Definitions

1. Hyponatremia

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| 435515 | Hypo-osmolality and or hyponatremia | Condition | SNOMED | NO | YES | NO |
| 4232311 | Hyponatremia | Condition | SNOMED | NO | YES | NO |

### Hypotension

Note: This algorithm follows Wernli et al. [[6](#_ENREF_6)]

Initial Event Cohort

People having any of the following: 

* a condition occurrence of Hypotension1
  + for the first time in the person's history

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.**

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

Appendix 1: Concept Set Definitions

1. Hypotension

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 4120275 | Drug-induced hypotension | Condition | SNOMED | NO | YES | NO |
| 317002 | Low blood pressure | Condition | SNOMED | NO | YES | NO |
| 314432 | Maternal hypotension syndrome | Condition | SNOMED | YES | YES | NO |
| 319041 | Orthostatic hypotension | Condition | SNOMED | NO | YES | NO |

### Hypothyroidism

Note: This algorithm requires the occurrences of 2 more diagnose codes, as recommended by Lu et al. [[7](#_ENREF_7)]

Initial Event Cohort

People having any of the following: 

* a condition occurrence of Hypothyroidism1

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

For people matching the Primary Events, include:

Having all of the following criteria:

* at least 2 occurrences of a condition occurrence of Hypothyroidism1

starting between 0 days Before and 90 days After event index date

Limit cohort of initial events to: **earliest event per person.**

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

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

Appendix 1: Concept Set Definitions

1. Hypothyroidism

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 140673 | Hypothyroidism | Condition | SNOMED | NO | YES | NO |

### Insomnia

Initial Event Cohort

People having any of the following:

* a condition occurrence of Insomnia1
  + for the first time in the person's history

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.**

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

Appendix 1: Concept Set Definitions

1. Insomnia

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 439708 | Disorders of initiating and maintaining sleep | Condition | SNOMED | NO | YES | NO |
| 436962 | Insomnia | Condition | SNOMED | NO | YES | NO |
| 4305303 | Sleep deprivation | Condition | SNOMED | NO | YES | NO |

### Nausea

Initial Event Cohort

People having any of the following:

* a condition occurrence of Nausea1

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

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

End Date Strategy

Date Offset Exit Criteria

This cohort defintion end date will be the index event's start date plus 1 days

Cohort Collapse Strategy:

Collapse cohort by era with a gap size of 30 days.

Appendix 1: Concept Set Definitions

1. Nausea

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| 30284 | Motion sickness | Condition | SNOMED | YES | YES | NO |
| 31967 | Nausea | Condition | SNOMED | NO | YES | NO |

### Open-**angle** glaucoma

Note: This algorithm follows Stein et al. [[8](#_ENREF_8)]

Initial Event Cohort

People having any of the following:

* a condition occurrence of Open-angle glaucoma1 for the first time in the person's history

with continuous observation of at least 365 days prior and 0 days after event index date, and limit initial events to: **earliest event per person.**

For people matching the Primary Events, include:

Having all of the following criteria:

* at least 1 occurrences of a condition occurrence of Open-angle glaucoma1
  + provider specialty is any of: Ophthalmology, Optometry, Optician

starting between 1 days After and 365 days After event index date

Limit cohort of initial events to: **earliest event per person.**

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

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

Appendix 1: Concept Set Definitions

1. Open-angle glaucoma

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 432908 | Glaucomatocyclitic crisis | Condition | SNOMED | YES | YES | NO |
| 441561 | Low tension glaucoma | Condition | SNOMED | NO | YES | NO |
| 4216823 | Open angle with borderline findings | Condition | SNOMED | YES | YES | NO |
| 441284 | Open-angle glaucoma | Condition | SNOMED | NO | YES | NO |
| 4072218 | Secondary open-angle glaucoma | Condition | SNOMED | YES | YES | NO |

### Unconfirmed open-**angle** glaucoma

Initial Event Cohort

People having any of the following:

* a condition occurrence of Open-angle glaucoma1
  + for the first time in the person's history

with continuous observation of at least 365 days prior and 0 days after event index date, and limit initial events to: **earliest event per person.**

For people matching the Primary Events, include:

Having all of the following criteria:

* at least 1 occurrences of a condition occurrence of Open-angle glaucoma1 starting between 1 days After and 365 days After event index date

Limit cohort of initial events to: **earliest event per person.**

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

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

Appendix 1: Concept Set Definitions

1. Open-angle glaucoma

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 432908 | Glaucomatocyclitic crisis | Condition | SNOMED | YES | YES | NO |
| 441561 | Low tension glaucoma | Condition | SNOMED | NO | YES | NO |
| 4216823 | Open angle with borderline findings | Condition | SNOMED | YES | YES | NO |
| 441284 | Open-angle glaucoma | Condition | SNOMED | NO | YES | NO |
| 4072218 | Secondary open-angle glaucoma | Condition | SNOMED | YES | YES | NO |

### Seizure

Note: This algorithm requires either inpatient or emergency room visits as recommended by Wu et al. [[9](#_ENREF_9)]

Initial Event Cohort

People having any of the following: 

* a condition occurrence of Seizure and seizure disorder1
  + for the first time in the person's history
  + 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.**

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

Appendix 1: Concept Set Definitions

1. Seizure and seizure disorder

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 380533 | Convulsions in the newborn | Condition | SNOMED | YES | YES | NO |
| 45757050 | Epilepsy in mother complicating pregnancy | Condition | SNOMED | YES | YES | NO |
| 377091 | Seizure | Condition | SNOMED | NO | YES | NO |
| 4029498 | Seizure disorder | Condition | SNOMED | NO | YES | NO |

### Stroke

Initial Event Cohort

People having any of the following:

* a condition occurrence of Stroke (ischemic or hemorrhagic)2

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

For people matching the Primary Events, include:

Having any of the following criteria:

* at least 1 occurrences of a visit occurrence of Inpatient or ER visit1

where event starts between all days Before and 1 days After index start date and event ends between 0 days Before and all days After index start date

Limit cohort of initial events to: **all events per person.**

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

End Date Strategy

Date Offset Exit Criteria

This cohort defintion end date will be the index event's start date plus 7 days

Cohort Collapse Strategy:

Collapse cohort by era with a gap size of 180 days.

Appendix 1: Concept Set Definitions

1. Inpatient or ER visit

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| 262 | Emergency Room and Inpatient Visit | Visit | Visit | NO | YES | NO |
| 9203 | Emergency Room Visit | Visit | Visit | NO | YES | NO |
| 9201 | Inpatient Visit | Visit | Visit | NO | YES | NO |

2. Stroke (ischemic or hemorrhagic)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| 372924 | Cerebral artery occlusion | Condition | SNOMED | NO | NO | NO |
| 375557 | Cerebral embolism | Condition | SNOMED | NO | NO | NO |
| 376713 | Cerebral hemorrhage | Condition | SNOMED | NO | NO | NO |
| 443454 | Cerebral infarction | Condition | SNOMED | NO | YES | NO |
| 441874 | Cerebral thrombosis | Condition | SNOMED | NO | NO | NO |
| 439847 | Intracranial hemorrhage | Condition | SNOMED | NO | NO | NO |
| 432923 | Subarachnoid hemorrhage | Condition | SNOMED | NO | NO | NO |

### Suicide and suicidal ideation

Initial Event Cohort

People having any of the following: 

* a condition occurrence of Suicide and suicidal ideation1
  + for the first time in the person's history
* an observation of Suicide and suicidal ideation1
  + for the first time in the person's history

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.**

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

Appendix 1: Concept Set Definitions

1. Suicide and suicidal ideation

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 439235 | Self inflicted injury | Condition | SNOMED | NO | YES | NO |
| 4181216 | Self-administered poisoning | Condition | SNOMED | NO | YES | NO |
| 444362 | Suicidal deliberate poisoning | Condition | SNOMED | NO | YES | NO |
| 4273391 | Suicidal thoughts | Condition | SNOMED | NO | YES | NO |
| 440925 | Suicide | Observation | SNOMED | NO | YES | NO |
| 4303690 | Intentionally harming self | Observation | SNOMED | NO | YES | NO |

### Tinnitus

Note: This algorithm follows Lee et al. [[10](#_ENREF_10)]

Initial Event Cohort

People having any of the following: 

* a condition occurrence of Tinnitus1
  + for the first time in the person's history

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.**

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

Appendix 1: Concept Set Definitions

1. Tinnitus

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 377575 | Tinnitus | Condition | SNOMED | NO | YES | NO |

### Ventricular arrhythmia and sudden cardiac death

Note: This algorithm follows the definition used by Leonard et al. [[11](#_ENREF_11)]

Initial Event Cohort

People having any of the following: 

* a condition occurrence of Ventricular arrhythmia and sudden cardiac death1
  + for the first time in the person's history
  + condition type is any of: Inpatient detail - primary, Inpatient header - primary, Primary Condition, Carrier claim detail - 1st position, Carrier claim header - 1st position, Inpatient detail - 1st position, Inpatient header - 1st position, Outpatient detail - 1st position, Outpatient 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: **earliest event per person.**

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

Appendix 1: Concept Set Definitions

1. Ventricular arrhythmia and sudden cardiac death

| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| --- | --- | --- | --- | --- | --- | --- |
| 321042 | Cardiac arrest | Condition | SNOMED | NO | YES | NO |
| 442289 | Death in less than 24 hours from onset of symptoms | Observation | SNOMED | NO | YES | NO |
| 441139 | Instantaneous death | Observation | SNOMED | NO | YES | NO |
| 4132309 | Sudden death | Observation | SNOMED | NO | YES | NO |
| 4185572 | Ventricular arrhythmia | Condition | SNOMED | NO | YES | NO |
| 437894 | Ventricular fibrillation | Condition | SNOMED | NO | YES | NO |
| 4103295 | Ventricular tachycardia | Condition | SNOMED | NO | YES | NO |

### Vertigo

Initial Event Cohort

People having any of the following:

* a condition occurrence of Vertigo1
  + for the first time in the person's history

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. Vertigo

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Concept Id** | **Concept Name** | **Domain** | **Vocabulary** | **Excluded** | **Descendants** | **Mapped** |
| 437496 | Epidemic vertigo | Condition | SNOMED | NO | YES | NO |
| 78162 | Peripheral vertigo | Condition | SNOMED | NO | YES | NO |
| 439383 | Vertigo | Condition | SNOMED | NO | YES | NO |
| 381035 | Vertigo of central origin | Condition | SNOMED | NO | YES | NO |

### Negative control outcomes

Negative controls are concepts known to not be associated with the target or comparator cohorts, such that we can assume the true relative risk between the two cohorts is 1. Negative controls are selected using a similar process to that outlined by Voss et al. [[12](#_ENREF_12)]. Person counts of all potential drug-condition pairs are reviewed in observational data; this person count data helps determine which pairs are even probable for use in calibration. Given the list of potential drug-condition pairs, the concepts in the pairs must meet the following requirements to be considered as negative controls: (1) that there is no Medline abstract where the MeSH terms suggest an association between the drug and the condition [[13](#_ENREF_13)], (2) that there is no mention of the drug-condition pair on a US Product Label in the “Adverse Drug Reactions” or “Postmarketing” section [[14](#_ENREF_14)], (3) there are no US spontaneous reports suggesting that the pair is in an adverse event relationship [[15](#_ENREF_15), [16](#_ENREF_16)], (4) that the OMOP Vocabulary does not suggest that the drug is indicated for the condition, (5) that the concepts are usable (i.e. not too broad, not suggestive of an adverse event relationship, not pregnancy related), and (6) the exact concept itself is utilized in patient level data (i.e. concepts that are not usually used within the data are usually indicative a broad concept that has a child that is more specific). The remaining concepts are “optimized”, meaning parent concepts remove children as defined by the OMOP Vocabulary (e.g. if both “Non-Hodgkin’s Lymphoma” and “B-Cell Lymphoma” we selected, child concept “B-Cell Lymphoma would be removed for its parent “Non-Hodgkin’s Lymphoma”). Once potential negative control candidates were selected, manual clinical review to exclude any pairs that may still be in a causal relationship or similar to the study outcome was be performed to select the top concepts by patient exposure. The final list can be found in Table 3.

|  |  |
| --- | --- |
| Acariasis | Ingrowing nail |
| Amyloidosis | Iridocyclitis |
| Ankylosing spondylitis | Irritable bowel syndrome |
| Aseptic necrosis of bone | Lesion of cervix |
| Astigmatism | Lyme disease |
| Bell's palsy | Malignant neoplasm of endocrine gland |
| Benign epithelial neoplasm of skin | Mononeuropathy |
| Chalazion | Onychomycosis |
| Chondromalacia | Osteochondropathy |
| Crohn's disease | Paraplegia |
| Croup | Polyp of intestine |
| Diabetic oculopathy | Presbyopia |
| Endocarditis | Pulmonary tuberculosis |
| Endometrial hyperplasia | Rectal mass |
| Enthesopathy | Sarcoidosis |
| Epicondylitis | Scar |
| Epstein-Barr virus disease | Seborrheic keratosis |
| Fracture of upper limb | Septic shock |
| Gallstone | Sjogren's syndrome |
| Genital herpes simplex | Tietze's disease |
| Hemangioma | Tonsillitis |
| Hodgkin's disease | Toxic goiter |
| Human papilloma virus infection | Ulcerative colitis |
| Hypoglycemic coma | Viral conjunctivitis |
| Hypopituitarism | Viral hepatitis |
| Impetigo | Visceroptosis |

**Table 3**. Negative control outcomes

For each negative control outcome, a patient enters the negative control outcome cohort at the occurrence of a diagnose code identified by the concepts listed above, or any one of its descendant codes.

### Positive control outcomes

In addition to negative control outcomes, we will also include synthetic positive control outcomes. These are outcomes based on the real negative controls, but where the true effect size is artificially increased to a desired effect size by injection of additional, simulated outcomes [[17](#_ENREF_17)]. To preserve confounding, these additional outcomes are sampled from predicted probabilities generated using a fitted predictive model. For each negative control outcome, three positive control outcomes will be generated with true relative risk is 1.5, 2, and 4. Using both negative and positive controls, we will fit a systematic error model and perform confidence interval calibration [[17](#_ENREF_17)].

## Covariates

### Propensity score covariates

Propensity scores (PS) will be used as an analytic strategy to reduce potential confounding due to imbalance between the target and comparator cohorts in baseline covariates. The propensity score is the probability of a patient being classified in the target cohort vs. the comparator cohort, given a set of observed covariates.

The types of baseline covariates used to fit the propensity score model will be:

* Demographics
  + Gender
  + Age group (5-year bands)
  + Index year
  + Index month
* Conditions
  + In prior 30d
  + In prior 365d
* Condition aggregation
  + SNOMED
* Drugs
  + In prior 30d
  + In prior 365d
  + Overlapping index date
* Drug aggregation
  + Ingredient
  + ATC Class
* Risk scores
  + Charlson comorbidity index
* Prior number of depression treatments (1,2,3,4, 5 or more)

All covariates that occur in fewer than 0.1% of the persons between the target and comparator cohorts combined will be excluded prior to model fitting for computational efficiency.

# Data Analysis Plan

## Calculation of time-at risk

Two time-at-risk periods will be used:

* On-treatment. Starting on the day of treatment initiation, and stopping at treatment end, allowing for a maximum gap of 30 days between prescriptions.
* Intent-to-treat: Starting on the day of treatment initiation and stopping at the end of observation.

## Model Specification

In this study, we compare the target cohort with the comparator cohort for the hazards of outcome during the time-at-risk by applying a Cox proportional hazards model.

The time-to-event of outcome among patients in the target and comparator cohorts is determined by calculating the number of days from the start of the time-at-risk window (the cohort start date), until the earliest event among 1) the first occurrence of the outcome, 2) the end of the time-at-risk window as defined in section 9.1 (i.e.. ‘on-treatment’ or ‘intent-to-treat’), and 3) the end of the observation period that spans the time-at-risk start.

Patients with the outcome observed prior to target or comparator cohort entry are excluded from consideration.

Propensity scores will be used as an analytic strategy to reduce potential confounding due to imbalance between the target and comparator cohorts in baseline covariates. The propensity score is the probability of a patient being classified in the target cohort vs. the comparator cohort, given a set of observed covariates. In this study, the propensity score is estimated for each patient, using the predicted probability from a regularized logistic regression model, fit with a Laplace prior (LASSO) and the regularization hyperparameter selected by optimizing the likelihood in a 10-fold cross, a starting variance of 0.01 and a tolerance of 2e-7. Covariates to be used in the propensity score model are listed in section 8.6.

In one analysis the target cohort and comparator cohorts will be stratified into ten quantiles of the propensity score distribution. A second analysis will use variable ratio matching based on the propensity score, using a caliper of 0.2 on the standardized logit scale. The final outcome model will apply a conditional Cox proportional hazard model, conditioned on the propensity score strata or matched sets.

Interactions between the treatment effect and the predefined subgroups will be evaluated in separate outcome models, one per subgroup. For efficiency reasons, only propensity score stratification will be used when investigating effect interactions

Incidence rates will be computed for each outcome in each exposure group, in both the on-treatment and intent-to-treat windows.

### Pooling effect estimates across databases

Effects will be pooled across databases using a random-effects meta-analysis. Estimates for negative and positive controls will be pooled before performing empirical calibration on the pooled estimates.

## Analyses to perform

### Comparative analyses

The following comparative analyses will be performed:

* 328 comparisons: 17 \* 16 = 272 comparisons between individual treatments, and 8 \* 7 = 56 comparisons between classes of treatments
* 23 outcomes of interest
* 2 time-at-risk definitions: on-treatment and intent-to-treat
* 2 models: Cox regression using propensity score stratification and Cox regression using propensity score matching
* 4 databases: CCAE, MDCD, MDCR, Optum

The total number of analyses for outcomes of interest is therefore 328 \* 23 \* 2 \* 2 \* 4 = 120,704 analyses.

We will also include 52 negative control outcomes, and 3 \* 52 positive control outcomes, so 208 control outcomes. The total number of control analyses is therefore 328 \* 208 \* 2 \* 2 \* 4 = 1,091,584 analyses.

Additionally, interaction effects will be computed with the 6 subgroups of interest. For efficiency reasons, this will only be computed when using propensity score stratification.

The number of analysis of subgroup interactions is therefore 328 \* 23 \* 2 \* 1 \* 4 \* 6 = 362,112 analyses.

For interaction effects, only negative control outcomes will be added. The total number of control analyses for interaction terms is therefore 328 \* 52 \* 2 \* 1 \* 4 \* 6 = 818,688 analyses.

### Descriptive analyses

The following incidence rate computations will be performed:

* 25 groups of interest: new users of the 17 individual treatments, and new users of the 8 treatment classes
* 23 outcomes of interest
* 2 time-at-risk definitions: on-treatment and intent-to-treat
* 4 databases: CCAE, MDCD, MDCR, Optum

The total number of analyses for outcomes of interest is therefore 25 \* 23 \* 2 \* 4 = 4,600 analyses

## Output

The output will be stored in the LEGEND evidence model, which is described elsewhere.

## Evidence Evaluation

We have executed diagnostics to determine if the analysis can be appropriately conducted. The diagnostics include:

* Propensity score distribution
* Covariate balance before and after propensity score matching
* Estimation for negative and positive controls, to assess residual error
* Negative and positive control exposures and outcomes will be used to evaluate the potential impact of residual systematic error in the study design, and to facilitate empirical calibration of the p-value and confidence interval for the exposures and outcome of interest.

Negative control outcomes in the context of this study are outcomes that are not believed to be caused by neither exposure in any comparison and where therefore the true hazard ratio is equal to 1. We will execute the same analysis used for the primary hypotheses to produce hazard ratio estimates for the negative controls. The distribution of effect estimates across all negative controls will be used to fit an empirical null distribution which models the observed residual systematic error. The empirical null distribution will then be applied to the target exposures and outcome of interest to calibrate the p-value [[18](#_ENREF_18)].

Positive control exposures and outcomes are pairs of exposures and outcomes where the hazard ratio is known to be of some magnitude greater than 1. We will synthesize positive controls by starting with the negative controls defined earlier, and adding additional, simulated outcomes during the time-at-risk until the desired true hazard ratio is achieved. The target hazard ratios are 1.5, 2 and 4. The negative and positive controls together will be used to estimate an empirical systematic error model, which will inform whether systematic error changes as a function of true effect size. The empirical systematic error model will then be applied to the target the target exposures and outcome of interest to calibrate the confidence interval [[17](#_ENREF_17)].

Empirical calibration serves as an important diagnostic tool to evaluate if the residual systematic error is sufficient to cast doubt on the accuracy of the unknown effect estimate. The calibration effect plot and calibration probability plots will be generated for review. We will report the traditional and empirically calibrated p-value and confidence interval for each negative control, as well as the hypothesis of interest.

# Study Diagnostics

## Sample Size and Study Power

This will be reported in the output (see the LEGEND data model).

## Cohort Comparability

This will be reported in the output (see the LEGEND data model).

## Systematic Error Assessment

This will be reported in the output (see the LEGEND data model).

# Strengths and Limitations of the Research Methods

Strength

* Cohort studies allow direct estimation of incidence rates following exposure of interest, and the new-user design can capture early events following treatment exposures while avoiding confounding from previous treatment effects. New use allows for a clear exposure index date.
* PS matching allow balancing on a large number of baseline potential confounders.
* Use of negative and positive control outcomes allow for evaluating the study design as a whole in terms of residual bias.

Limitations

* Even though many potential confounders will be included in this study, there may be residual bias due to unmeasured or misspecified confounders.

# 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.

# Management and Reporting of Adverse Events and Adverse Reactions

This study uses coded data that already exist in an electronic database. In this type of database, it is not possible to link (i.e., identify a potential causal association between) a particular product and medical event for any individual. Thus, the minimum criteria for reporting an adverse event (i.e., identifiable patient, identifiable reporter, a suspect product, and event) are not available and adverse events are not reportable as individual adverse events reports. The study results will be assessed for medically important results.

# Plans for Disseminating and Communicating Study Results

The study results will be posted on the OHDSI website after completion of the study. At least one paper describing the study and its results will be written and submitted for publication to a peer-reviewed scientific journal. The results will also be presented at the OHDSI 2018 Symposium.

# References

1. Schuemie, M.J., et al., *Improving reproducibility by using high-throughput observational studies with empirical calibration.* Philos Trans A Math Phys Eng Sci, 2018. **376**(2128).

2. Udo, R., et al., *Validity of diagnostic codes and laboratory measurements to identify patients with idiopathic acute liver injury in a hospital database.* Pharmacoepidemiol Drug Saf, 2016. **25 Suppl 1**: p. 21-8.

3. Mody, R., et al., *Prevalence and risk of developing comorbid conditions in patients with chronic constipation.* Curr Med Res Opin, 2014. **30**(12): p. 2505-13.

4. McCoy, T.H., Jr., et al., *Underreporting of Delirium in Statewide Claims Data: Implications for Clinical Care and Predictive Modeling.* Psychosomatics, 2016. **57**(5): p. 480-8.

5. Lanteigne, A., et al., *Serotonin-norepinephrine reuptake inhibitor and selective serotonin reuptake inhibitor use and risk of fractures: a new-user cohort study among US adults aged 50 years and older.* CNS Drugs, 2015. **29**(3): p. 245-52.

6. Wernli, K.J., et al., *Risks Associated With Anesthesia Services During Colonoscopy.* Gastroenterology, 2016. **150**(4): p. 888-94; quiz e18.

7. Lu, M.C., et al., *Higher Risk of Thyroid Disorders in Young Patients with Type 1 Diabetes: A 12-Year Nationwide, Population-Based, Retrospective Cohort Study.* PLoS One, 2016. **11**(3): p. e0152168.

8. Stein, J.D., T.S. Blachley, and D.C. Musch, *Identification of persons with incident ocular diseases using health care claims databases.* Am J Ophthalmol, 2013. **156**(6): p. 1169-1175 e3.

9. Wu, C.S., et al., *Comparative risk of seizure with use of first- and second-generation antipsychotics in patients with schizophrenia and mood disorders.* J Clin Psychiatry, 2016. **77**(5): p. e573-9.

10. Lee, C.F., et al., *Increased risk of tinnitus in patients with temporomandibular disorder: a retrospective population-based cohort study.* Eur Arch Otorhinolaryngol, 2016. **273**(1): p. 203-8.

11. Leonard, C.E., et al., *Antidepressants and the risk of sudden cardiac death and ventricular arrhythmia.* Pharmacoepidemiol Drug Saf, 2011. **20**(9): p. 903-13.

12. Voss, E.A., et al., *Accuracy of an automated knowledge base for identifying drug adverse reactions.* J Biomed Inform, 2017. **66**: p. 72-81.

13. Winnenburg, R., et al., *Leveraging MEDLINE indexing for pharmacovigilance - Inherent limitations and mitigation strategies.* J Biomed Inform, 2015. **57**: p. 425-35.

14. Duke, J., J. Friedlin, and X. Li, *Consistency in the safety labeling of bioequivalent medications.* Pharmacoepidemiol Drug Saf, 2013. **22**(3): p. 294-301.

15. Evans, S.J., P.C. Waller, and S. Davis, *Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports.* Pharmacoepidemiol Drug Saf, 2001. **10**(6): p. 483-6.

16. Banda, J.M., et al., *A curated and standardized adverse drug event resource to accelerate drug safety research.* Sci Data, 2016. **3**: p. 160026.

17. Schuemie, M.J., et al., *Empirical confidence interval calibration for population-level effect estimation studies in observational healthcare data.* Proc Natl Acad Sci U S A, 2018. **115**(11): p. 2571-2577.

18. Schuemie, M.J., et al., *Interpreting observational studies: why empirical calibration is needed to correct p-values.* Stat Med, 2014. **33**(2): p. 209-18.

# Appendix A. Differences with previous depression study

Below is a summary of the differences between this study and the previous depression study [[1](#_ENREF_1)].

**Indication**

Instead of requiring a diagnosis of Major Depressive Disorder (including descendants) prior to treatment, we now require a diagnosis of Depressive Disorder (including descendants) prior to treatment. The rationale is that this will allow us to include other databases whose source codes do not translate to Major Depressive Disorder, but do translate to Depressive Disorder.

**Exposures**

Next to the individual ingredients, we have now also include drug classes as exposures. The reason being that some effects are class effects, and we would like to observe those directly, rather than to infer them from the ingredient effects in the classes.

**Outcomes**

* Open-angle glaucoma: added a second definition that does not require second diagnose to be by ophthalmologist (called 'Unconfirmed open-angle glaucoma'). The reason is to include databases that do not store physician specialty, or have no specialist records (e.g. primary care databases).
* Stroke: use updated definition from upcoming hypertension study, which also includes hemorrhagic type, and doesn't require explicit link to visit but instead infers from overlapping dates. The reasons are to be consistent between depression and hypertension studies, and to include databases that do not explicitly link diagnoses to visits.
* Suicide: added concept 'Intentionally harming self' to definition. This new code fits the original definition and is used in some databases not included in the original study.
* Decreased libido: use updated definition from upcoming hypertension study, which additionally includes 'Lack of libido' and 'Lack or loss of sexual desire'. The reasons are to be consistent between depression and hypertension studies, and to include databases not part of the original study.
* Acute Myocardial Infarction: use updated definition from upcoming hypertension study, which uses dates to link visits rather than rely on explicit links (which may not exist in all data sources). Also, the diagnose is no longer required to be a primary diagnosis.
* Diarrhea: use updated definition from upcoming hypertension study, which includes several additional concepts such as enteritis.
* Gastrointestinal hemorrhage: use updated definition from upcoming hypertension study, which uses dates to link visits rather than rely on explicit links (which may not exist in all data sources). Also, the diagnose is no longer required to be a primary diagnosis.
* Hyponatremia: use updated definition from upcoming hypertension study, which includes additional concepts such as hyponatremia, and drops the definition based on measurement values.
* Nausea: use updated definition from upcoming hypertension study, which is identical to the original definition except for allowing multiple occurrences of nausea (not currently used since the analysis limits to first occurrence).
* Vertigo: use updated definition from upcoming hypertension study, which includes additional concepts such Epidemic Vertigo.

**Subgroups**

In contrast to the previous study, this study considers interactions of an observed effect with predefined subgroups:

* Renal impairment
* Hepatic impairment
* Pregnant women
* Children (age < 18)
* Elderly (age >=65)
* Gender = female

**Methods**

Whereas the previous study only included propensity score stratification, this study will also use variable ratio propensity score matching.

**Databases**

The previous study was executed against four US insurance claims databases (CCAE, MDCD, MDCR, Optum). All four are also included in this study, with updated data, and several others have been added.