

OMOP Tools and Their Application

Name	Application
Common Data Model (CDM)	http://omop.fnih.org/ETLProcess The CDM intends to facilitate observational analyses of disparate healthcare databases. The CDM defines table structures for each of the data entities (e.g., Persons, Visit Occurrence, Drug Exposure, Condition Occurrence, Observation, Procedure-Occurrence, etc.). It includes all observational data elements that are relevant to identifying drug exposures and defining condition occurrence. The CDM includes both the dictionary of terms and the entity domain tables.
Standardized Terminologies	http://omop.fnih.org/vocabularies The Standard Terminology facilitates the OMOP methodological research by making it possible that methods can be systematically applied to produce meaningfully comparable results across different data sources with different coding schemes. It enables a mechanism of transforming raw data into standardized data. It also plays a role in searching and querying the transformed data in the CDM, as well as browsing and navigating the hierarchies of classes and abstractions inherent in the transformed data, and interpreting the results returned by those operations.
Health Outcomes of Interest (HOI)	http://omop.fnih.org/HOI OMOP established an open-source library of 10 HOI definitions for use in observational studies. A summary is available of the operational definitions of the HOIs under OMOP study. Where there was lack of consensus in the appropriate definition for an outcome, multiple alternative definitions were created; therefore an HOI may have more than one definition. The alternative definitions cover inclusive (broad) and specific (narrow) definitions, and make use of combinations of diagnoses with diagnostic procedures, therapeutic procedures, lab tests and lab results.
Observational Medical Dataset Simulator (OSIM)	http://omop.fnih.org/OSIM Open-source software application that allows users to create simulated data sets that conform to the OMOP CDM. The simulation creates hypothetical persons with fictitious drug exposure and conditions, with known characteristics that represent the types of scenarios expected in real observational sources.
Observational Source Characteristics Analysis Report (OSCAR)	http://omop.fnih.org/OSCAR An SAS program that creates descriptive statistics, allowing the following functions: summarizing available data from a given source within the OMOP CDM; providing context for interpreting and analyzing findings of drug safety studies; facilitating comparisons between data sources; enabling comparison of overall database to specific subpopulations of interest; and supporting validation of transformation from raw data to OMOP CDM.
Natural History Analysis (NATHAN)	http://omop.fnih.org/NATHAN An SAS program extension of OSCAR, creating a standardized summary providing some context and expected rates of drug utilization and condition occurrence to facilitate the interpretation of benefit and risk information and generate a standardized report summarizing characteristics about the population of interest, including demographic factors (age and sex); comorbid conditions and concomitant medications; and health service utilization before, during, and after the event onset.

Name	Application
Regularized Identification of Cohorts (RICO)	http://omop.fnih.org/MethodsLibrary A procedure that standardizes patient cohort selection. Standard “cohort definitions” are created by using criteria specified in input parameters. These cohort definitions are input into RICO, and patients meeting the criteria can be automatically and rapidly selected from any database conforming to the OMOP CDM.
Generalized Review of OSCAR Unified Checking (GROUCH)	http://omop.fnih.org/GROUCH GROUCH is a program that produces a summary report for each data source of warnings of implausible and suspicious data observed from the OSCAR summary. It identifies potential issues across all OMOP CDM tables, including potential concerns with all drug exposures and all conditions. GROUCH allows for data quality review of specific drugs (such as the ingredients that comprise the OMOP drugs of interest) or specific conditions (including population-level prevalence of the health outcomes of interest, and unexpected gender-specific rates, such as males with pregnancy, and females with prostate cancer).
Analysis Methods	http://omop.fnih.org/MethodsLibrary A library of methods and specifications, developed for the OMOP CDM, to address the analysis problems of Monitoring of Health Outcomes of Interest and Identification of Non-Specified Conditions. <ul style="list-style-type: none"> • Disproportionality Analysis (DP)- OMOP Research Team • Univariate Self-Controlled Case Series (USCCS) - OMOP Research Team • Observational Screening (OS)- ProSanos Corporation • Multi-Set Case Control Estimation (MSCCE)- OMOP Research Team • Bayesian Logistic Regression (BLR)- OMOP Research Team • Case Control Surveillance (CCS)- Lilly • IC Temporal Pattern Discovery (ICTPD)- Uppsala Monitoring Centre • Case-Crossover (CCO)- University of Utah • HSIU Population-Based Method - Indiana University • Maximized Sequential Probability Ratio Test (MSPRT) - Harvard Pilgrim • Conditional Sequential Sampling Procedure (CSSP) - Harvard Pilgrim • High-Dimensional Propensity Score (HDPS) - OMOP Research Team • Incident User Design (IUD-HOI) - M. Alan Brookhart