SENTINEL MODULAR PROGRAMS

Querying Tools: Overview of Functionality and Technical Documentation

Prepared by the Sentinel Operations Center
November 1, 2018
Version: 7.0.0

Sentinel is sponsored by the U.S. Food and Drug Administration (FDA) to monitor the safety of FDA-regulated medical products. Sentinel is one piece of the Sentinel Initiative, a multi-faceted effort by the FDA to develop a national electronic system that complements previously existing methods of safety surveillance. Sentinel Collaborators include Data and Academic Partners that provide access to health care data and ongoing scientific, technical, methodological, and organizational expertise. The Sentinel Coordinating Center is funded by the FDA through the Department of Health and Human Services (HHS) Contract number HHSF223201400030I. This project was funded by the FDA through HHS Mini-Sentinel contract number HHSF223200910006I.

Table of Contents

Modification History

Version	Date	Modification	By
7.0.0	11/1/2018	• Added the ability for Type 4 pregnancy analyses to evaluate maternal and/or infant outcomes following maternal exposure when used with the PSA tool.	Sentinel Operations Center
6.0.0	10/1/2018	• Multiple enhancements to Type 2 analyses, including 1) ability to evaluate secondary events in the pre-defined observation window in relation to the primary episode; 2) ability to examine concomitant use of two different treatment episodes. • Added the ability to evaluate product utilization and switching patterns via creation of a new cohort identification strategy (Type 6) • Ability to stratify by covariates and option to create user-defined strata	Sentinel Operations Center
5.4.4	08/07/2018	• Upversioned due to bug fix	Sentinel Operations Center
5.4.3	06/14/2018	• Upversioned due to bug fix	Sentinel Operations Center
5.4.2	06/13/2018	• Upversioned due to bug fix	Sentinel Operations Center
5.4.1	05/22/2018	• Upversioned due to bug fix	Sentinel Operations Center
5.4.0	05/16/2018	• Added the ability to optionally output the distribution of codes associated with index events for all analysis types • Multiple enhancements to Type 4 pregnancy analyses, including 1) allowing users to include first or last only medical product use episodes during pregnancy; 2) “blackout” medical product exposures that occur within specified time period during pregnancy; 3) raw and adjusted code counts; 4) ability to define pre-pregnancy period instead of 90 day default. • Replaced outputs for Dispensings with Adjusted code counts and Raw code counts for Types 1, 2, 5	Sentinel Operations Center
5.3.1	04/18/2018	• Altered Type 5 output to provide episode length by individual episode (EpisodeLength), as well as cumulatively across all episodes (CumEpisodeLength) • Restricted the PS distribution to the current look	Sentinel Operations Center

Version	Date	Modification	By
5.3.0	03/30/2018	• Added ability to create a never-exposed cohort in a Type 2 request • Added ability to use never-exposed cohort as the comparator in a propensity score analysis • Added ability to perform multiple factor matching	Sentinel Operations Center
5.2.1	02/02/2018	• Removed default requirement for post-index enrollment criteria when cohort exclusion, covariate, most frequent untilization, and high-dimensional propensity score evaluation windows extend beyond the index date. Post-index enrollment requirements must now be specified by the requester.	Sentinel Operations Center
5.2.0	01/25/2018	• Added ability to create report for Type 1 and Type 2 requests • Redesigned parameter definitions for comorbidity score • Added ability to count multiple occurrences of codes in inclusion/exclusion and covariate criteria	Sentinel Operations Center
5.1.2	12/12/2017	• Upversioned due to bug fix	Sentinel Operations Center
5.1.1	12/07/2017	• Restricted covariate codes to valid enrollment spans	Sentinel Operations Center
5.1.0	11/20/2017	• Added the ability to output strata by regions • Added two new parameters to the Geography Lookup Table (formerly known as “Zip Code Lookup Table”) • Updated QRP with new ICD-10 inclusive pregnancy algorithm	Sentinel Operations Center
5.0.5	10/17/2017	• Upversioned due to bug fix	Sentinel Operations Center
5.0.4	10/16/2017	• Upversioned due to bug fix	Sentinel Operations Center
5.0.3	09/25/2017	• Upversioned due to bug fix	Sentinel Operations Center
5.0.2	09/18/2017	• Upversioned due to bug fix	Sentinel Operations Center
5.0.1	09/14/2017	• Upversioned due to bug fix	Sentinel Operations Center
5.0.0	08/21/2017	• Integrated MP7 and MP8 into QRP • Made Type 2 and Type 4 consistent with Type 5	Sentinel Operations Center
4.1.3	07/13/2017	• Upversioned due to bug fix	Sentinel Operations Center

Version	Date	Modification	By
4.1.2	07/12/2017	• Upversioned due to bug fix	Sentinel Operations Center
4.1.1	07/05/2017	• Upversioned due to bug fix	Sentinel Operations Center
4.1.0	06/26/2017	• Added the ability to allow subgroup analyses by RACE or HISPANIC • Added YEAR values to Table 1 output • AGESTRAT values can now be specified by scenario (vary across GROUP values) • Enabled the creation of PS distributions locally	Sentinel Operations Center
4.0.0	05/23/2017	• Made SEX, RACE, HISPANIC, ENRDAYSAFT available to all Types • Moved ENRDAYSFLOOR parameter from COHORTFILE to the TYPE4FILE • Moved GEOG parameter from COHORTFILE to Type 1 and Type 2 files • Renamed SubGroup input parameter to "StockGroup" and RawSubGroup to "RawStockGroup" in the combo tool • Removed concept of "categories" for Type 4 analyses • Redesigned Type 4 output • Removed PXDXSupplyFunc • Covariates are now able to be defined within the Combo Tool • CARESETTINGPRINCIPAL allowable values updated • New section titled "Creation and Retention of First Valid Episodes" added	Sentinel Operations Center
3.3.6	03/02/2017	• Updated riskdiff output to include only informative events; updated propensity score distribution figures	Sentinel Operations Center
3.3.5	02/13/2017	• Added ability to output unmatched and matched propensity score distributions in .pdf format	Sentinel Operations Center
3.3.4	02/09/2017	• Upversioned due to bug fix	Sentinel Operations Center
3.3.3	01/12/2017	• Upversioned due to minor bug fix for attrition table	Sentinel Operations Center
3.3.2	12/12/2016	• Upversioned to allow duplicate NDC codes across covariates	Sentinel Operations Center
3.3.1	12/09/2016	• Removed KeepAllDum as a parameter in the Comorbfile	Sentinel Operations Center

Version	Date	Modification	By
3.3.0	12/06/2016	• Added ability to truncate follow-up at query end date • Added ability to output reason for follow-up time censoring	Sentinel Operations Center
3.2.0	11/22/2016	• Added ability to stratify output by geographic location • Added ability to turn off "envelope" macro	Sentinel Operations Center
3.1.0	10/19/2016	• Added new prospective surveillance section and options for using the propensity score matched design	Sentinel Operations Center
3.0.3	08/15/2016	• Updated HDPS to include two new ICD-10 data dimensions	Sentinel Operations Center
3.0.2	07/20/2016	• Added ability to output additional information for patients identified in a self-controlled risk interval design (SCRI) cohort: event information for patients not meeting post-exposure enrollment requirements.	Sentinel Operations Center
3.0.1	07/14/2016	• Added ability to specify a maximum episode duration for a "Type 2" analysis	Sentinel Operations Center
3.0.0	07/14/2016	• Added ability to define complex inclusion/exclusion criteria algorithms outside of the Combo Tool • Added ability to define varying covariate assessment windows • Added ability to explicitly define covariates in propensity score estimation models (i.e., no forced variable inclusion) • Added ability to return additional propensity score estimation diagnostic output	Sentinel Operations Center
2.2.3	06/16/2016	• Up-versioned due to bug fix	Sentinel Operations Center
2.2.2	06/13/2016	• Up-versioned due to bug fix	Sentinel Operations Center
2.2.1	04/20/2016	• Up-versioned due to bug fix	Sentinel Operations Center
2.2.0	04/15/2016	• Modified to include drug use during pregnancy "Type 4" analysis	Sentinel Operations Center
2.1.1	02/19/2016	• Added ability to specify covariates and generate baseline "Table 1" output for evaluation prior to specifying comparative, adjusted analyses • Updated lab lookup table for SCDM 5.0.1 compliance	Sentinel Operations Center
2.1.0	12/14/2015	• Added risk set level analysis for propensity score matched design	Sentinel Operations Center

Version	Date	Modification	By
2.0.8	07/21/2015	• Up-versioned due to minor bug fix	Sentinel Operations Center
2.0.7	07/10/2015	• Up-versioned due to minor bug fix	Sentinel Operations Center
2.0.6	07/07/2015	• Up-versioned due to minor bug fix	Sentinel Operations Center
2.0.5	04/13/2015	• Up-versioned due to minor bug fix	Sentinel Operations Center
2.0.4	03/31/2015	• Up-versioned due to minor bug fix	Sentinel Operations Center
2.0.3	03/19/2015	• Changed output variable names in *[RUNID]* _mstr Output for Type 3 analyses • Made minor bug fixes	Sentinel Operations Center
2.0.2	02/20/2015	• Moved output of two datasets to SOC folder from DPLOCAL folder: *[RUNID]*_varinfo_*[COMP_ORDER]*_*[PERIODID]* and *[RUNID]*_estimates_*[COMP_ORDER]*_ *[*PERIODID*]* • Made minor bug fixes	Sentinel Operations Center
2.0.1	01/30/2015	• Up-versioned due to minor bug fix	Sentinel Operations Center
2.0	01/16/2015	• Modified to include self-controlled risk interval design cohort identification strategy • Modified to change variable matching ratio ceiling from 100 to 10	Sentinel Operations Center
1.4.1	12/18/2014	• Up-versioned due to minor change to macro headers and comments	Sentinel Operations Center
1.4	10/14/2014	• Up-versioned due to minor bug fix	Sentinel Operations Center
1.3	10/02/2014	• Modified to make 1:100 variable ratio matching optional	Sentinel Operations Center
1.2	09/16/2014	• Up-versioned due to minor code modifications	Sentinel Operations Center
1.1	09/05/2014	• Modified to allow combo tool module updated to process combinations of combination items	Sentinel Operations Center
1.0	07/31/2014	• Original published version	Sentinel Operations Center

I. OVERVIEW

Sentinel modular programs (MPs) allow rapid implementation of standard queries across the Sentinel Distributed Database (SDD). MPs are designed to run against the Sentinel Common Data Model (SCDM).[1] They are written in SAS and can be customized using various parameter settings that define exposures, outcomes, covariates, inclusion/exclusion criteria, date ranges, age ranges, and other implementation details.

II. AVAILABLE MODULAR PROGRAMS

Several modular programs are available to execute Sentinel data queries. The **Cohort Identification and Descriptive Analysis (CIDA) tool** identifies and extracts cohorts of interest from the SDD based on the specification of a number of requester-defined options (e.g., exposures, outcomes, continuous enrollment requirements, incidence criteria, inclusion/exclusion criteria, relevant age groups, demographic criteria such as sex or race). The CIDA tool calculates descriptive statistics for the cohort(s) of interest and outputs datasets that may be useful for additional analyses (e.g., to calculate unadjusted and adjusted effect estimates and 95% confidence intervals).

The CIDA tool generates output containing information on exposures, outcomes, and covariates that can be used as input to the **Propensity Score Analysis (PSA) tool**. The PSA tool uses the information output by the CIDA tool to estimate a propensity score based on predefined covariates and/or via a high-dimensional propensity score approach. The PSA tool then matches individuals in an exposed cohort to individuals in a comparator cohort based on propensity score and calculates hazard ratios, incidence rate differences and 95% confidence intervals.

The CIDA tool generates output containing information on exposures and outcomes that can be used as input to the **Multiple Factor Analysis (MFM) tool**. The MFM tool uses the information output by the CIDA tool to find an exact match between an exposure and comparator cohort based on predefinied options. The MFM tool calculates hazard ratios, incidence rate differences and 95% confidence intervals.

III. COHORT IDENTIFICATION AND DESCRIPTIVE ANALYSIS TOOL

The main purpose of the CIDA tool is to identify and extract cohorts of interest from the SDD based on the specification of requester-defined options (e.g., exposures, outcomes, continuous enrollment requirements, incidence criteria, inclusion/exclusion criteria, relevant age groups and demographic criteria such as sex and race). The CIDA tool may be used alone or in conjunction with additional tools that perform more complex adjustment for confounders.

[1] See https://www.sentinelinitiative.org/sentinel/data/distributed-database-common-data-model for more information about the SCDM.

A. COHORT IDENTIFICATION STRATEGIES

There are six cohort identification strategies currently available with the CIDA tool:

1. Extract information to calculate background rates: The program identifies an exposure, outcome, or medical condition, and calculates the rate of that event in the SDD. Output includes the number of individuals with the exposure/outcome/medical condition, eligible members, and eligible member-days. Rates are reported overall and stratified by requester-defined age group, sex, year, and year-month. An attrition table is also provided upon request.

2. Extract information on exposures and follow-up time: The program identifies an exposure of interest, determines exposed time as either requester-defined number of days after exposure initiation or based on drug dispensing days supply, and looks for the occurrence of HOIs during exposed time. Output metrics include number of exposure episodes and number of individuals, number of HOIs, and days at-risk. Events per person-day at risk are reported overall and stratified by requester-defined age group, sex, year, and year-month. Incidence rate ratios (IRRs) can be calculated using two identified cohorts (e.g., exposed vs. active-comparator cohort). Unadjusted IRRs and IRRs adjusted by age group, sex, year, and Data Partner are reported upon request. An attrition table is also provided upon request.

The user also has the option to characterize cohorts with concomitant use of medical products and look for the occurrence of health outcomes of interest (HOI) using the Concomitant Use Tool. Users may also characterize multiple events within an episode of medical product use through use of the Multiple Events Tool and characterize overlap between two separate treatment episodes using the Overlap Tool.

The exposures and follow-up time cohort identification strategy is designed to be compatible with the PSA tool. For this cohort identification strategy the CIDA tool will 1) extract covariates of interest during requester-defined evaluation windows for the propensity score estimation model; and 2) output an analytic dataset containing the necessary information for the PSA tool to execute.

3. Extract information for a self-controlled risk interval design: The program identifies an exposure of interest, identifies a requester-defined risk and control window relative to the exposure date, and examines the occurrence of HOIs during the risk and control windows. Output metrics include number of exposure episodes, exposed individuals, individuals with an HOI in the risk and/or control windows, and censored individuals, overall and stratified by requester-defined age group, sex, year, year-month, and time-to-event in days. An attrition table is provided upon request.

4. Extract information for medical product use during pregnancy: The program identifies live births, computes pregnancy episodes based on those live birth events, and assesses the use of specific medical products both during pregnancy episodes and in a comparator group of women likely to not have delivered a live birth during the same time frame. Output includes the number of pregnancy episodes stratified by year, maternal age, and existence of a pre-term or postterm pregnancy code. Medical product use is reported for both pregnancy episodes and comparator episodes according to trimester of use, gestational week, maternal age, and calendar year of delivery.

The medical product use during pregnancy cohort identification strategy is designed to be compatible with the PSA tool. When used with the PSA tool, maternal and infant health outcomes of interest are evaluated. An exposure pregnant cohort and a comparator or unexposed pregnant cohort can be assessed. For this cohort identification strategy, the CIDA tool will 1) extract covariates of interest during requester-defined evaluation windows for the propensity score estimation model; and 2) output an analytic dataset containing the necessary information for the PSA tool to execute.

5. Extract information for medical product utilization: The program identifies the "first valid" exposure episode (i.e., the first episode during the query period that meets cohort entry criteria) as the index date, and then includes all subsequent exposure episodes. Output metrics include the number of patients, episodes, dispensings, and days supply by sex, age group and month of study start (for the first patient episode or all observed episodes during the query period); number of episodes by episode number, episode length, sex and age group, reason(s) for censoring; number of episode gaps by gap number, gap length, sex and age group.

6. Extract information on manufacturer-level product utilization and switching patterns: The program identifies product groups by user-defined lists of product codes (e.g., NDCs) grouped together to represent distinct manufacturer-level products and then characterizes patterns of drug use. Output metrics include counts of users and dispensings, days supplied per dispensing, episode duration, as well as time to uptake. The CIDA tool also performs a product switching analysis that evaluates patient-level switching behavior between manufacturer-level product groups.

While many options available in the CIDA tool are specific to the cohort identification strategy employed, some are not. The next sections describe functionality common across strategies, and are followed by descriptions of functionality specific to the cohort identification strategy selected.

B. UNIVERSAL OPTIONS

1. Defining the Query Period, Age and Demographic Criteria

All CIDA tool requests must specify a query start date and query end date. Available data before the query start date may be used to determine if specified enrollment and incidence criteria are met and evaluate inclusion/exclusion criteria and presence/absence of covariates. For the exposures and follow-up cohort identification strategy, exposure episodes may also extend beyond the query end date, as long as the episode was initiated during the query period and all enrollment requirements are still met (requesters may optionally turn this functionality off, to censor episodes at query end date). CIDA tool requests can also restrict cohorts to a requester-defined age range or demographic criteria such as sex and race, and stratify output by requester-defined age groups. Index date (e.g., exposure initiation date) is used to calculate patient age.

For the pregnancy episodes cohort identification strategy (Section F), query start date and query end date binds the delivery date only. Available data before the query start date may be used to determine pregnancy start date, if specified enrollment and incidence criteria are met and evaluate inclusion/exclusion criteria. Delivery date is used to calculate maternal age in the pregnancy episodes cohort identification strategy.

2. Enrollment Requirements

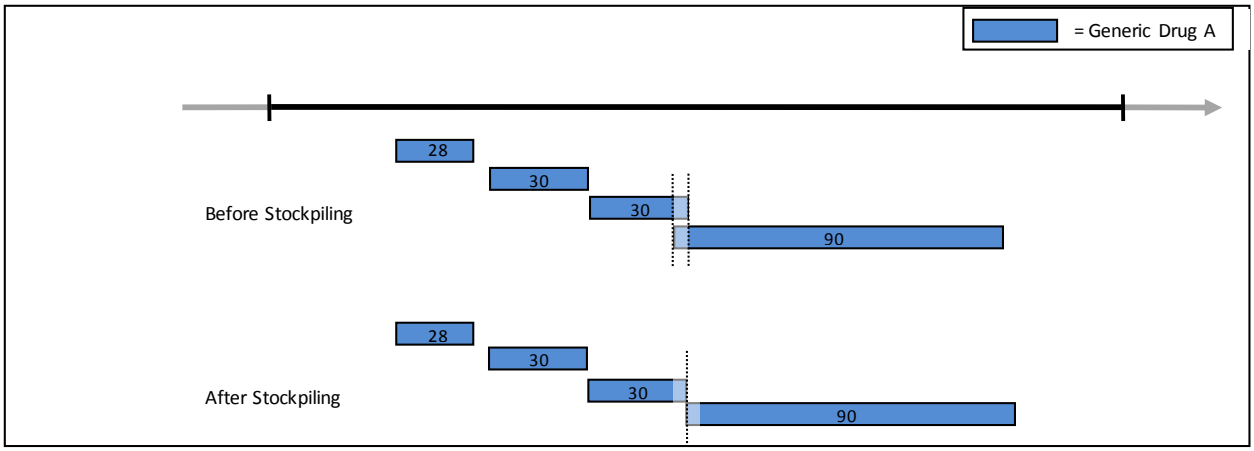
All data used by the CIDA tool to select cohorts of interest must be observed during valid enrollment periods of a specific coverage type. Requesters must select the type of coverage required based on whether medical, drug, or both medical and drug coverage are required during enrollment periods. This may be driven by the query of interest (e.g., if an exposure is defined using outpatient pharmacy dispensings and the HOI is defined using ICD diagnosis codes, requesters will want to ensure that periods with both medical and drug coverage are considered). Requesters can also require to exclude members for whom medical charts cannot be requested for the entire study period.

Once coverage type and chart availability restriction are established, continuous enrollment periods are constructed by bridging consecutive enrollment periods of the specified coverage type. Requesters may specify a maximum enrollment gap that will instruct the CIDA tool to "bridge" consecutive enrollment periods that are separated by (no more than) the specified number of days. These continuous enrollment periods define the time period that patients are eligible to contribute information used to create the cohort.

3. National Drug Code Processing and the Stockpiling Algorithm

National Drug Codes (NDCs) used to define exposures, HOIs, inclusion/exclusion criteria and covariates are pre-processed by the CIDA tool. In the rare case that an NDC has days supplied or amount supplied value equal to or lower than zero, the program will, by default, disregard the dispensing. Additionally, because members may refill their drug prescriptions before the end of days supply of the prior prescription, a stockpiling algorithm is used to account for dispensings with overlapping days of supply. Since this early-refill pattern may artificially reduce the length of exposed time (and therefore days at-risk), the dispensing date of the subsequent overlapping dispensing is adjusted. Drug dispensings are typically stockpiled by generic name (e.g., if the exposure of interest is all statins, atorvastatin, fluvastatin, lovastatin, etc. will be stockpiled independently). Figure 1 illustrates the stockpiling algorithm and how dispensing dates are adjusted by generic name.

Figure 1. Stockpiling Algorithm to Adjust Dispensing Dates



Note that while stockpiling is performed by default, augmenting how and when drug dispensings are stockpiled is possible upon request.

4. Defining Complex algorithms

The CIDA tool is integrated with the Combo tool, a re-usable SAS macro that allows requesters to define events (i.e., exposures, HOIs, and inclusion/exclusion criteria) using complex algorithms. Utilizing the Combo tool, the CIDA tool can do the following to define a single event: combine any NDCs, diagnoses, procedures, encounter types, enrollment episodes and lab values (using "and" and "or" joins); use same-day, same-encounter, or time intervals (*e.g.,* diagnoses X and procedure Y within 2 weeks of each other); and define a specific exposure length for any codes that comprise an event definition (similar to days of supply or length of stay). Detailed documentation of Combo tool functionality can be found in a separate document: 'Sentinel Toolkit Combo Tool Documentation'.

5. Inclusion/Exclusion Criteria

The CIDA tool allows the application of additional inclusion/exclusion criteria for cohort selection. Inclusion/exclusion criteria can be defined using any combination of NDCs, procedure and/or diagnosis codes, and laboratory result values found in the SCDM. Procedure and diagnosis codes can be restricted to those observed in specific care settings (e.g., inpatient, outpatient) and diagnosis codes can be restricted by position (e.g., principal discharge diagnosis, secondary diagnosis). Each inclusion/exclusion criteria can further be defined by the number of days the code occurs.

Additional inclusion/exclusion criteria are assessed during a requester-defined number of days before, on, or after the index date. This allows assessments to be restricted, for example, to patients with a specific indication for treatment.

Patients are required to have continuous enrollment in the coverage type specified during the lookup period prior to index date for assessment of *exclusion* criteria but not inclusion criteria. If this condition is not met, the *exposure episode* is excluded from analysis. Patients are not required to have continuous enrollment post-index date for neither exclusion nor inclusion criteria.

6. Covariate Assessment

The CIDA tool can extract covariate information for a requester-defined number of days around the index date for all cohort identification strategies with the exception of the product utilization and switching cohort identification strategy.

Covariates can be defined using NDCs, procedure codes and/or diagnosis codes found in the SCDM. If NDCs are used, dispensings are processed via the stockpiling algorithm. Procedure and diagnosis codes can be restricted to those observed in specific care settings (e.g., inpatient, outpatient), and diagnosis codes can be restricted by position (e.g., principal discharge diagnosis, secondary diagnosis).

Patients are required to have continuous enrollment in the coverage type specified during covariate evaluation window prior to index date. If this condition is not met, the *exposure episode* is excluded from analysis. Continuous enrollment is not required post index date.

For Type 4 PSA analyses, the index date for the covariate assessment window can be defined by the user as pregnancy start date, date of exposure initiation, or delivery date/child birth date. Enrollment is required during this requester defined window prior to index date.

When a baseline table is requested for Type 4 analysis, the covariate assessment window is not assessed for continuous enrollment. The requester would need to ensure continuous enrollment during any period assessing covariates.

Note that the user also has the ability to use the Combo Tool (and combo codes) to define covariates or define covariates with laboratory result values.

The user has the option of specifying for the CIDA tool to generate the baseline covariate table (i.e., "Table 1") for the cohorts of interest. In order to achieve this, the user has to define a baseline period (i.e., a covariate lookback window) and a list of covariates of interest. CIDA will generate an output table containing the baseline prevalence of covariates of interest.

7. Charlson/Elixhauser Combined Comorbidity Score

The CIDA tool can calculate a combined Charlson/Elixhauser comorbidity score[2] for all patients in the cohort in all cohort identification strategies with the exception of the manufacturer-level product utilization and switching cohort identification strategy. The score is calculated based on comorbidities observed during a requester-defined window around the exposure episode start date (index date).

For requests that will use the PSA tool, the raw score is available to be used as a covariate to estimate the propensity score.

8. Medical Utilization Metrics

The CIDA tool can calculate medical utilization metrics for all patients in the cohort for all cohort identification strategies with the exception of the manufacturuer-level product utilization and switching cohort identification strategy. Medical utilization is defined as a number of encounters with the health system observed during a requester-defined number of days around the exposure episode start date (index date). Metrics can be calculated overall (total number of visits) or by encounter type (e.g., number of inpatient stays, number of emergency department visits, etc.).

For requests that will use the PSA tool, calculation of medical utilization is available to be used as a covariate to estimate the propensity score. Five metrics will be calculated and can be used as covariates to estimate the propensity score: number of 1) inpatient stays, 2) outpatient visits, 3) emergency department visits, 4) institutional stays, and 5) other ambulatory encounters (e.g., telemedicine, email consult). These five metrics correspond to the available encounter types in the SCDM.

9. Drug Utilization Metrics

The CIDA tool can calculate drug utilization metrics for all patients in the cohort for all cohort identification strategies with the exception of the manufacturer-level product utilization and switching cohort identification strategy. Metrics calculated are 1) number of dispensings; 2) number of unique generics dispensed; and 3) number of unique drug classes dispensed during a requester-defined number of days around the episode start date (index date).

For requests that will use the PSA tool, the calculation of drug utilization is available to be used as a covariate to estimate the propensity score.

10. Stratification of Results

The CIDA tool can stratify select results from all cohort identification strategies by age, sex, year, month, race, and certain geographic information. In all cohort identification strategies except the manufacturer-level product utilization and switching strategy the requester may also stratify by defined covariates. Custom strata may be defined in the CIDA tool from lists of valid stratification variables specific to each method of cohort identification.

[2] Gagne JJ, Glynn RJ, Avorn J, Levin R, Schneeweiss S. A combined comorbidity score predicted mortality in elderly patients better than existing scores. J Clin Epidemiol. 2011 Jul;64(7):749-59. doi: 10.1016/j.jclinepi.2010.10.004. Epub 2011 Jan 5.

C. EXPOSURES AND FOLLOW-UP TIME COHORT IDENTIFICATION STRATEGY

The exposures and follow-up time cohort creation strategy defines episodes of new use of a medical product of interest and evaluates the occurrence of HOIs. There are numerous requester options, including defining new use, exposed time, and episode censoring rules.

1. Identifying Exposure and Creating Exposure Episodes

An exposure can be defined using any set of NDCs, procedure and/or diagnosis codes, and laboratory result values found in the SCDM. Procedure and diagnosis codes can be restricted to those observed in specific care settings (e.g., inpatient, outpatient) and diagnosis codes can be restricted by position (e.g., principal discharge diagnosis, secondary diagnosis). For example, exposure to a drug product dispensed in the outpatient setting can be defined as observation of one or more NDCs in the pharmacy dispensing table, whereas exposure to a vaccine can be defined based on observation of specific procedure codes in the procedure table.

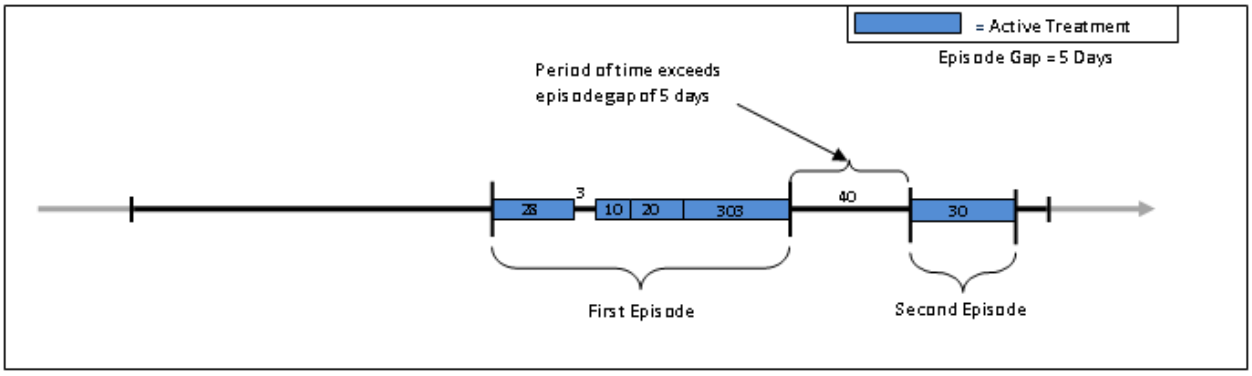
The CIDA tool queries the SDD and extracts all codes indicative of exposure during the query period. NDCs are processed and those with a part of their days supply outside enrollment episodes are truncated to constrain the supply within eligibility. Dispensing dates are modified using the stockpiling algorithm and supply is truncated again to make sure they are still in eligibility periods (stockpiling can push claims outside enrollment period).

After dispensing dates are adjusted using the stockpiling algorithm, exposure episodes are created. Exposure episodes can be defined in one of two ways: a) using outpatient pharmacy dispensing days supplied to create a sequence of continuous exposure, and b) defining a specific number of days after exposure initiation as exposed time.

a) Creating Exposure Episodes using Dispensing Days Supplied

An exposure episode using outpatient pharmacy dispensing days supplied is defined as a sequence of treatment that ends when interrupted by a gap in days supply greater than a requester-defined episode gap. Consider an example where five outpatient pharmacy dispensings of the exposure of interest are observed during the query period (Figure 2).

Figure 2. Exposure Episode Creation and Episode Gap



In this example, the CIDA tool is instructed to allow an episode gap of five days between drug dispensings. The "active treatment" (in blue) corresponds to the days supply for dispensings. Four dispensings make up the first exposure episode since there is only a three-day interruption in exposure

between the first and second dispensing; that gap is "bridged" to create a single exposure episode. The fifth dispensing, observed 40 days after the end of the previous dispensing's end of supply, will initiate a second exposure episode. Requesters also have the option to force a specified days suppy to a given code.

b) Creating Exposure Episodes as Defined Number of Days

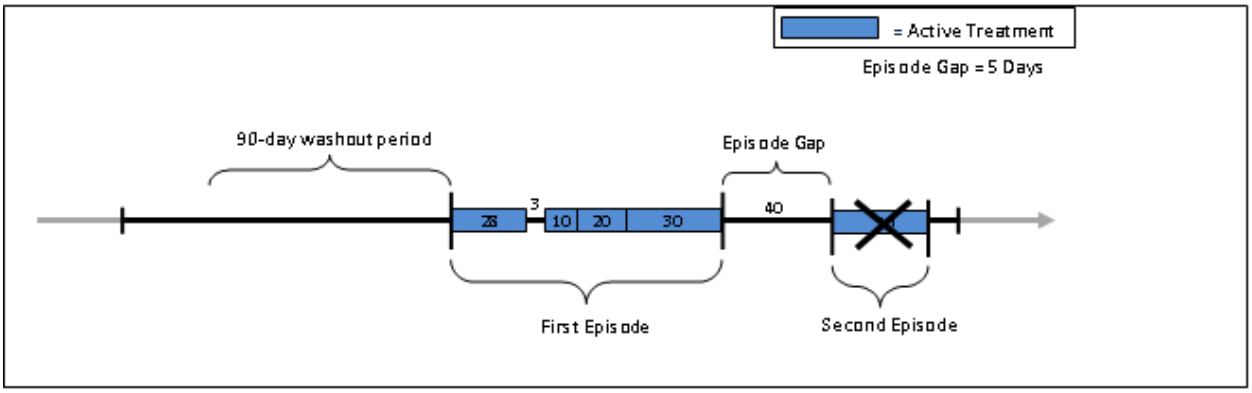
An exposure episode defined as a number of days after exposure initiation allows exposed time to be defined at the discretion of the requester. The same duration is specified for all exposed patients. Note, however, that due to censoring this does not necessarily mean that all patients have the same exposure duration.

c) Incident Medical Product Exposure

The CIDA tool is designed to identify both incident and prevalent use of a medical product of interest. To define new use, requesters must specify a number of days before the exposure index date that the patient must be free of the medical product of interest (i.e., a washout period). Users may determine whether the lookback period scans for evidence of a dispensings' days supply or for a dispensing date. If a new user cohort is requested, only exposure episodes meeting the requester-defined incidence definition are included in the cohort. Patients are required to have continuous enrollment in the coverage type specified during the washout period to ensure that new use can be assessed.

Consider the previous example where we observe two exposure episodes for a member during the query period (Figure 2). Suppose the requester requires a 90-day washout period to define new use of the exposure. In this case, given that the second exposure episode begins 40 days after the end of the first episode's last day of supply, the second episode is deemed ineligible for inclusion in the cohort (Figure 3).

Figure 3. Determining Incident Medical Product Exposure

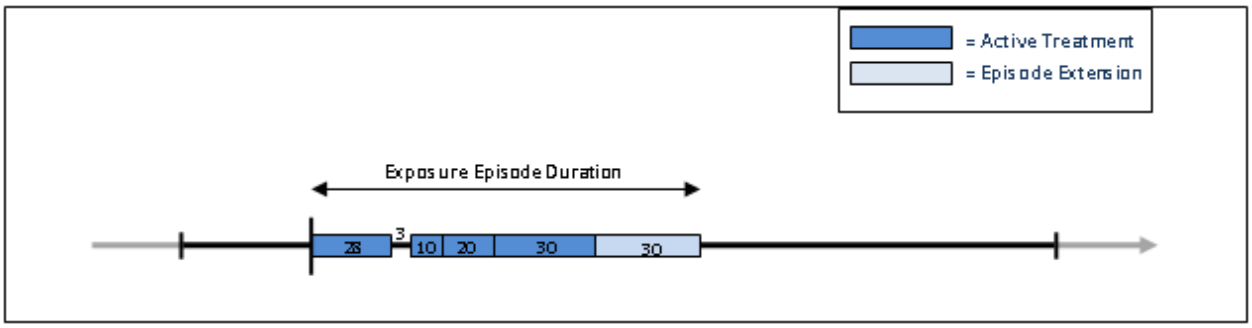


Instead of defining the number of washout days to assess incidence, requesters also have the option to require that a patient never have evidence of the exposure of interest during their entire available enrollment history (to identify the first observed use of the medical product). Note that while this option is available, interpretation may be challenging as available enrollment history is variable across patients.

d) Exposure Episode Extension

After medical product exposure incidence is assessed, the CIDA tool has the option to extend exposure episodes (and thus exposed time) by a requester-defined number of days. Note that this option is only relevant for exposure episodes created based on dispensing days supply. Figure 4 depicts a scenario where the episode extension is set to 30 days.

Figure 4. Exposure Episode Creation and Episode Extension



When an episode extension is used, the end date of the episode is adjusted to reflect the additional days of extension. These additional days contribute to days at risk metrics, and HOIs observed during extension days will be attributed to the exposure.

Note that if an episode extension extends into another exposure episode (*i.e.*, episode extension > allowable episode gap), the episode extension is truncated, and no "bridging" of exposure episodes occurs.

e) Exposure Episode Duration and Censoring

The exposure episode start date (or index date) is the date associated with the first code of interest used to define the episode. The exposure episode end date corresponds to the earliest of 1) end of exposure episode, 2) end of enrollment, 3) end date of SDD, or 4) occurrence of requester-defined censoring criteria (including an option to censor of query end date). The length of an exposure episode is defined as the difference between the episode end date and the episode start date plus one.

Exposure episodes may be censored based on requester-defined criteria. This could be based on the observation of any NDC, procedure code, diagnosis code, or laboratory result value of interest, or based on medical utilization like the occurrence of a hospitalization. If censoring criteria are observed during an exposure episode, the episode is truncated at the date of the observed criterion.

f) Number of Valid Exposure Episodes per Patient

Requesters have the ability to specify the number of valid exposure episodes each patient can contribute to the final cohort. Options include:

- Include only the first valid exposure episode during the query period: selects the first valid exposure episode during the query period that meets all requester criteria. Required option if output of the CIDA tool will be used for propensity score matched analyses.
- Include all valid exposure episodes during the query period: selects all valid exposure episodes during the query period that meet all requester criteria.
- Include all valid exposure episodes during the query period until an event occurs: selects all valid exposure episodes during the query period until an HOI is observed during exposure.

2. Identifying Health Outcome of Interest (HOI)

An HOI can be defined using any combination of NDCs, procedure and/or diagnosis codes, and laboratory result values found in the SCDM. Procedure and diagnosis codes can be restricted to those observed in specific care settings (e.g., inpatient, outpatient) and diagnosis codes can be restricted by position (e.g., principal discharge diagnosis, secondary diagnosis).

The CIDA tool queries the SDD and extracts all codes indicative of the HOI for exposed patients. If NDCs are used to define the HOI, they will undergo additional processing with the stockpiling algorithm.

a) Days at-Risk and Blackout Period

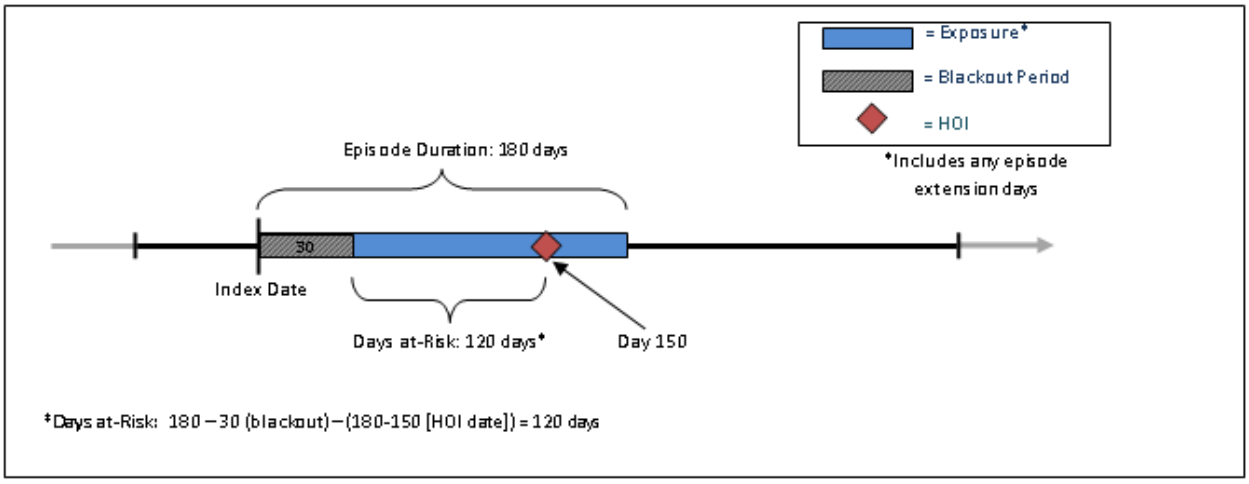
Days at-risk for the HOI are calculated as the number of days each patient is exposed and at-risk for the HOI. Days at-risk start accumulating at the exposure episode start date (or index date). Days at-risk stop accumulating at the earliest of 1) end of exposure episode (including episode extension days); 2) end of enrollment; 3) end date of SDD; 4) occurrence of requester-defined censoring criteria (including censoring at query end date); or 5) HOI date.

The CIDA tool has an additional option to include a "blackout" period after the exposure index date. This is a period of time on and after the index date during which HOIs are not attributed to the exposure. When a blackout period is defined, the effective start of the exposure episode remains the same, but days at-risk do not start accumulating until the end of the blackout period.

For example, if an episode starts on January 1st, 2010 and the blackout period is equal to 5 days, any HOIs occurring between January 1st, 2010 and January 5th, 2010 will not be attributed to the exposure episode. The episode start date, however, will remain January 1st, 2010. Note that if an HOI is observed during the blackout period, the *exposure episode* is discarded from the analysis.

Figure 5 illustrates HOI assessment and days at-risk accumulation when a blackout period is specified.

Figure 5. HOI Ascertainment, Blackout Period, and Days at Risk



In Figure 5, the number of days at risk begins accumulating at the end of the blackout period and extends until the occurrence of the HOI.

b) HOI Incidence Assessment

In addition to determining exposure episode incidence, requesters can also specify criteria for a new occurrence of an HOI. Specifying HOI incidence criteria instructs the CIDA tool to evaluate a requester-defined number of days before the *exposure episode index date* (not the HOI date) to determine if the HOI is "new." If an HOI is observed during the requester-defined number of days before the exposure index date, the *exposure episode* is discarded from analysis.

Patients are required to have continuous enrollment in the coverage type specified during the HOI incidence assessment period. If this condition is not met, the *exposure episode* is excluded from analysis.

3. Additional Enrollment Requirements

By default, the CIDA tool will require continuous enrollment in the coverage type specified before the index date for the longest duration defined by the following: 1) exposure washout period, 2) HOI washout period, 3) exclusion criteria lookup period (if specified), 4) most frequent use window (if specified) and 5) covariate evaluation window (if specified). Requesters have the option to define a longer enrollment period duration before index date.

By default, the CIDA will not require continuous enrollment in the coverage type specified for any evaluation windows that are after the index date. Users must explicitly define post-index enrollment requirements.

4. Optional Incidence Rate Ratio (IRR) Calculation

A standalone analysis tool compatible with the CIDA tool exposures and follow-up time cohort identification strategy output is available to calculate incidence rate ratios (IRRs) for exposure and comparator cohorts. The IRR tool utilizes a Poisson regression and a large sample approximation for calculation of the IRR; the tool may not be robust against samples with small HOI rates.

The tool outputs a comparison of two requester-defined cohorts from the CIDA tool output and provides both the unadjusted and adjusted IRRs and the corresponding 95% confidence intervals. The requester can adjust for any combination of age, sex, year of exposure, and Data Partner site. One output table is generated that contains: number of new users, person-years of follow-up, number of HOIs, incidence per 1,000 persons, and incidence rate per 1,000 person-years. Person-years of follow-up are estimated using person-days at risk standardized into years. Incidence per 1,000 persons is calculated as the number of HOIs divided by the number of new users, standardized to 1,000 persons. Incidence rate per 1,000 person-years is calculated as the number of HOIs divided by person-years of follow-up, standardized to 1,000 person-years.

5. Creation and Retention of First Valid Episodes

The CIDA tool goes through the following steps when determining a valid episode:

1. Creates a list of all potential index dates. These are dates with cohort defining codes that meet washout criteria. Additionally, in analyses where exposed time is requester-defined number of days, that exposed time is treated as a washout.
2. Using the list created in step (1), a list of all potential index dates that do not meet incidence criteria gets created.
3. Using the list created in step (1), a list of all potential index dates that meet both inclusion and exclusion criteria gets created.
4. Using the list created in step (1), episodes get created.
    a. When exposed time is assessed using dispensings' days supply, episodes are created by bridging claims (using the allowed number of days between two consecutive claims to consider them as part of the same treatment episode). Subsequent claims are automatically bridged. Due to this, except in the scenario for exposure extension, episodes don't get censored because of the start of a subsequent episode.
    b. When exposed time is requester-defined number of days, then the number of days after exposure initiation is what is considered "exposed time." After creating episodes, exposure extension is applied and follow-up truncated if the exposure extension overlaps a subsequent episode.
5. Creation of a master list of potential episodes by taking episodes from step (1), removing those that are in step (2), and removing those that aren't in steps (3 and 4).
6. Using this master list of patients, CIDA a) Applies enrollment criteria; b) Truncates episodes based on whether codes should be used to assess cohort index incidence only or enrollment end date or query end date or death or end of Data Partner data; c) Applies Minimum episode duration and maximum episode duration (NOTE: this is applied to the episode after exposure extension has been applied); and d) Removes episodes with no at risk time because of blackout.
7. Minimum day supply criteria is applied. This is not done in step 6 because actual dispensing information (without episode gaps/extensions) needs to be gathered.
8. Finds all events and removes episodes with an event during the event washout.
9. Applies the requirement of how many index dates an individual can contribute to the cohort (only the first valid index date or all valid index dates per individual during the query period). This is the very last criteria that gets applied to the cohort selection. All valid episodes are gathered, and if specified by the requester, then restricted to only the first valid index date. This means that if an individual has 5 index dates in step (1), and the first 3 aren't valid (no inclusion criteria, doesn't meet enrollment requirements, etc.), the 4th index date is still eligible for inclusion when only the first valid index date per individual during the query period is requested.

6. Creation of a Never-Exposed Cohort and Identification of Events

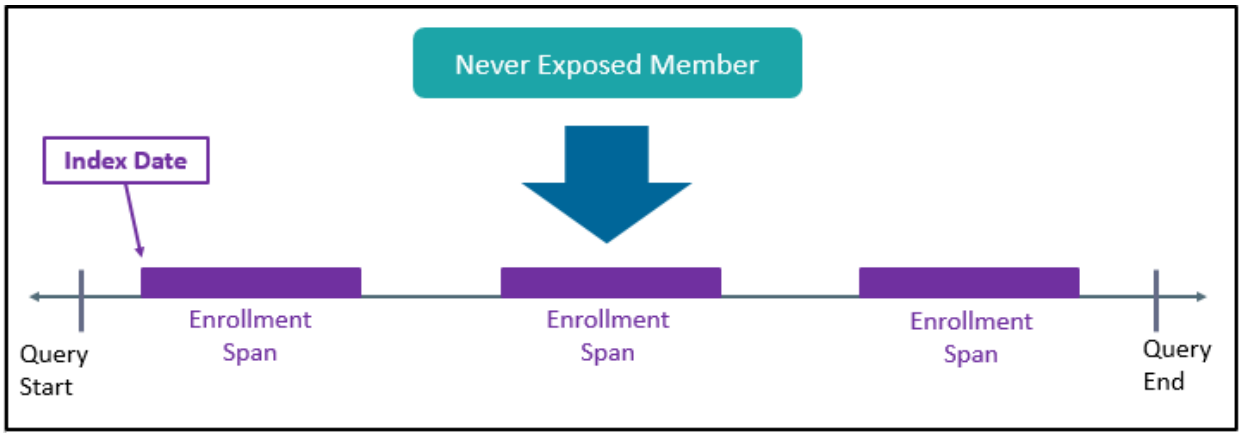
After identifying an exposed cohort in a Type 2 (exposures and follow-up time) analysis, the CIDA tool creates a never-exposed cohort by bridging all enrollment spans that meet the demographic criteria and coverage type and removing enrollment spans with:

1. Days that do not meet lookback and lookforward enrollment criteria
2. Days outside of the query period
3. Days outside age-group criteria
4. Days that do not meet inclusion and exclusion criteria
5. Days that do not meet washout criteria
6. Days that do not meet follow-up event/washout criteria
7. Days that do not have at least 1 day of follow-up during query period after taking the blackout period into account

Then, CIDA will remove any members with evidence of an index defining code (i.e., exposure) at any point during the member's enrollment history. The result will be a list of enrollment spans during the query period that meet all enrollment criteria, index incidence, event incidence, demographic criteria, inclusion/exclusion criteria, and do not have evidence of an index-defining code.

After identifying all never-exposed enrollment spans, CIDA will take the 1st enrollment span per member (PATID) and set the index date to the 1st eligible day. Figure 6. Determine a Never-Exposed Member's Index Date and Enrollment Time illustrates how the index date is determined.

Figure 6. Determine a Never-Exposed Member's Index Date and Enrollment Time



CIDA will remove episodes that do not meet the minimum episode duration (MinEpisDur) requirement.

If exposure is defined using outpatient pharmacy dispensing days supplied, the episode end date will be the earliest of:

- End of query period (if censor_qry = Y)
- DP end date (if censor_dp = Y)
- Disenrollment (end of eligible enrollment span)
- Evidence of death (if censor_dth = Y)
- The maximum episode duration has been reached (MaxEpisDur)
- Occurrence of requester defining censoring criteria

If exposure is defined using user-defined number of days of follow-up (ITTDAYS), the episode end date will be the earliest of:

- End of user defined number of days of follow-up (ITTDAYS)
- End of query period (if censor_qry = Y)
- DP end date (if censor_dp = Y)
- Disenrollment
- Evidence of death (if censor_dth = Y)
- Occurrence of requester defining censoring criteria

Once the never-exposed cohort has been defined, members will be followed for events. CIDA will output the number of events and the first event date. Time to event (TTE) will be censored at occurrence of the first event. If the censoring output table is requested for the exposed cohort, the never-exposed cohort will be included in the censor table.

If requested, CIDA will extract covariates, medical utilization, drug utilization, calculate comorbidity score, and output a baseline table for the never-exposed cohort.

If requested, CIDA will perform MFU analysis, determine covariate profile and extract HDPS files for the never-exposed cohort.

Group name will be the group name of the exposed cohort concacentated with the suffix "_nvrexp." For example, if the group name is DrugA, the never-exposed group name is DrugA_nvrexp.

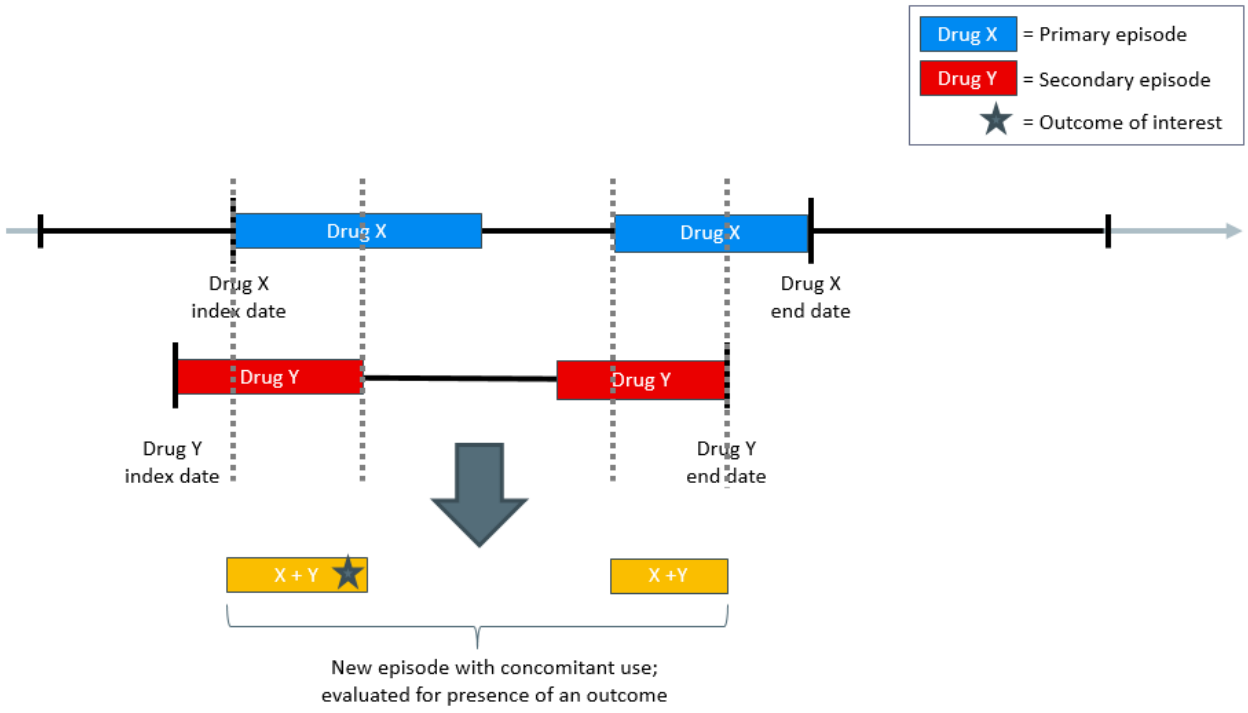
7. Identifying Episodes of Concomitant Use

a) Overview

Using the Concomitant Use Tool, requesters may create concomitant treatment episodes by looking at primary and secondary treatment episodes and bringing together treatment durations where the two occur concurrently. They may then assess the health outcomes of interest (HOI) during periods of concomitant use. This is achieved by using the start and end dates of the primary and secondary episodes to capture periods of concomitant use and to create a concomitant use episode.

The requester also has the ability to examine either the first concomitant episode or all concomitant episodes. Another feature of this tool is that the requester can indicate whether to always require the primary exposure to be initiated before the secondary exposure, restrict concomitant episodes to those where the primary and secondary episodes are initiated on the same day or not enforce any restrictions. Once the concomitant episodes are created, the program can then examine HOI that happened during those episodes. The user is able to set an event blackout for concomitant period by specifying a period at the start of a concomitant treatment episode during which valid events found by the concomitant algorithm are to be ignored. Figure 7 below illustrates evaluation for outcomes of interest during concomitant use episodes.

Figure 7. Evaluation of Outomes During Concomitant Treatment Use



In the figure above, Drug X represents the primary episode and Drug Y is the secondary episode. Days for which treatments X and Y overlap create concomitant use episodes. This tool can then look for HOI that occur during the concomitant treatment. In this example, the outcome of interest occurs once.

b) Output

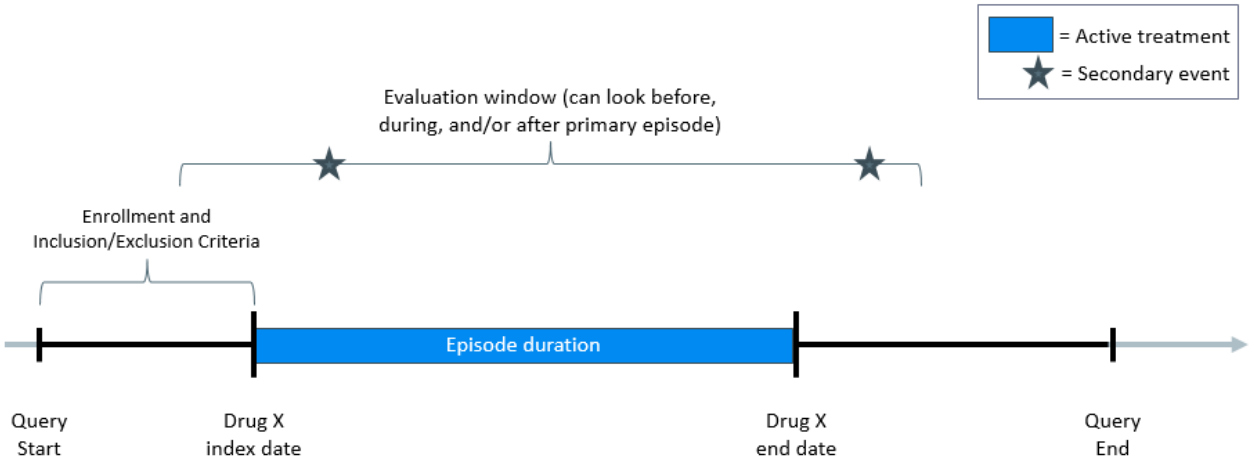
Concomitant Use Tool automatically generates an output table that characterizes the concomitant use of two medical products by producing metrics such as the number of individuals who had concomitant use episodes, number of concomitant use exposure episodes, dispensings, days supply, concomitant use episodes with HOI, total number of all HOI in all concomitant use exposure episodes, days at-risk for an HOI.

8. Identifying Multiple Events

a) Overview

The Multiple Events Tool allows the requester to specify a primary treatment episode, define an observation window relative to that primary episode, and evaluate the occurrence of multiple secondary events. Events can be defined as an interval (i.e., an episode) or as a single point in time. The tool gives users the flexibility to specify the observation window to be before, during or after the primary treatment episode. Secondary cohort events are only considered if they fall in a requester-defined observation window. Figure 8 below illustrates multiple events assessment.

Figure 8. Evaluation of Multiple Events



In this figure, Drug X represents the primary episode. The evaluation window for the secondary event is set based on index and episode end dates of Drug X. In this example, the secondary event occurs twice, once during the primary episode and the second time in the post-primary episode window. The Multiple Events tool enforces enrollment criteria through the entire defined observation window.

b) Adherence

Requesters have an option of specifying multiple criteria to determine overall treatment adherence via requester-defined primary episode duration, minimum number of secondary events, time to first secondary event, and secondary event gap. Any number of combinations of these metrics can be used to define adherence.

c) Output

The Multiple Events Tool automatically generates a table of primary and secondary episode characteristics, such as number of patients with multiple events, number of multiple event episodes, total duration of primary episode, number of secondary episodes, primary episodes with no/at least one secondary episode(s), time to user-specified secondary episode. Adherence measures such as the number of episodes and patients that meet adherence definition are also included if the user requests the optional multiple events adherence.

9. Identifying and Characterizing Treatment Overlap

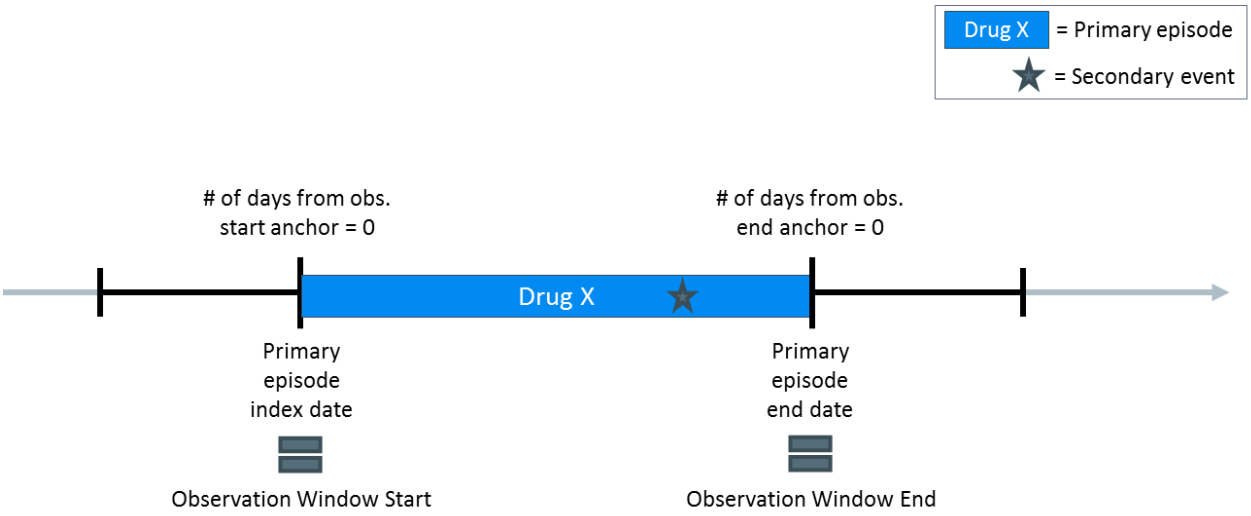
a) Overview

The Overlap Tool can be used to characterize the overlap between primary and secondary treatment episodes during the observation window. The observation window is user-defined relative to the first primary treatment episode, during which the occurrence of secondary episodes are evaluated. The tool gives users the flexibility to specify the observation window to be before, during or after the primary treatment episode. Secondary episodes are only considered if they fall in a requester-defined observation window. This additional analysis is an add-on to and is relevant for the Exposures and Follow-up time Cohort Identification Strategy only.

Using this tool, requesters may define a primary treatment episode and an observation window around that primary episode, and then assess the secondary event (episode or a single point in time). Along with overlap between primary and secondary treatment episodes during the observation window, the user is also able to optionally assess for "adherence" to user-defined thresholds for the % or days overlap between two treatment episodes.

Figure 9 below illustrates an example of how the observation period can be created.

Figure 9. Creation of the Observation Window, Example 1

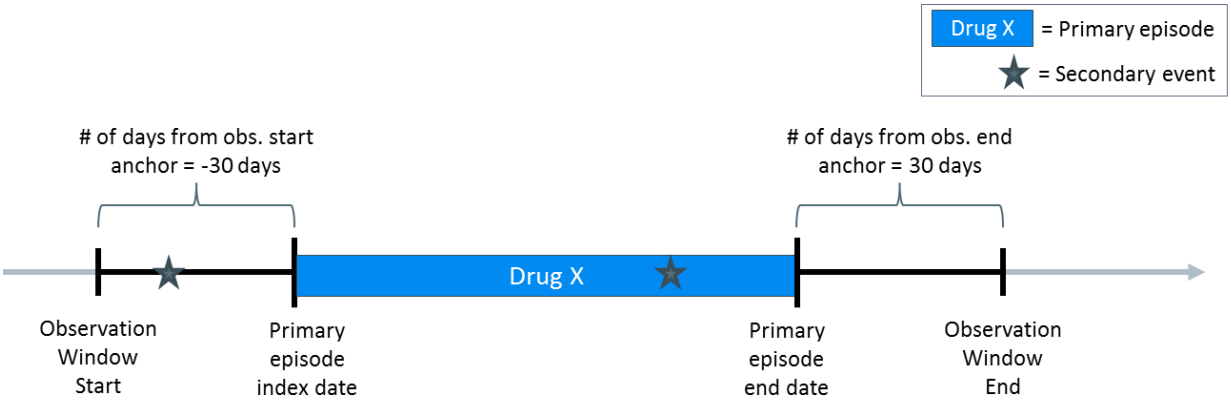


In order to define the observation window, the user has to decide on what to anchor the calculation of the observation window start and end dates. There are two anchors that need to be specified: an anchor for observation start and an anchor for observation end. For each one, the user has two options to pick from: primary episode index or end date.

In Figure 9 above, observation start is anchored on primary episode index date and observation end is anchored on primary episode end date. After setting the anchors, the user can then specify the number of days from those anchors to look for the secondary event. In this example, those number of days are set to zero, which is why the observation window ends up being from primary episode's index date through end date. There is one secondary event captured in this example.

Alternative setup is shown in Figure 10 below, where after setting the anchors to be primary episode index and end dates, the user specified the number of days from observation start anchor to be -30 and the number of days from observation end anchor to be 30. The Overlap Tool forces enrollment criteria through the entire defined observation window.

Figure 10. Creation of the Observation Window, Example 2

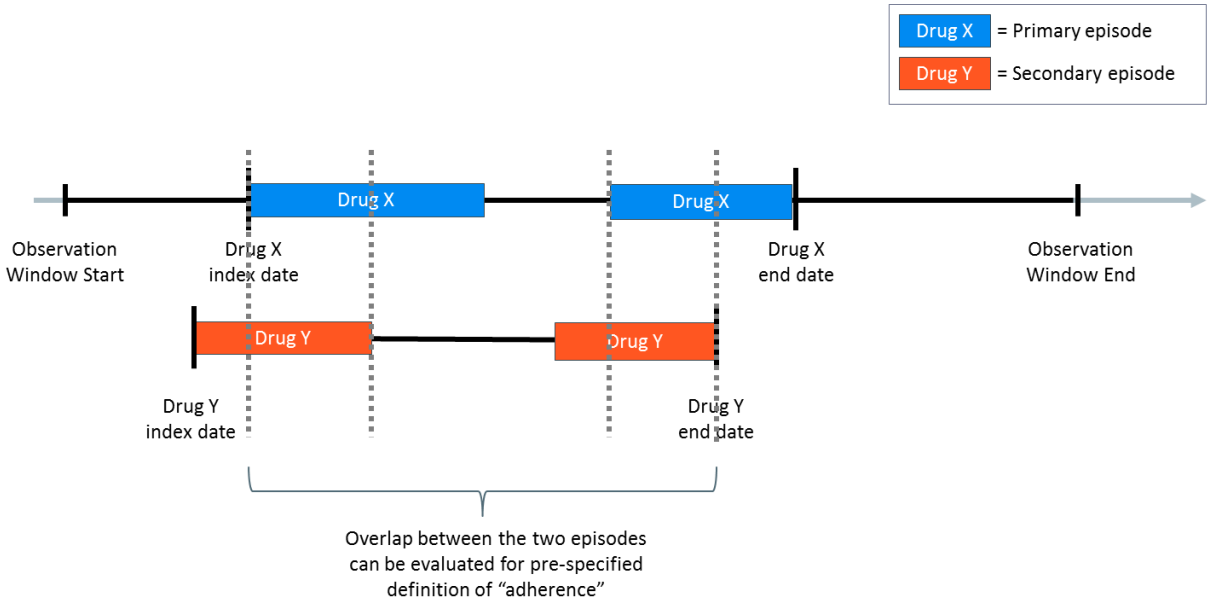


In this 2nd example, the program will evaluate for the presence of the secondary event one month prior to primary episode index through one month after primary episode end date. There are two such secondary events captured in this example.

b) Adherence

Requesters also have an option to define adherence to user-defined expectations of patient behavior. This will allow for assessment of adherence to clinical recommendations as defined by specified % or days overlap between the primary and secondary treatment episodes. The requester may specify multiple criteria to determine overall adherence to overlapping treatments. In **Figure 11** adherence may be based on minimum % overlap between the two episodes Drug X and Drug Y.

Figure 11. Evaluation of Overlap and Adherence



Depending on how the observation window is defined, the % overlap will produce the % overlap between observation window and secondary episode (if the observation window falls outside of the primary episode) or it will produce the % overlap between the primary and secondary episodes (if the observation window includes the primary episode). Multiple criteria to measure adherence can be

specified and any number of combinations of minimum and maximum overlap between secondary episodes and observation window can be used to define adherence.

c) Output

The Overlap Tool automatically generates a table of primary and secondary episode characteristics, such as number of patients with overlap episodes, number of overlap exposure episodes, days supply for primary episode dispensings, days supply for secondary episode dispensings, number of index defining codes for primary and separately for secondary episodes, total duration of primary episode, number of overlap days, number of primary episodes with no/at least one secondary episode(s), the number of users with no/at least one secondary episode(s), adherence measures (if requested), overall number of individuals eligible to have a primary episode and the total number of days those individuals are eligible to have an index date.

D. SELF-CONTROLLED RISK INTERVAL (SCRI) DESIGN COHORT IDENTIFICATION STRATEGY

The self-controlled risk interval (SCRI) design cohort identification strategy defines new use of a medical product of interest, identifies a risk and control window relative to exposure, and examines the occurrence of HOIs. Risk and control windows may be of the same or different duration, and the control window may be specified before exposure or after the risk window. To avoid bias by contraindication, requesters specifying a control window before exposure should have confidence that the occurrence of an HOI does not influence receipt of treatment.[3]

Two cohorts are identified using the self-controlled design: an exposure cohort and an analytic cohort. The exposure cohort includes patients with the exposure of interest that meet cohort inclusion criteria; the analytic cohort is a subset of the exposure cohort that includes patients that also have an HOI during the risk and/or control windows and sufficient post-exposure continuous enrollment.

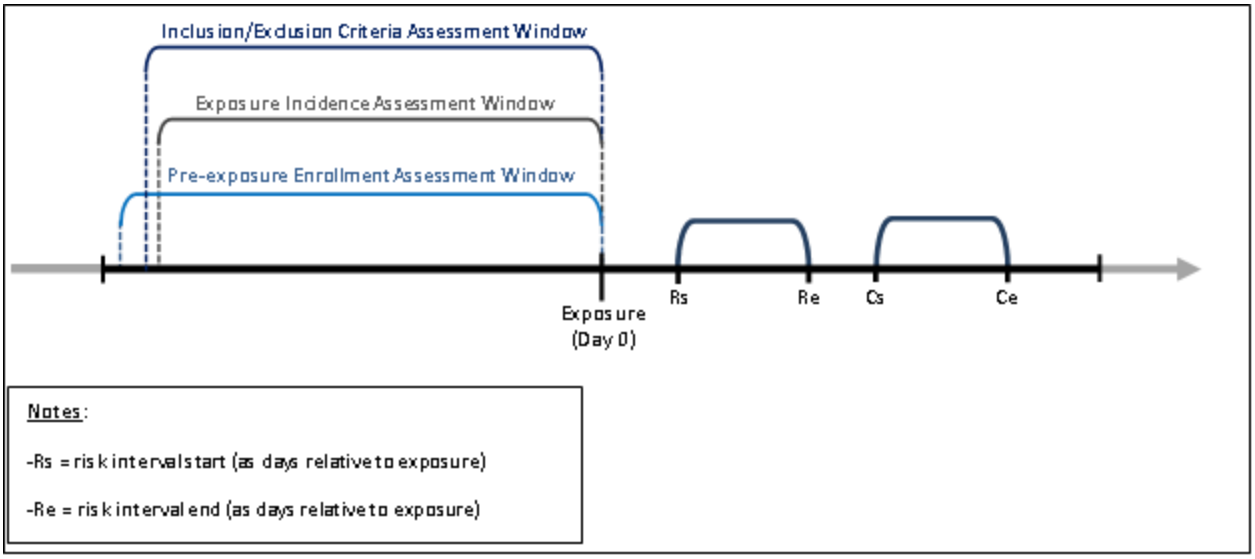
While the analytic cohort is the one of interest for prospective surveillance activities, delineating an exposure cohort allows requesters to characterize users of medical products of interest regardless of the occurrence of an HOI. The sections below provide additional details on requirements for exposure and analytic cohort entry.

1. Exposure Cohort

To be included in the exposure cohort, patients must have a valid exposure of interest. Valid means that all pre-exposure enrollment, incidence, and inclusion/exclusion criteria specified by the requester are met. Criteria are assessed during a requester-defined number of days before exposure; assessment window duration may vary by criterion (Figure 12).

[3] Platt R, Archdeacon P, Bell C, et al. Mini-Sentinel Methods: Prospective Routine Observational Monitoring Program Tools (PROMPT) Users' Guide, Version 1.0. Appendix 2, Section 1. June 13, 2014. Available at: https://www.sentinelinitiative.org/sites/default/files/Methods/Mini-Sentinel_PROMPT_Users-Guide_0.pdf [Accessed January 9, 2015].

Figure 12. Defining an Exposed Cohort



An exposure can be defined using any set of NDCs, procedure and/or diagnosis codes, and laboratory result values found in the SCDM. Procedure and diagnosis codes can be restricted to those observed in specific care settings (e.g., inpatient, outpatient) and diagnosis codes can be restricted by position (e.g., principal discharge diagnosis, secondary diagnosis). For example, exposure to a drug product dispensed in the outpatient setting can be defined as observation of one or more NDCs in the pharmacy dispensing table, whereas exposure to a vaccine can be defined based on observation of specific procedure codes in the procedure table.

The CIDA tool queries the SDD and extracts all codes indicative of exposure during the query period. NDCs are processed and dispensing dates are modified using the stockpiling algorithm. To define new medical product use, requesters may specify a number of days before the exposure index date that the patient must be free of the medical product of interest (i.e., exposure incidence assessment period). If a new user cohort is requested, only exposures meeting the requester-defined incidence definition are included in the cohort.

a) Pre-exposure Enrollment Requirements

At minimum, patients in the exposure cohort must be continuously enrolled in the coverage type specified for the duration of the 1) exposure incidence assessment period, 2) pre-exposure evaluation period for exclusion criteria, and 3) HOI incidence assessment period. If desired, requesters may specify longer pre-exposure enrollment requirements.

2. Analytic Cohort

To be included in the analytic cohort, patients in the exposure cohort must have an incident HOI in the risk or control window.

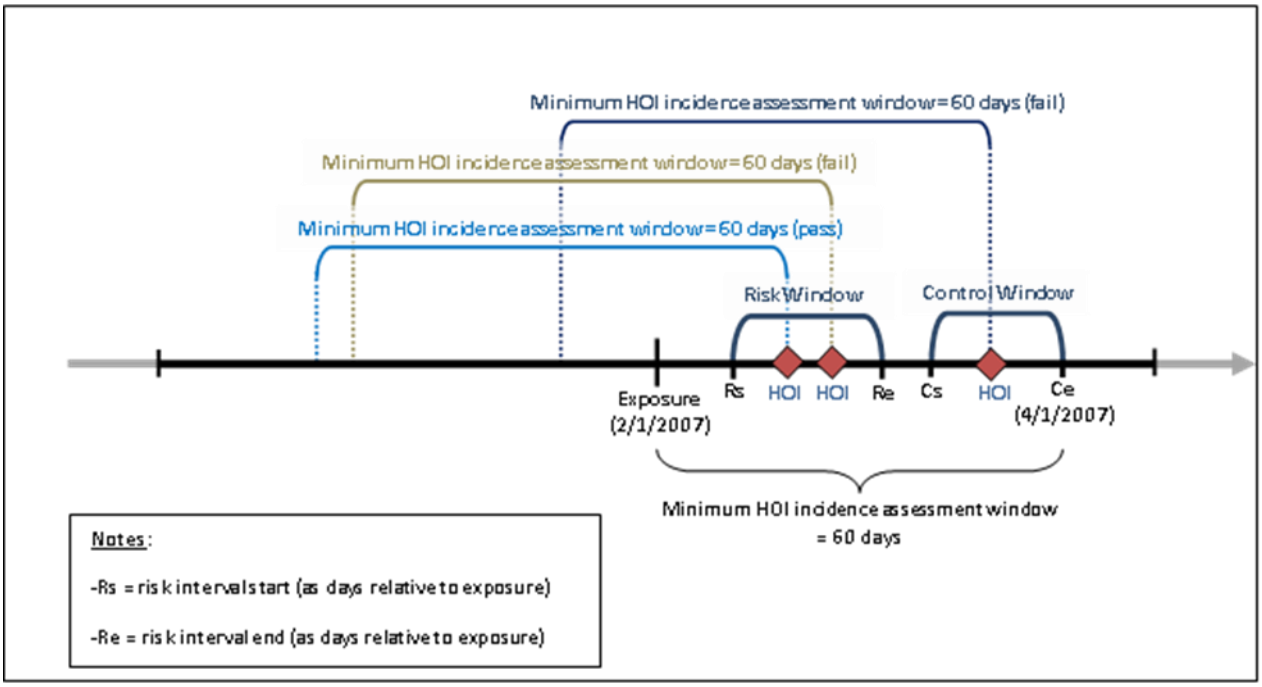
a) HOI Incidence Assessment

An HOI can be defined using any set of NDCs, procedure and/or diagnosis codes, and laboratory result values found in the SCDM. Procedure and diagnosis codes can be restricted to those observed in specific care settings (e.g., inpatient, outpatient) and diagnosis codes can be restricted by position (e.g., principal discharge diagnosis, secondary diagnosis). Unlike the exposures and follow-up cohort identification strategy, the SCRI design requires that HOI incidence is assessed relative to the HOI date (rather than the exposure date). This means that incidence criteria are only assessed if an HOI is observed during the risk or control window. Note also that the CIDA tool enforces a minimum HOI incidence assessment period for the self-controlled risk interval design, to ensure that patients can only contribute an incident HOI to either the risk or control window (but not both). The minimum HOI incidence assessment period is:

[Maximum (Risk interval end date, Control interval end date, Exposure date)] – [Minimum (Control interval start date, Risk interval start date, Exposure date)] + 1 days in duration.

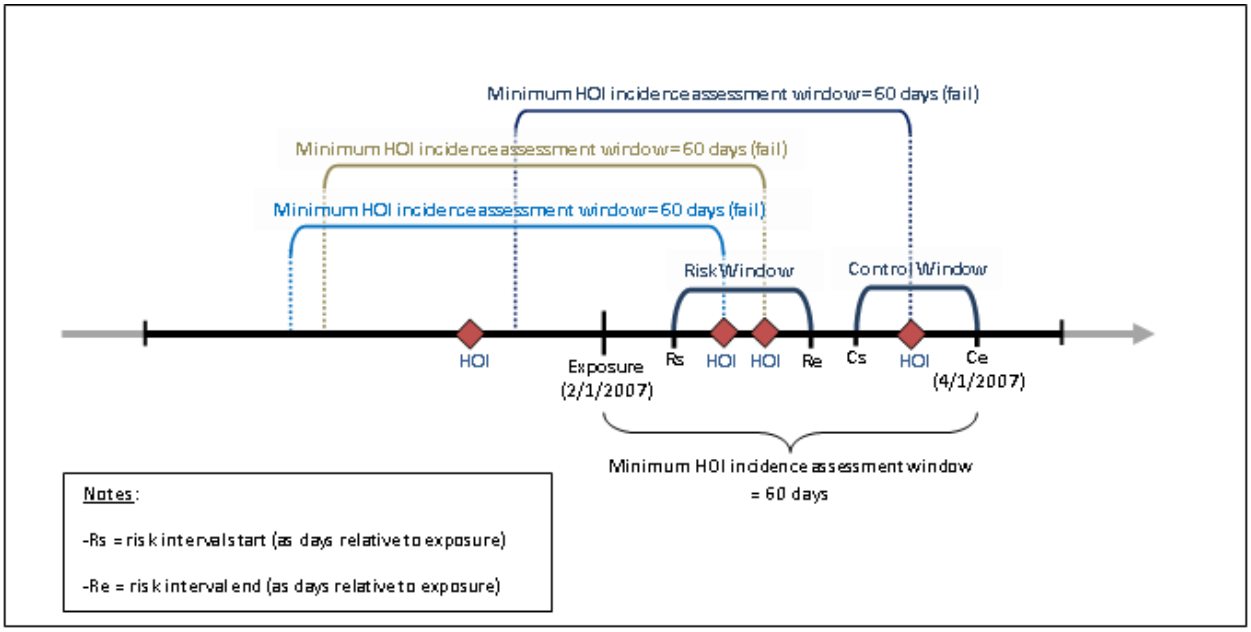
Figure 13 displays a scenario where three HOIs are observed during the risk and control windows. Because the minimum HOI incidence assessment period is 60 days (4/1/2007 – 2/1/2007 + 1 days) only the first HOI can potentially be considered an incident HOI. In this example, the patient will contribute an HOI to the risk window for the analysis.

Figure 13. Minimum HOI Incidence Assessment Period (HOI in Risk Window)



However, if the first HOI did not meet incidence criteria (Figure 14), the patient would not be included in the analytic cohort (as there are no valid HOIs in the risk or control window).

Figure 14. Minimum HOI Incidence Assessment Period (No Valid HOI)

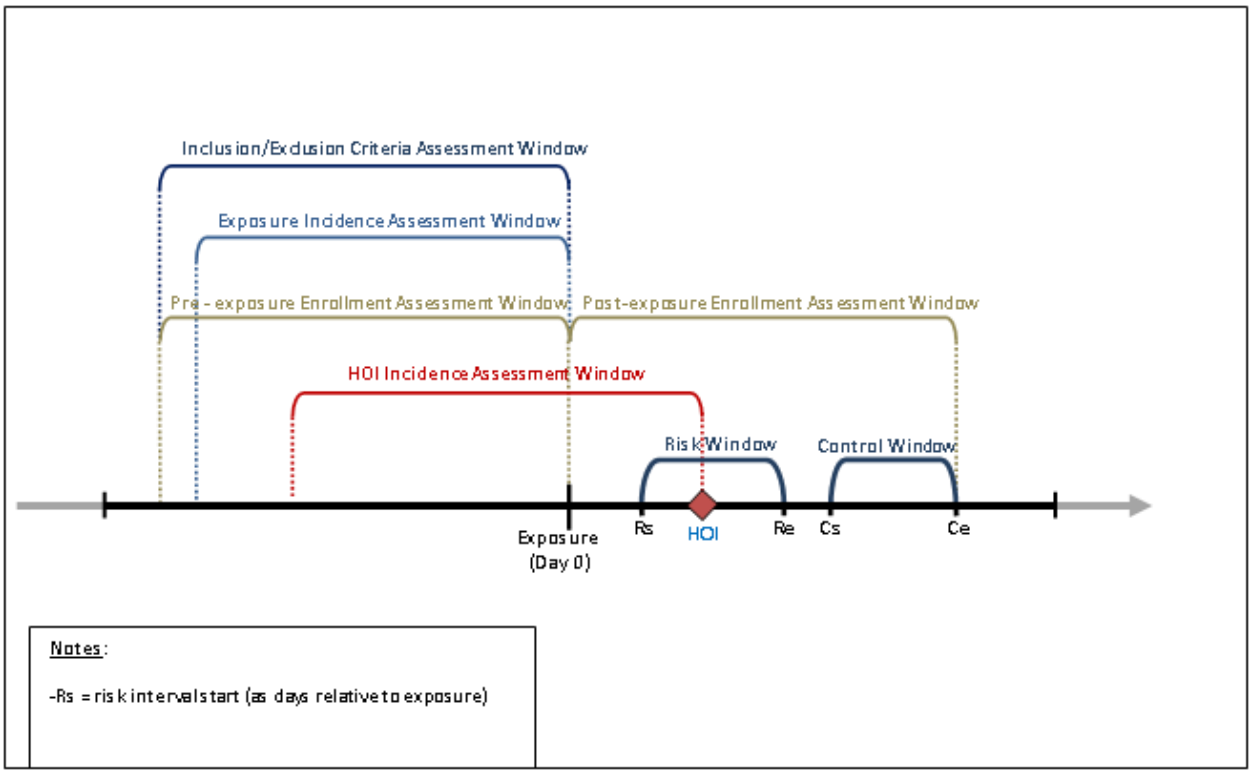


b) Post-exposure Enrollment Requirements

Patients must also have sufficient post-exposure continuous enrollment to contribute to the analytic cohort. If the control window is after the risk window, patients must be continually enrolled from the exposure date through the end of the control window. If the control window is before exposure, patients must be continually enrolled from the control window start date to the risk window end date.

Figure 15 displays the requirements for analytic cohort entry, using an observed incident HOI in the risk window as an example (i.e., patients may also enter the analytic cohort if an HOI is observed in the control window).

Figure 15. Defining an Analytic Cohort (Example with HOI in Risk Window)



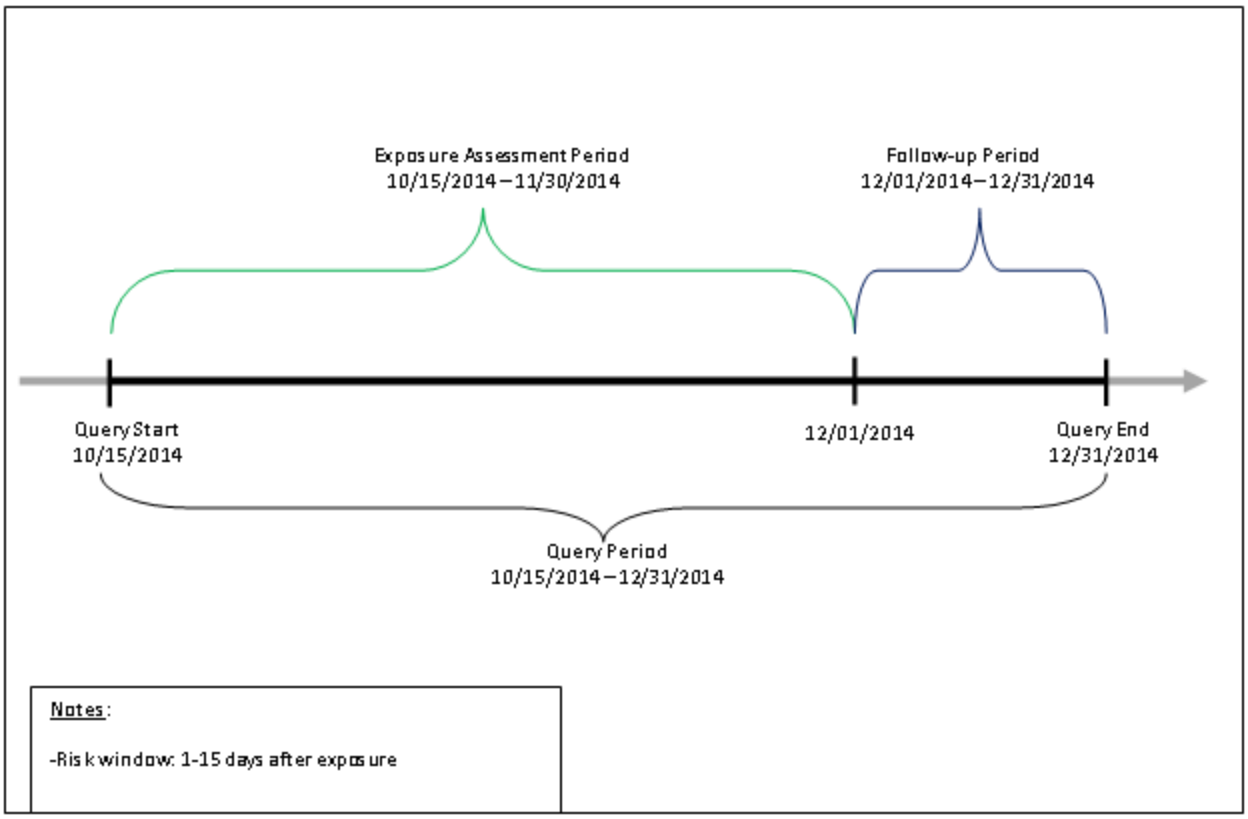
3. Characterizing Exposed Patients Excluded from Analytic Cohort

Patients in the exposure cohort may be excluded from the analytic cohort due to 1) insufficient post-exposure enrollment during risk and control windows, 2) absence of an incident HOI in the risk or control window, 3) insufficient enrollment prior to risk window start date or control window start date or 4) indication of death (optional). For all patients excluded from the analytic cohort, the CIDA tool will characterize the reason for exclusion. This allows requesters to determine whether patients were excluded due to insufficient follow-up time or lack of valid HOIs.

4. Exposure Assessment and Follow-up Periods

Because the SCRI design requires post-exposure eligibility, the query period and the time period for which patients may contribute an exposure to the cohort are not the same. Consider an example where a Data Partner's database contains information through December 31, 2014. A requester is examining exposure to a new vaccine with a risk window 1-15 days after exposure, a control window 20-31 days after exposure, and a query start date of October 15, 2014. Since patients are required to be enrolled for the duration of the follow-up period (i.e., during the risk and control windows), the latest date a patient can potentially contribute an exposure to the cohort is November 30, 2014 (with 31 days post-exposure follow-up). Therefore, it is useful in this design to differentiate the exposure assessment period (in the above example, the query period start date through November 30, 2014) and the follow-up assessment period (December 1, 2014 -December 31, 2014) as distinct time periods within the query period, since no patient will be able to contribute an exposure from December 1, 2014 -December 31, 2014 (Figure 16).

Figure 16. Exposure Assessment and Follow-up Periods



Note that the exposure assessment period will always be rounded down to the latest complete month. That is, if in the example above the required follow-up period duration was 25 days, the exposure assessment period would still end on 11/30/2014 (as 12/6/2014 would be rounded down to 11/30/2014).

5. Data Completeness

Another consideration for the SCRI design is the concept of Data Partner data completeness. In the above example, it was noted that the Data Partner's database contained information through December 31, 2014. While it is useful to know the latest date for which information is available for a Data Partner, a more useful date is one that represents a date where information is likely to be complete (i.e., all or most claims have been adjudicated, no information is missing due to lags in processing or receipt of information, etc.). In the context of the SCRI design cohort identification strategy, setting an appropriate query end date and Data Partner data completeness date improves confidence that HOIs that occur during the follow-up period are captured by the program. Determining Data Partner data completeness involves a combination of examining available data and working with partner sites to understand potential lags in data receipt and adjudication.

E. BACKGROUND RATE CALCULATION COHORT IDENTIFICATION STRATEGY

The background rate cohort identification strategy identifies prevalent or incident use of an event of interest (i.e., exposure to a medical product or occurrence of an HOI) during the requester-defined query period. This type of request is often used in preparation for a more in-depth analysis, to determine the rate of new use of a medical product, or the prevalence/incidence of an HOI in the SDD. There are several requester options, including defining incidence and additional inclusion/exclusion criteria.

1. Identifying Events

An event can be defined using any set of NDCs, procedure and/or diagnosis codes, and laboratory result values found in the SCDM. Procedure and diagnosis codes can be restricted to those observed in specific care settings (e.g., inpatient, outpatient), and diagnosis codes can be restricted by position (e.g., principal discharge diagnosis, secondary diagnosis). For example, exposure to a drug product dispensed in the outpatient setting can be defined as observation of one or more NDCs in the pharmacy dispensing table; occurrence of an HOI can be defined based on observation of specific diagnosis codes in the diagnosis table.

The CIDA tool queries the SDD and extracts all codes indicative of the event during the query period. NDCs identified undergo additional processing through the stockpiling algorithm.

2. Event Incidence

If an incidence rate calculation is needed, requesters must specify the criteria to define an incident occurrence of the event of interest. Specifying incidence criteria instructs the CIDA tool to evaluate a requester-defined number of days before the event date (index date) to determine if the event is "new."

Patients are required to have continuous enrollment in the coverage type specified during the incidence assessment period. If this condition is not met, the event is excluded from analysis.

3. Number of Valid Events per Patient

Requesters have the ability to specify the number of valid events each patient can contribute to the final cohort. Options include:

- Include the first valid event per patient that meets all requester criteria during the query period.
- Include all valid events per patient that meet all requester criteria during the query period.

4. Eligible Patients and Eligible Days

In order to calculate prevalence and incidence, the program calculates both the number of eligible patients and the number of eligible days for the denominator(s).

a) Eligible Patients

For prevalent cohorts, the number of eligible patients includes those patients enrolled in the coverage type and demographic criteria specified for at least one day during the query period. For incident cohorts, the number of eligible patients includes those patients enrolled in the coverage type and demographic criteria specified for at least the washout period number of days plus one day during the

query period. If additional inclusion/exclusion criteria are specified, the patient must also satisfy those requirements to be included.
Eligible Days
Eligible days for prevalent and incident cohorts are calculated as all the days during the query period that an eligible patient is eligible for inclusion in the cohort.

F. PREGNANCY EPISODES COHORT IDENTIFICATION STRATEGY

The pregnancy episodes cohort identification strategy identifies live births and calculates the length of the pregnancy episode based on ICD-10-CM codes indicative of weeks of gestation, ICD-10-CM and ICD-9-CM codes for preterm and postterm deliveries. A comparator episode from a woman of the same age, meeting all inclusion/exclusion criteria and likely to not have a live birth is also selected. This type of request is used to characterize medical product use during a requester-specified number of days prior to the start of pregnancy, and during each trimester. There are several requester options, including defining how to calculate pregnancy start date (index date) and defining how to categorize medical product use.

When used with the PSA tool, maternal and infant health outcomes of interest following maternal exposure can be evaluated. Users have the option of identifying 1) a pregnant cohort with an exposure of interest, 2) a pregnant cohort with a comparator or interest, and/or 3) an unexposed pregnant cohort. Maternal outcomes can be assessed for all pregnancy episodes and infant outcomes can be assessed for live births included in the Sentinel Common Data Model (SCDM) Mother-Infant Linkage (MIL) Table. There are several requester options, including specifying which date to use for covariate evaluation period and propensity score risk-set creation.

1. Identifying Live Births

Live births can be identified by one of two strategies:

1. A live birth can be defined using any set of NDCs, procedure and/or diagnosis codes, and laboratory result values found in the SCDM. Procedure and diagnosis codes can be restricted to those observed in specific care settings (e.g., inpatient, outpatient), and diagnosis codes can be restricted by position (e.g., principal discharge diagnosis, secondary diagnosis).

   By default, live births are defined using a code list developed by the Medication Exposure in Pregnancy Risk Evaluation Program (MEPREP) and other Sentinel and non Sentinel pregnancy-related projects, but it is adaptable as needed. Identification of live births based on the original work in MEPREP has been updated to incorporate ICD-10-CM and ICD-10 Procedure Coding System (ICD-10-PCS) codes.

2. A Live birth can also be defined using the the Sentinel Common Data Model (SCDM) Mother-Infant Linkage (MIL) Table. The Mother-Infant Linkage Table identifies live birth deliveries and matches mother patient IDs to infant patient IDs when possible. All identified live birth deliveries can be evaluated or deliveries can be restricted to those with an infant match. If there is a discrepancy between the delivery date and the infant's birth date, the infant's birthday from the MIL table will replace the delivery date if a user-defined tolerance (i.e., number of days between birth date and delivery date) in not exceeded. The pregnancy episode is excluded from the analysis when tolerance is exceeded. Only singleton births are included.

2. Calculating Start of Pregnancy Index Date and Length of Pregnancy Episode

In the other cohort identification strategies, the index date is determined via the occurrence of a code of interest. In the pregnancy episodes cohort identification strategy, delivery date is determined by identifying a live birth, and the index date (start of pregnancy episode) must be calculated by subtracting the length of the pregnancy episode from the date of live birth. The length of the pregnancy episode can be defined by using a list of prioritized gestational age codes. Overall, this algorithm prioritizes ICD-10 codes over ICD-9-CM codes to assign gestational age. These codes can be any set of NDCs, procedure and/or diagnosis codes, and laboratory result values found in the SCDM. Procedure and diagnosis codes can be restricted to those observed in specific care settings (e.g., inpatient, outpatient), and diagnosis codes can be restricted by position (e.g., principal discharge diagnosis, secondary diagnosis).

Each gestational age code is assigned an associated pregnancy duration and priority. If a code is observed within a requester defined window before and after delivery, the duration associated with the code is used to calculate pregnancy duration. If multiple codes are observed, priority is used to determine appropriate duration. If no codes are observed, the requester defines the number of days used to define pregnancy duration.

By default, pregnancy duration is defined using an algorithm developed by the Medication Exposure in Pregnancy Risk Evaluation Program (MEPREP) and other Sentinel and non Sentinel pregnancy-related projects, but it is adaptable as needed.[4,5,6] Identification of gestational age based on the original work in MEPREP was updated to incorporate ICD-10-CM codes.

3. Identifying Medical Products of Interest and Creating Medical Product Episodes

Medical products of interest can be defined using any set of NDCs, procedure and/or diagnosis codes, and laboratory result values found in the SCDM. Procedure and diagnosis codes can be restricted to those observed in specific care settings (e.g., inpatient, outpatient), and diagnosis codes can be restricted by position (e.g., principal discharge diagnosis, secondary diagnosis).

To create episodes, NDCs are processed and dispensing dates are modified using the stockpiling algorithm. After dispensing dates are adjusted using the stockpiling algorithm, exposure episodes are created. Exposure episodes can be defined in one of two ways: a) using outpatient pharmacy dispensing days supplied to create a sequence of continuous exposure, and b) defining a specific number of days after exposure initiation as exposed time. Requesters may optionally restrict evaluation of medical product use to "new use" by requiring a specified period of time (i.e., a "washout" period) before exposure initiation where the individual has no evidence of prior medical product use.

[4] Kawai AT, Li L, Kulldorff M, et al. Mini-Sentinel CBER/PRISM Surveillance Protocol Influenza Vaccines and Pregnancy Outcome. Version 2. https://www.sentinelinitiative.org/sites/default/files/Drugs/Assessments/Mini-Sentinel_PRISM_Influenza-Vaccines-and-Pregnancy-Outcomes-Protocol_0.pdf 2014; Accessed July 7, 2014.
[5] Andrade SE, Davis RL, Cheetham TC, et al. Medication Exposure in Pregnancy Risk Evaluation Program. Matern Child Health J. 2012 Oct;16(7):1349-54. doi: 10.1007/s10995-011-0902-x.
[6] Li Q, Andrade SE, Cooper WO, et al. Validation of an algorithm to estimate gestational age in electronic health plan databases. Pharmacoepidemiol Drug Saf. 2013 May;22(5):524-32. doi:10.1002/pds.3407

a) Timing of Medical Product Exposure During Pregnancy Episodes

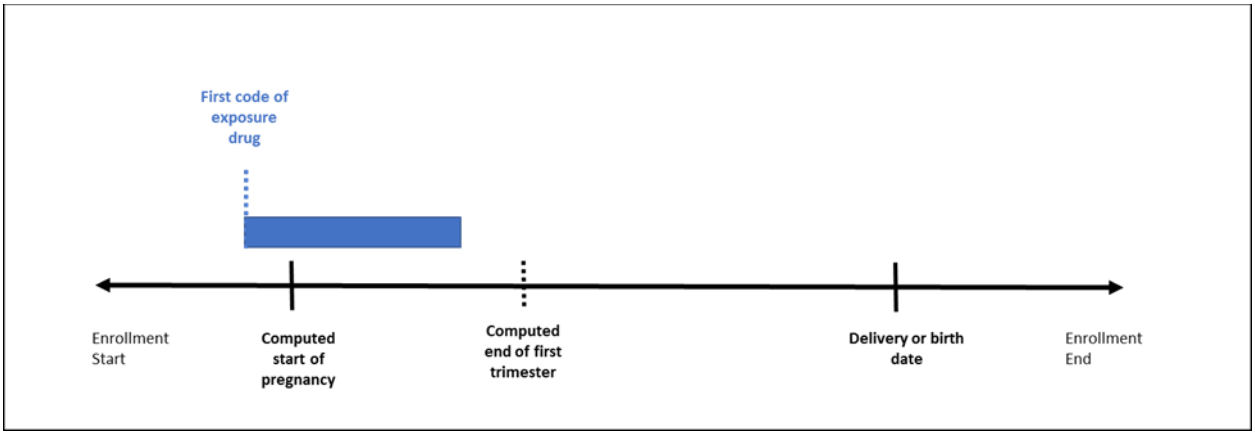
Exposure to the medical products of interest can be assessed during a user-defined pre-pregnancy period before the estimated pregnancy start or during the pregnancy. To facilitate more precise estimates of timing of exposure, pregnancy episodes are divided into trimesters and gestational age in weeks based on calculated time since pregnancy start. The pregnancy period of interest is considered exposed if the dispensing date plus days supply and any stockpiling overlap with the pregnancy period of interest.

b) Medical Product Exposure for Pregnancy Cohorts Processed by PSA Tools

Exposed and comparator pregnancy cohorts are selected based on timing of exposure, as defined by the requester. Date of exposure is considered the first occurrence of exposure during the defined pregnancy period (pre-pregnancy period, trimester of interest, or gestational week of interest). If the exposure episode occurs prior to the defined period of interest, but overlaps the period, the date of exposure will be set to the beginning of the period.

Consider an example where exposed orcomparatorpregnancy cohort selection is based on exposure during the first trimester (Figure 17). In this example, the exposure date would be the computed start of pregnancy.

Figure 17. Exposed Cohort Selection Criteria for Patient Exposed During First Trimester

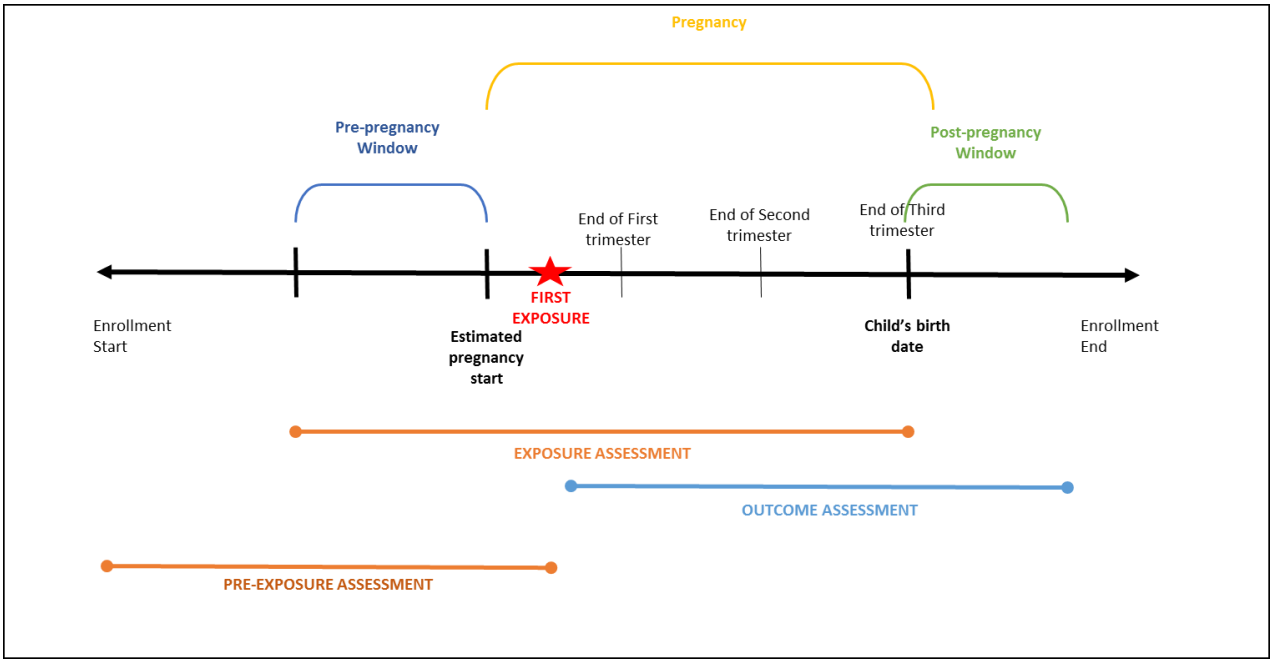


4. Identifying Health Outcome of Interest (HOI)

An HOI can be defined using any combination of NDCs, procedure, and/or diagnosis codes found in the SCDM. Procedure and diagnosis codes can be restricted to those observed in specific care settings (e.g., inpatient, outpatient) and diagnosis codes can be restricted by position (e.g., principal discharge diagnosis, secondary diagnosis). If NDCs are used to define the HOI, they will undergo additional processing with the stockpiling algorithm.

Both maternal HOIs and infant HOIs can be assessed following maternal exposure for pregnancy cohorts designed for further processing with the PSA tool. The CIDA tool queries the SDD and extracts all codes indicative of the HOI for mothers in the pregnancy cohort or for matched infants. Infant HOI assessment is limited to live births included in the SCDM MIL Table. HOIs are assessed during a requester-defined number of days relative to pregnancy start, exposure initiation, and/or delivery date (Figure 18).

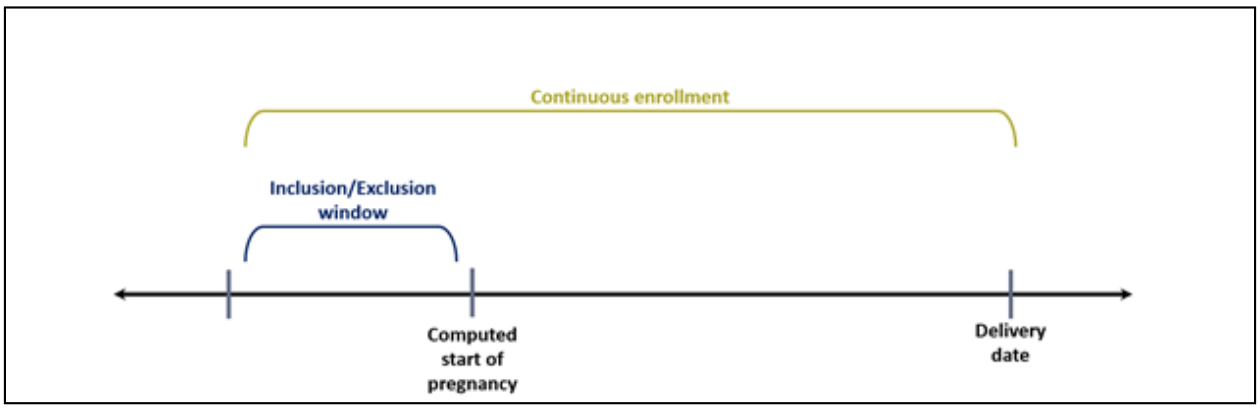
Figure 18. Assessment Windows for a Patient with a First Trimester Exposure



5. Eligible Pregnancy Episodes

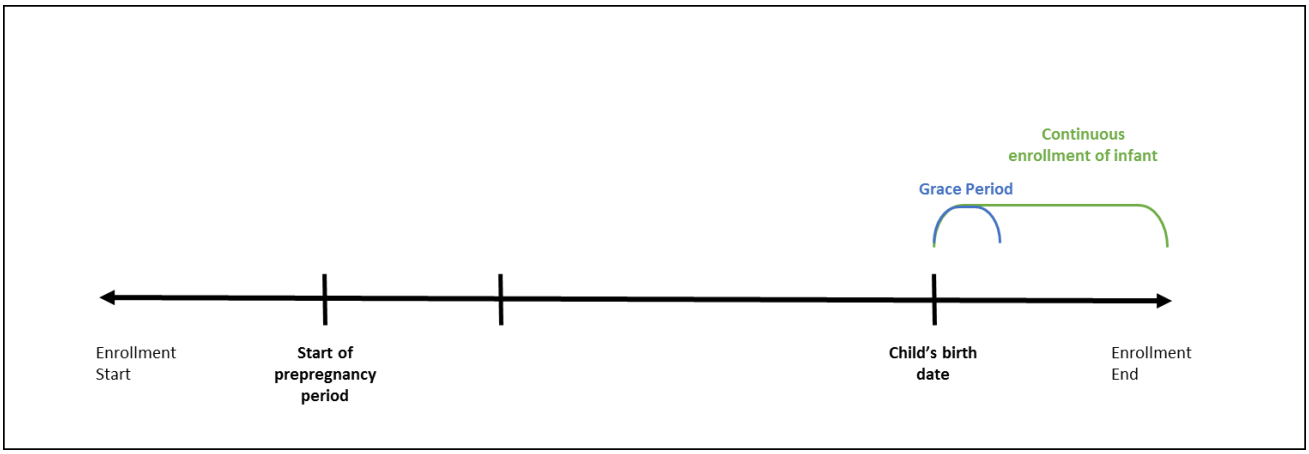
Requester-specified continuous enrollment criteria are assessed relative to delivery date. Requesters must ensure continuous enrollment 1) during the entire pregnancy episode, and 2) during any pre-index date period assessing cohort exclusion criteria (Figure 19). These requirements impose stricter (lengthier) continuous enrollment requirements compared to other cohort identification strategies. In order to evaluate implications of these strict enrollment requirements, requesters will be able to evaluate output for 1) deliveries meeting all cohort inclusion criteria, and 2) deliveries meeting all cohort inclusion criteria *except* specified continuous enrollment criteria.

Figure 19. Ensuring Appropriate Pre-delivery Continuous Enrollment Requirements



For pregnancy cohorts designed for further processing with the PSA tool and with an infant match from the SCDM MIL table, infant enrollment criteria can also be assessed. An infant's enrollment may not begin immediately at birth. Therefore, users are able to specify a post-birth grace period for infants' enrollment start; that is, a user-specified number of days between an infant's birth and enrollment start where an infant is considered to be continuously enrolled (Figure 20).

Figure 20. Ensuring Appropriate Post-delivery Continuous Enrollment Requirements for Infant



6. Number of Valid Pregnancies per Patient

All valid pregnancies per patient that meet requester criteria during the query period are identified.

7. Inclusion/Exclusion Criteria

The CIDA tool allows the application of additional inclusion/exclusion criteria for cohort selection. Inclusion/exclusion criteria can be defined using any combination of NDCs, procedure and/or diagnosis codes, and laboratory result values found in the SCDM. Procedure and diagnosis codes can be restricted to those observed in specific care settings (e.g., inpatient, outpatient) and diagnosis codes can be restricted by position (e.g., principal discharge diagnosis, secondary diagnosis). Each inclusion/exclusion criteria can further be defined by the number of days the code occurs.

Additional inclusion/exclusion criteria are assessed during a requester-defined number of days before, on, or after the pregnancy start date. These criteria determine which pregnancy episodes are included.

a) Inclusion and Exclusion Criteria for Pregnancy Cohorts Processed by PSA Tool

For pregnancy cohorts designed for further processing with the PSA tool, additional inclusion and exclusion criteria can be assessed once the Pregnant Exposed, Pregnant Unexposed, and/or Pregnancy Comparator cohorts are created. Inclusion and exclusion criteria can be assessed based on the mother's and/or infant's claim. The lookup period is determined based on a user-specified anchor date (e.g., pregnancy start date, exposure initiation date (if after pregnancy start), or delivery date of admission). The anchor date and length of lookup period can vary for each inclusion/exclusion criteria.

The mother is required to have continuous enrollment in the coverage type specified during the lookup period prior to the anchor date for assessment of *exclusion* criteria but not inclusion criteria. If this condition is not met, the *pregnancy episode* is excluded from analysis. Continuous enrollment is not

required post-anchor date for exclusion or inclusion criteria. Episodes are not excluded based on infant enrollment because inclusion/exclusion criteria for the infant always looks forward. For infants not enrolled at birth date, criteria are assessed at the start of enrollment (Figure 21 and Figure 22).

Figure 21. Inclusion/Exclusion Assessment Window for Infants with Enrollment at Birth Date

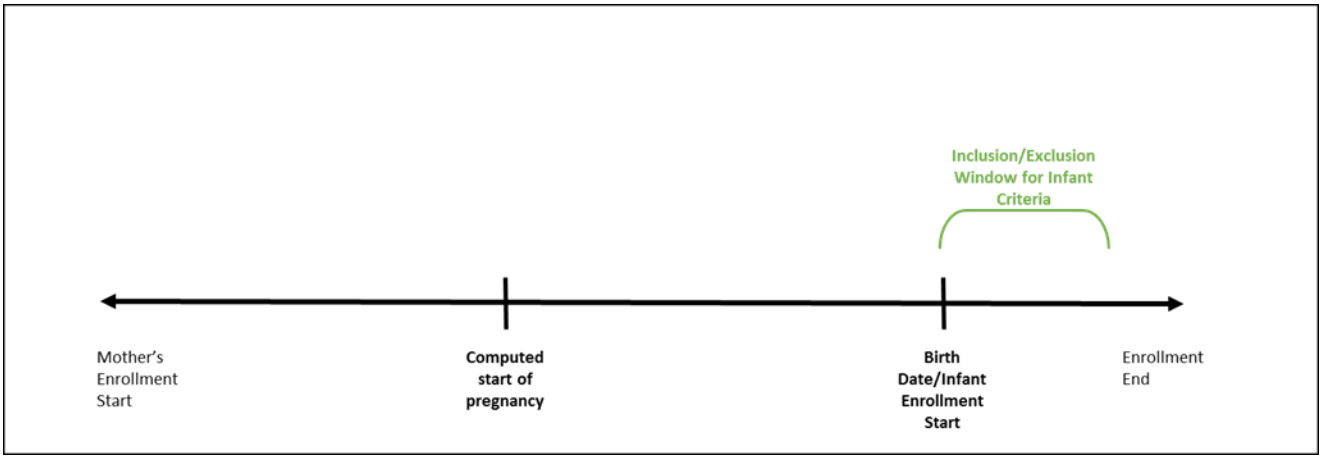
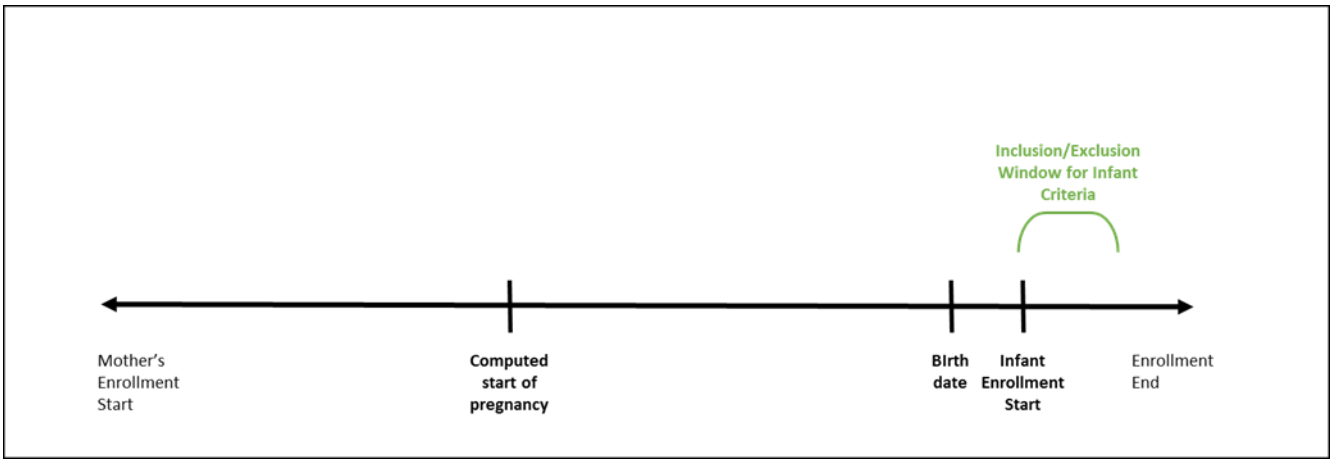


Figure 22. Inclusion/Exclusion Assessment Window for Infants with Enrollment After Birth Date



8. Identifying Non-Pregnant Comparator Episodes

For each identified pregnancy episode, a comparator episode is selected in order to compare medical product use during a defined period for which a woman is likely to not be pregnant. Pregnancy episodes are matched within Data Partner, to the first enrollment episode without a live birth that meets all inclusion/exclusion criteria, is the same (integer) age, and where pregnancy index date overlaps the eligible enrollment span. For example, if a woman has a pregnancy episode from 1/1/2010 to 10/1/2010, she is matched to a woman with a non-pregnant period from 1/1/2010 to 10/1/2010, who meets all inclusion/exclusion criteria and is the same age on 10/1/2010. Women and comparator episodes are allowed to be used multiple times as controls, and women with a pregnancy episode are allowed to contribute a separate comparator episode. This is an optional cohort for Type 4 descriptive analyses and cannot be used to assess HOIs or with the PSA tool.

G. MEDICAL PRODUCT UTILIZATION COHORT IDENTIFICATION STRATEGY

The medical product utilization cohort identification strategy is used to characterize patterns of drug use. This strategy creates episodes of medical product exposure and outputs metrics characterizing patient use and dispensing patterns (e.g., the distribution of days supply per dispensing used to create the treatment episode, the distribution of treatment episode length, reason(s) for treatment episode censoring, and the number of gaps between treatment episodes).

1. Identifying Exposure and Creating Exposure Episodes

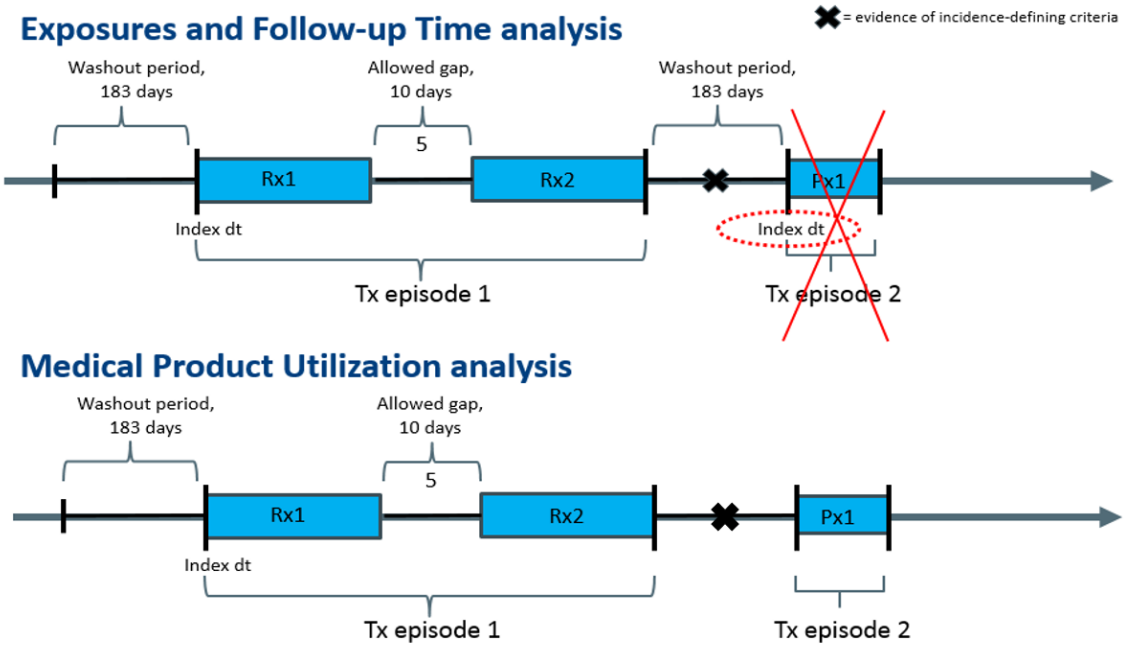
An exposure can be defined using any set of NDCs, procedure and/or diagnosis codes, and laboratory result values found in the SCDM. Procedure and diagnosis codes can be restricted to those observed in specific care settings (e.g., inpatient, outpatient) and diagnosis codes can be restricted by position (e.g., principal discharge diagnosis, secondary diagnosis). For example, exposure to a drug product dispensed in the outpatient setting can be defined as observation of one or more NDCs in the pharmacy dispensing table, whereas exposure to a vaccine can be defined based on observation of specific procedure codes in the procedure table.

The CIDA tool queries the SDD and extracts all codes indicative of exposure during the query period. NDCs are processed and dispensing dates are modified using the stockpiling algorithm.

After dispensing dates are adjusted using the stockpiling algorithm, exposure episodes are created. Exposure episodes can be defined using outpatient pharmacy dispensing days supplied to create a) sequence of continuous exposure, and b) defining a specific number of days after exposure initiation as exposed time. The CIDA tool allows the option to censor exposure episodes based on requester-defined criteria. This could be based on the observation of any NDC, procedure code, diagnosis code, or laboratory result value of interest, or based on medical utilization like the occurrence of a hospitalization. If censoring criteria are observed during an exposure episode, the episode is truncated at the date of the observed criterion.

Unlike the Exposures and Follow-up Time cohort identification strategy, the Medical Product Utilization cohort identification strategy only defines one index date per patient. The first valid treatment episode that meets incidence and pre-index enrollment requirements is identified, and then all subsequent treatment episodes are included for evaluation. Incidence and enrollment criteria is not assessed for episodes subsequent to the first index-defining treatment episode.

Figure 23. Difference Between Exposures and Follow-up and Medical Product Utilization Cohort Identification Strategies



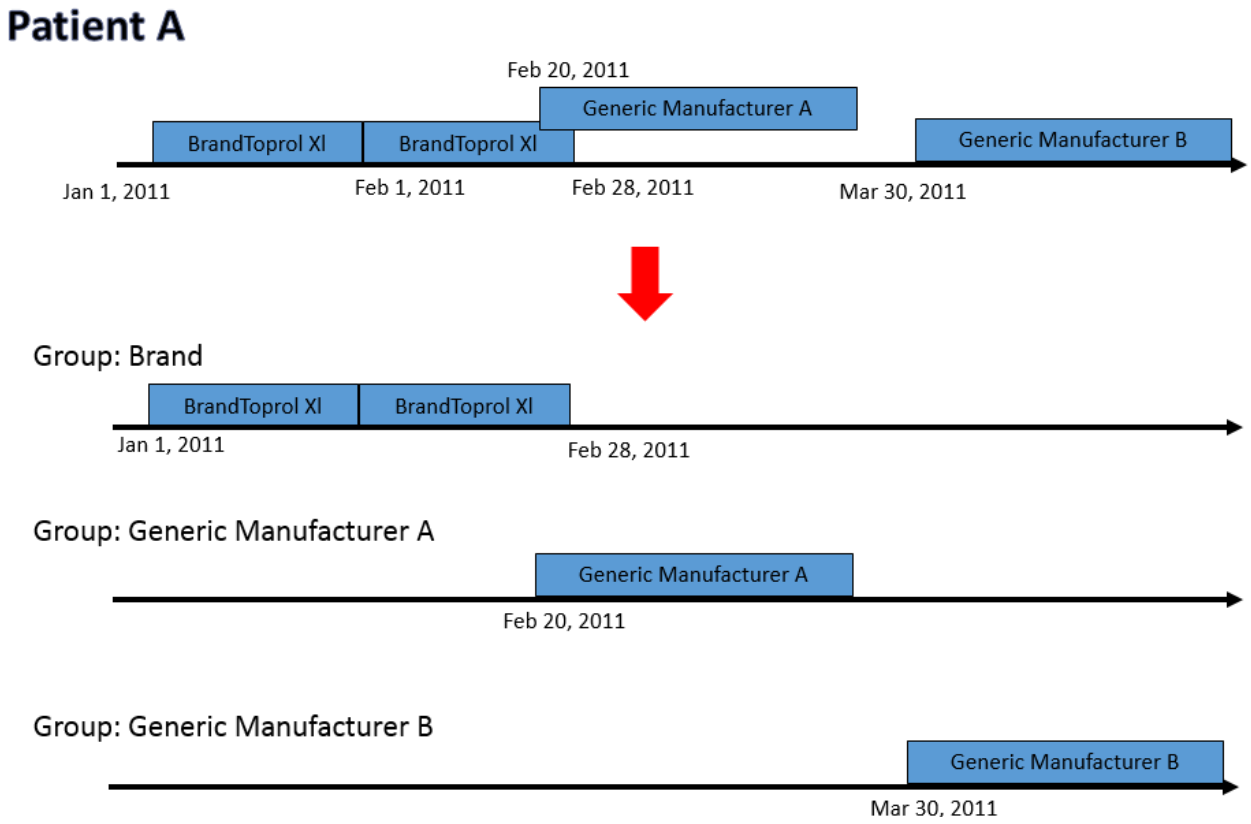
H. MANUFACTURER-LEVEL PRODUCT UTILIZATION AND SWITCHING PATTERNS COHORT IDENTIFICATION STRATEGY

The manufacturer-level product utilization and switching patterns cohort identification strategy is used to characterize patterns of drug use at the manufacturer-level. Product groups are identified by user-defined lists of product codes (e.g., NDCs) grouped together to represent distinct manufacturer-level products (Figure 24). CIDA then identifies treatment episodes constructed from these grouped codes. These treatment episodes serve as the basis for both utilization and switching analyses.

The CIDA tool performs a product utilization analysis for each manufacturer-level exposure group and calculates counts of incident and prevalent users, dispensings, days supplied per dispensing, episode duration, as well as time to uptake. The CIDA tool also performs a product switching analysis that evaluates patient-level switching behavior between manufacturer-level product groups. The tool optionally keeps or discards episodes based on one or more user-specified criteria. Details on both utilization and switching analyses are provided below.

While all exposure group treatment episodes are included in the calculation of utilization metrics, the user defines which of the exposure groups to evaluate for switching.

Figure 24. Example Patient with Episodes in Multiple Product Groups



1. Defining Episode Start and Follow-up

The product utilization and switching patterns cohort identification strategy allows new dates to be specified to build treatment episodes and compute durations, time-to product uptake, and other time-related metrics. At the product code (e.g., NDC) level, the following dates may be specified by the requester: product approval date, product marketing start date, or other product-related date that is requester-defined (e.g., regulatory event date). For each exposure group, the computed marketing start date may be specified by the requester. This date represents the first observed (minimum) dispensing date amongst all valid users within the exposure group within each data partner queried.

2. Product Utilization

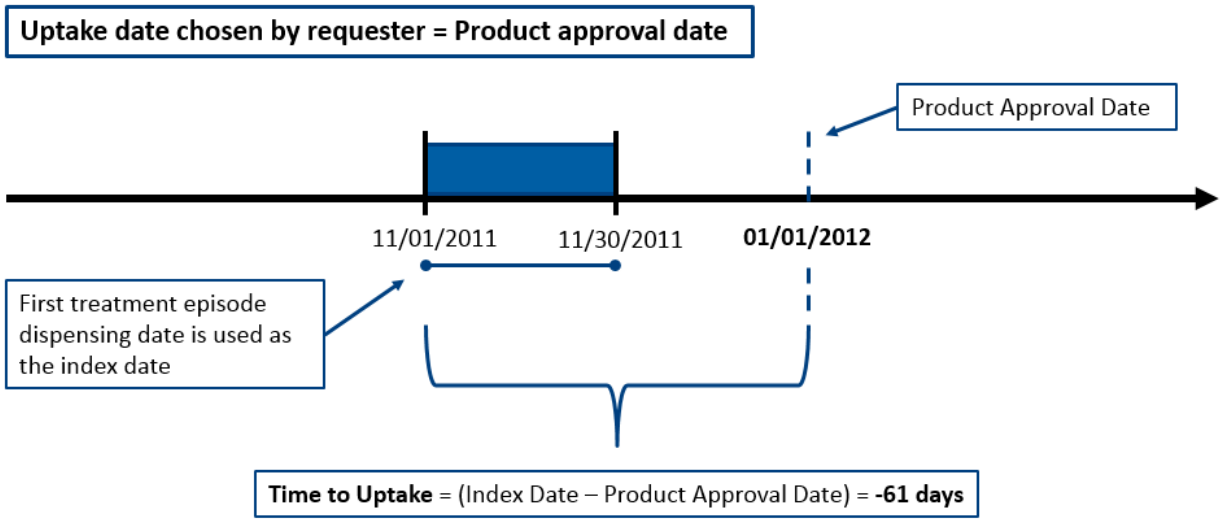
The product utilization portion of this cohort identification strategy computes and reports tables and figures on seven different types of utilization metrics:

1. Number of users (incident and prevalent) at time of index date
2. Number of prevalent users during month of use
3. Number of dispensings at time of index date
4. Number of dispensings at month of dispensing
5. Descriptive statistics for days supplied per dispensing
6. Descriptive statistics for episode duration, including reasons for censoring
7. Descriptive statistics for time to uptake

Time to uptake is computed from either product approval date, product marketing date, other product date (requester-defined date), or computed marketing start date. The computed marketing start date is exposure group-specific, while the product approval and marketing dates are specific to the drug or medical code (e.g., NDC). Time to uptake is calculated from whichever of these dates are specified to a patient's first dispensing date of a product that starts the first treatment episode, as a count of days. If multiple drug or medical codes (e.g., NDCs) with disparate approval or marketing dates are placed together in an exposure group, the tool uses the date for the first product code (e.g., NDC) in a patient's episode.

Time-to-uptake may be computed as a negative value. This occurs when a patient's index date occurs prior to product approval date, product marketing date, or other requester-defined product date (Figure 25). Requesters can decide to either report or discard these negative values in the utilization tables. If negative duration episodes are discarded from time-to-uptake tables, the numbers between that table and the other utilization table may not match.

Figure 25. Negative Time to Uptake



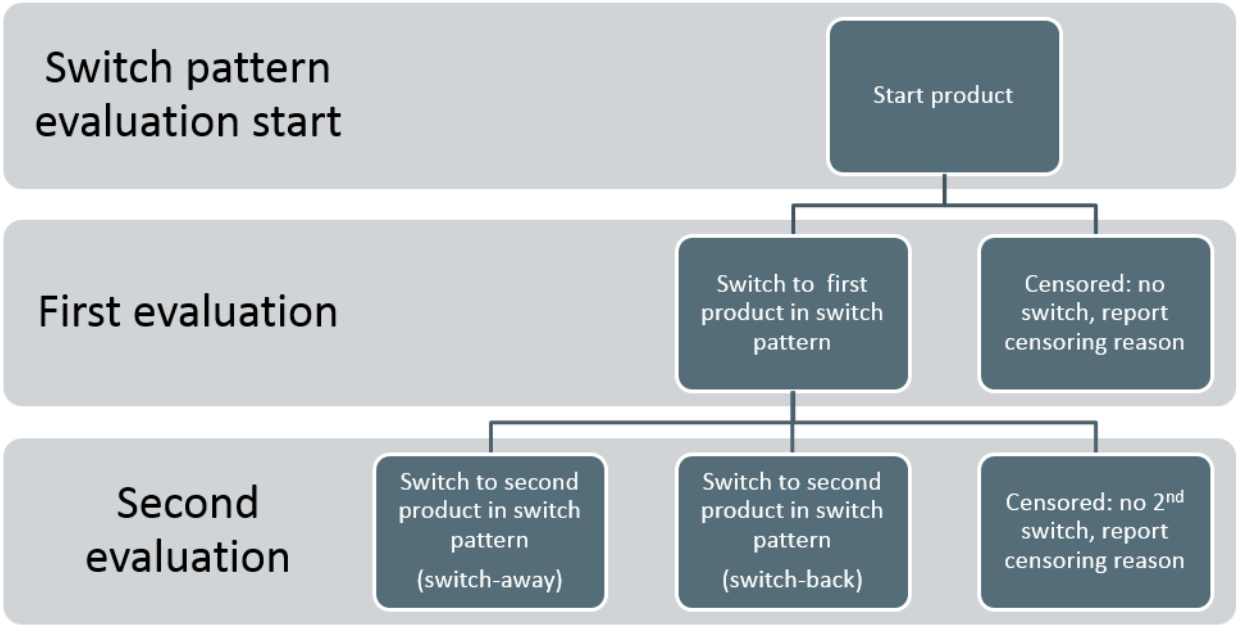
3. Product Switching

Treatment episodes are evaluated for product switching in designated exposure groups. Multiple switch patterns can be specified, but each must be specified separately. The following three types of switching patterns can be evaluated (Figure 26):

1. Switching (e.g., Product A→ Product B)
2. Switch-backs (e.g., Product A → Product B → Product A)
3. Switch-aways (e.g., Product A → Product B → Product C)

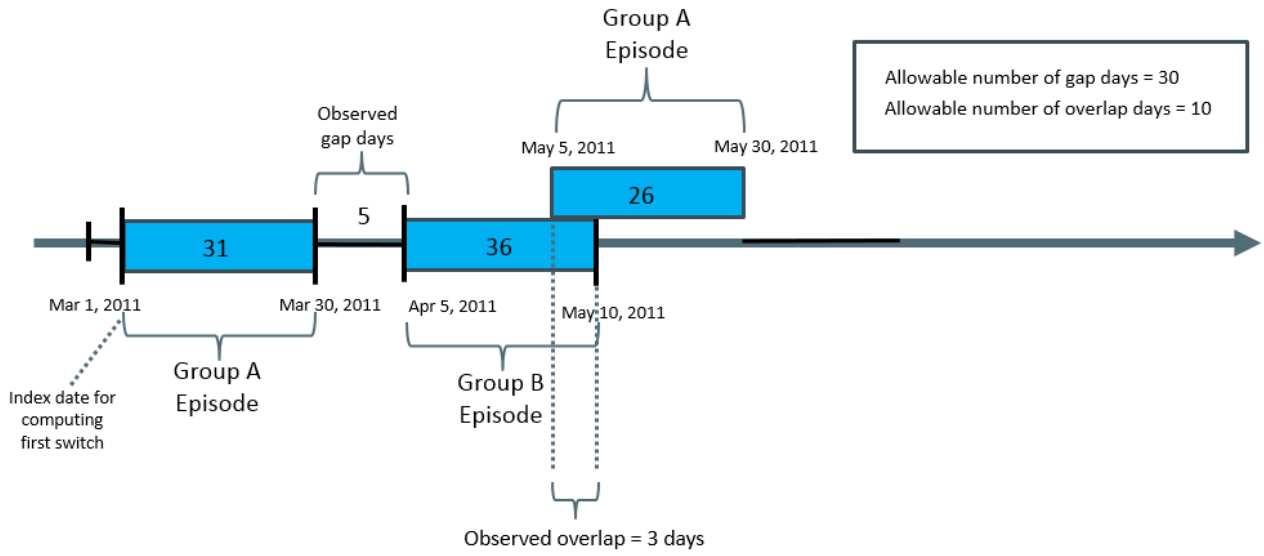
CIDA captures up to two switches per switch pattern (Figure 26). In order for an episode to be identified as a switch from the "start product," that episode's dispensing date should be on or after the dispensing date of the "start product." However, depending on how the date for the "start product" is defined, negative values for time to first switch are possible.

Figure 26. Product Switching Analysis Cohort Identification Strategy



Treatment episodes being evaluated for switching may overlap each other or have gaps between them. The CIDA tool allows the requester to specify overlap and gap thresholds that episodes must meet in order to qualify as a valid switch. Based on the allowable number of gap days and overlap days/percentage selected, Figure 27 ilustrates a valid switch-back (Group A → Group B → Group A).

Figure 27. Switching Episodes in the Product Switching Tool with Overlap and Gap Thresholds



The switching patterns tool allows the user to evaluate switch patterns only during the time in which a product switch was possible (e.g., that a generic product in the switch pattern was actually approved or on the market). It also makes it possible to flexibly compute durations. For example, it allows users the capability to answer these types of questions:

- Amongst *all* brand initiators, how many switched to a generic and how long *from their brand initiation* did that switch occur?
- Amongst brand initiators *who were taking that brand medication at the time a generic was on the market*, how many switched to that generic and *how long from the generic start marketing date* [*or how long from the start of their brand initiation*] did that switch occur?

The following switching metrics are reported by the switching pattern tool:

- Frequency distributions:
  - Time to 1st switch (among those with at least one switch)
  - Time to 2nd switch (among those with a second switch)
  - Patients who switch, by number of months to first-switch (amongst those with at least one switch pattern)
  - Patients who switch, by number of months to second-switch (two-switch pattern only)
  - Censor reason (non-switch)
- Kaplan-Meier curves

The CIDA tool allows the option to keep or discard treatment episodes based on one or more requester-specified criteria, in the following computational order of operations: a) switch inclusion and exclusion criteria, b) switch pattern criteria, and c) the number of valid switch pattern episodes each patient can contribute to the final cohort. These three criteria are described below in more detail.

a) Inclusion/Exclusion Criteria

First, the CIDA tool allows the requester to specify inclusion or exclusion criteria that must be met in order for a switch pattern episode to be retained. Inclusion and exclusion criteria can be defined using any combination of NDCs, procedure and/or diagnosis codes, and laboratory result values found in the SCDM. Procedure and diagnosis codes can be restricted to those observed in specific care settings (e.g., inpatient, outpatient) and diagnosis codes can be restricted by position (e.g., principal discharge diagnosis, secondary diagnosis). These criteria are assessed during a requester-defined number of days before, on, or after the dispensing date of the first treatment episode of a switching pattern.

b) Switch Pattern Criteria

Second, the requester has the option to retain or discard switch pattern episodes based on whether the first treatment episode of the pattern satisfies requester-defined switch pattern cohort inclusion date criteria. As previously described, the cohort inclusion date can be specified as either the product approval date, product marketing date, other requester-defined date, or the computed marketing start date. If a switch inclusion date is specified, then observed patterns of switching will only be counted as such if the date occurs on or before the last day of the first treatment episode of the pattern, inclusive of the gap tolerance value provided.

Appropriate values should be specified for the query start date and switch pattern cohort inclusion date in order to capture product utilization and switching within a calendar time period that is reasonable given particular products' approval or market entry dates. For example, if a user wants to capture use of brand products at a time when generic(s) were approved or on the market, the user could set the query

start date and/or switch pattern cohort inclusion date to the time of generic product approval to dictate at what calendar time point the tool will start looking for product use and potential switching.

The switch pattern cohort inclusion date is used in one of these two ways:

1. As cohort entry only (but the dispensing of the first episode of the pattern will be used as the index date to calculate time to first switch)
2. As both cohort entry AND index date

The algorithm to determine cohort entry is: switch pattern cohort inclusion date ≤ [episode end date + gap tolerance]. Therefore, in order to be included in the switch cohort, the value of the switch pattern cohort inclusion date must occur on or before the end of the first episode inclusive of the gap tolerance value.

Duration to first switch is computed as:

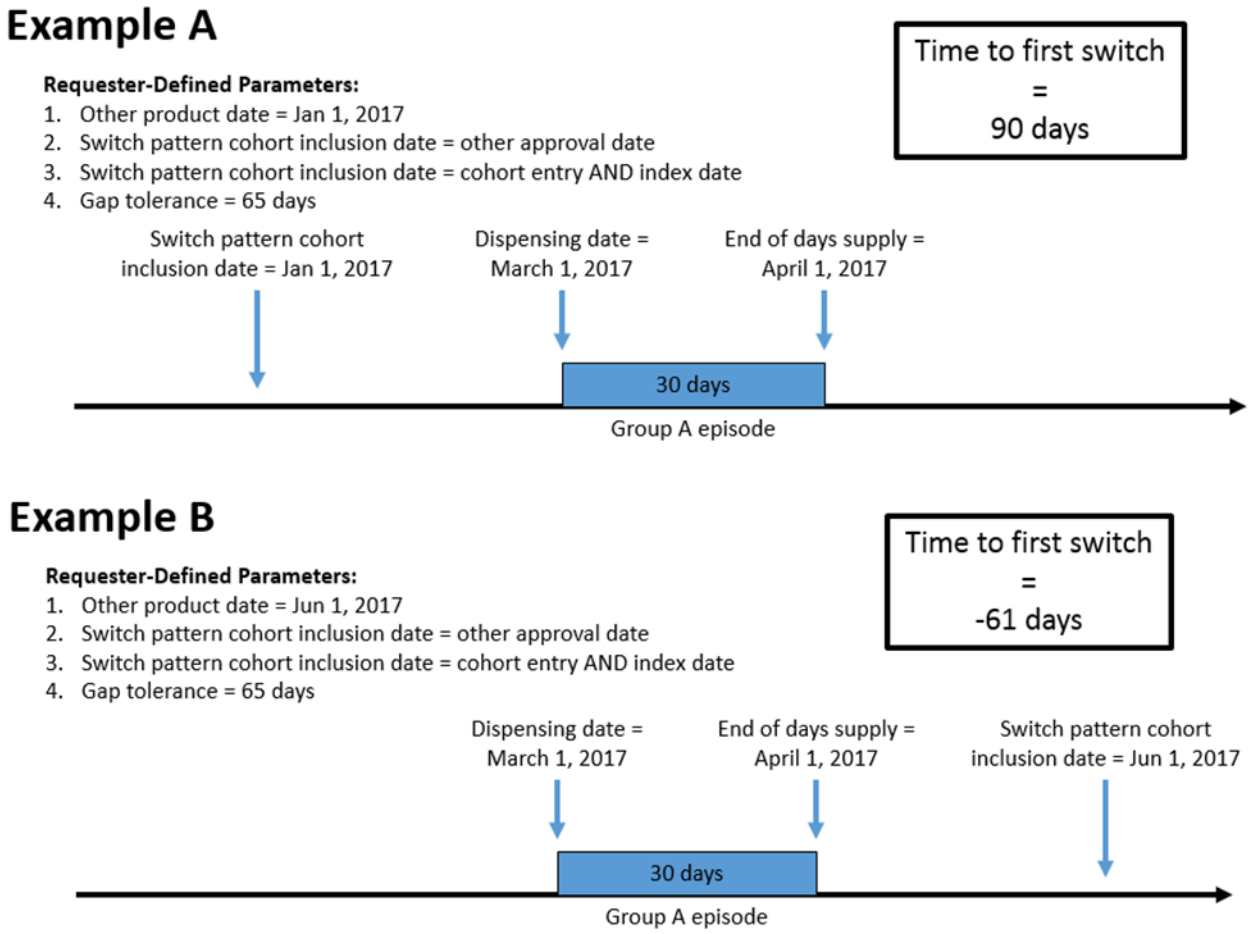
1. Time FROM first episode of pattern dispensing date, or
2. Time FROM switch pattern cohort inclusion date

TO second episode of the pattern dispensing date

Duration to second switch will be computed as time from second episode of the pattern dispensing date to third episode of pattern dispensing date.

When a switch pattern cohort inclusion date is specified as the switch cohort entry date AND as initial switch step index date, for episodes that do not switch, the time to first switch calculation will be from the switch pattern cohort inclusion date to the censoring date for the initial episode. These time-to-first-switch computations could result in a negative duration. Since the tool allows a gap when calculating switch inclusion, it is then possible to get negative time to 1st switch values if the switch pattern cohort inclusion date occurs after the end of the 1st episode, but prior to the end of the gap tolerance. For an example of this, please see Figure 28. In Example A and B below, the switch pattern cohort inclusion date is used as cohort entry and index date. The switch pattern being evaluated is Group A → Group B. In both examples, there is a Group A dispensing observed from 03/01/2017 to 04/01/2017 and no Group B dispensing. In Example A, the requester-defined other product date is 01/01/2017. As a result, the time to first switch is calculated as (04/01/2017) – (01/01/2017) = 90 days. In Example B, the requester-defined other product date is 06/01/2017. As a result, the time to first switch is calculated as (04/01/2017) – (06/01/2017) = -61 days.

Figure 28. Time to First Switch When There Is No Switch Observed



Requester-defined gap and overlap tolerance thresholds can be defined and used to determine whether an observed switch pattern qualifies as a switch or not. The gap tolerance is expressed as a number of allowable days between the two dates in the switch pattern, while the overlap tolerance is expressed as either a number of allowable days or as a percent of the first product group episode duration. Figure 29 shows an example of an observed gap in manufacturer-level exposure group episodes being assessed for switch pattern behavior. To assess for meeting the criteria as a "product-switch" from Group A to Group B, the switch pattern will be assigned a requester-specified value for an allowable gap and an allowable overlap in dispensing. In this example, the allowable gap for the evaluation of switching patterns between Group A and Group B was specified as 10 days. The observed gap was 5 days. This observed dispensing pattern would therefore qualify as a switch, since the observed gap of 5 days is below the requester-specified threshold allowable gap of 10 days. If the person has evidence of a death in the gap period, then the censoring reason would be end of episode and any treatment episode after that observed death would not be counted.

Figure 29. Observed versus Allowable Gap Assessment for Determination of Switch-Pattern Qualification

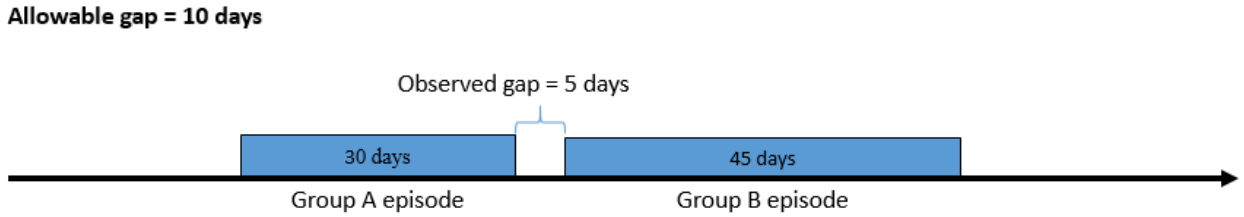


Figure 30 shows an example of an observed overlap in manufactuer-level exposure group episodes being assessed for switch pattern behavior. In this example, the allowable overlap for the evaluation of switching patterns between Group A and Group B was specified as 10 days. The observed overlap was 5 days. This observed dispensing pattern would therefore qualify as a switch, since the observed overlap of 5 days is below the requester-specified threshold allowable overlap of 10 days.

Figure 30. Observed versus Allowable Overlap (Expressed in Days) Assessment for Determination of Switch-Pattern Qualification

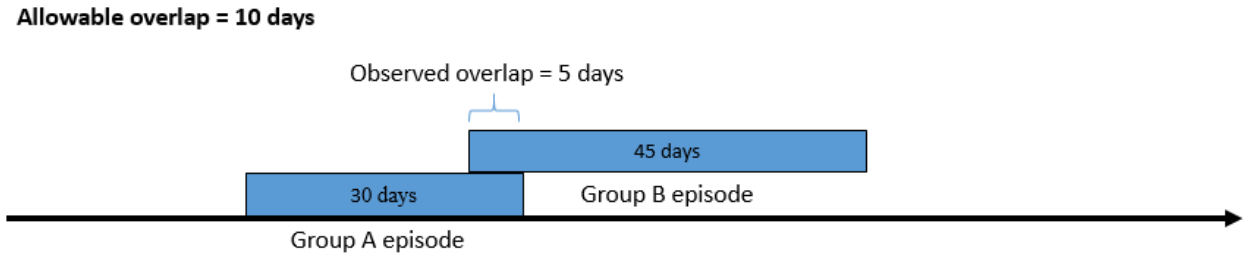
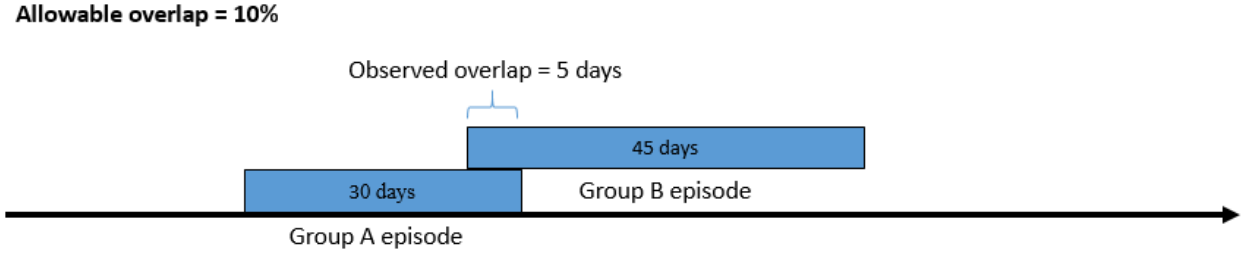


Figure 31 shows an example of an observed overlap in product group episodes being assessed for switch pattern behavior. In this example, the allowable overlap for the evaluation of switching patterns between Group A and Group B was specified as 10 percent of the first product-group episode duration. The observed overlap was 5 days of a 30 day episode, or 16.7 percent. This observed dispensing pattern would therefore not qualify as a switch, since the observed overlap of 16.7 percent is above the requester-specified threshold allowable overlap of 10 percent.

Figure 31. Observed versus Allowable Overlap (Expressed as a Percent) Assessment for Determination of Switch-Pattern Qualification



c) Number of Valid Switch Patterns per Patient

Finally, requesters have the ability to specify the number of valid switch pattern episodes each patient can contribute to the final cohort. Requesters may choose to retain either the first valid switch pattern only or all valid switch pattern episodes during the query period.

A patient may have multiple episodes of product switching. For example:

- Product A user from 01/01/2011 to 01/31/2011, switches to Product B on 02/01/2011 through 2/28/2011, and then switches back to Product A on 03/01/2011. This is an example of a switch-back.
- This same person could have another period of Product A use 2 years later from 01/01/2013 to 01/31/2013, switches to Product B on 02/01/2013 through 02/28/2013, then is censored for end of enrollment on 2/28/2013. This is an example of a switch.

Specifying the capture of only the FIRST switch pattern episode per patient, along with setting the query start date to a calendar time that greatly precedes a time at which products in a designated switch pattern were on the market, will likely result in capture of product use of the first product in the switch pattern at a time when no switch product was available to switch to. For example, setting a query start date date to 01/01/2005 to look for brand and generic product utilization and switching, when a generic product did not become available until 01/01/2011, and looking only for the FIRST product switch episode per patient may result in capturing only brand users who initiated brand on 01/01/2005 and ended use of that brand product on 03/30/2005 for a reason other than a switch.

IV. PROPENSITY SCORE ANALYSIS (PSA) TOOL

A. OVERVIEW

The PSA tool performs effect estimation by comparing exposure propensity-score matched parallel new user cohorts or comparing a new user cohort to a never-exposed cohort. Propensity score estimation and matching are conducted within each Sentinel Data Partner site via distributed programming code; data are returned to the Sentinel Operations Center (SOC), aggregated, and used to calculate effect estimates.

Propensity scores may be estimated using requester-defined covariates and/or empirically identified covariates via a high dimensional propensity score (hdPS) approach. Patients in exposed and comparator cohorts are matched in 1:1 or variable 1:*n* ($n \leq 10$) ratios within a requester-defined caliper.

As the PSA tool functions in a distributed database environment, propensity scores are estimated at each Data Partner site separately. Additionally, as the PSA tool is designed to support sequential analysis, patients are matched in each monitoring period and propensity scores are estimated for each monitoring period.

Note that empirical selection of covariates and fixed and variable ratio propensity score matching functionality was developed based on macros from the Pharmacoepidemiology Toolbox. The Pharmacoepidemiology Toolbox is developed and maintained by the Division of Pharmacoepidemiology and Pharmacoeconomics in the Department of Medicine at Brigham and Women's Hospital (http://www.drugepi.org/dope-downloads/).

B. CIDA TOOL REQUIREMENTS AND OUTPUT PRE-PROCESSING

Both exposure and comparator (active or never-exposed) cohorts for analysis are identified by the CIDA tool. These cohorts are created within the CIDA tool and are further processed by the PSA tool. If the PSA tool is executed, requesters must specify the following in the CIDA tool:

- Use the exposures and follow-up time cohort identification strategy
    - Can define exposed time using episodes created by outpatient pharmacy dispensings days supplied or requester-defined number of days after exposure initiation
- Allow only one exposure period per patient for exposure and comparator cohorts
- For analyses assessing pregnancy cohorts, allow only one pregnancy episode for exposure and comparator/unexposed cohorts. Also, allow only one exposure episode per pregnancy episode.
- Exclude same day initiators of the exposure and comparator product of interest
    - This requires specifying an exclusion criterion on the index date for both cohorts
- Specify covariate list and evaluation window(s) for estimating propensity score

The PSA tool will also perform the following pre-processing steps on the CIDA tool output:

- If the same patient is identified in the exposure and comparator cohort by the CIDA tool, the patient is only retained in the cohort of earliest exposure
- If in the CIDA tool output a patient initiates treatment with the exposure and comparator product on the same day, the patient is discarded from analysis
    - This should be specified by the CIDA tool, but the PSA tool will automatically check to see if the exposure and comparators of interest are initiated on the same day and exclude the patient from analysis.

C. PROPENSITY SCORE ESTIMATION

A propensity score is estimated for every patient using logistic regression with exposure as the dependent variable and potential confounders as independent variables. Each patient's predicted probability of exposure (given their observed covariates) is their propensity score.

Requesters may specify covariates for inclusion in the propensity score estimation model and/or empirically identify covariates via a hdPS approach.

1. Requester-defined Covariates

Requesters may define a list of binary covariates to include in the propensity score estimation model (e.g., history of diabetes, heart failure, etc.). If a clinical concept can be defined using any combination of NDCs, diagnosis and/or procedure codes, it can be included as a binary covariate in the propensity score estimation model.

Requesters may also choose to add any of the following categorical, continuous, or count metrics to the propensity score estimation model:

1. Age (continuous)
2. Sex
3. Race
4. Hispanic
5. Time period (i.e., monitoring period for sequential analyses)
6. Year of exposure

7. Comorbidity score
8. Medical utilization – number of inpatient stays
9. Medical utilization – number of institutional stays
10. Medical utilization – number of emergency department visits
11. Medical utilization – number of outpatient visits
12. Health care utilization – number of other ambulatory encounters (e.g., telemedicine, email consults)
13. Drug utilization – number of dispensings
14. Drug utilization – number of unique generics dispensed
15. Drug Utilization – number of unique drug classes dispensed

2. Empirically Identified Covariates

Requesters may also empirically-identify covariates for inclusion in the propensity score estimation model via a hdPS approach. The hdPS option allows for selection of empirically identified covariates based on the potential for confounding the exposure/outcome association under investigation. There are several requester options available for hdPS estimation:

- Number of covariates to consider for inclusion in the hdPS model for each *data dimension*. There are seven data dimensions considered: 1) drug class; 2) ICD-9-CM diagnosis codes; 3) ICD-10-CM diagnosis codes; 4) ICD-9-CM procedure codes; 5) ICD-10-CM procedure codes; 6) Healthcare Common Procedure Coding System (HCPCS) codes; and 7) Current Procedural Terminology (CPT) codes.
- Maximum number of covariates to include in the hdPS model. Note that this requester-defined maximum is overwritten by the smallest number of new users in either the exposure or comparator cohort.
- Method for ranking/prioritizing covariates for inclusion in the hdPS model. Options include:
    - Exposure association ranking (default): yields a variable list in which the variables are selected as ranked by the strength of the relationship between covariate and exposure. This is most suitable for cases where there are fewer than 150 exposed outcomes.
    - Outcome association ranking: yields a variable list in which the variables are selected as ranked by the strength of the relationship between confounder and the outcome. This is most suitable for disease risk scores.
    - Bias ranking: yields a variable list in which variables are selected as ranked by the Bross bias formula.[7]

D. PROPENSITY SCORE MATCHING

In propensity score matched analyses, patients in an exposed group are matched to patients in a comparator group with similar propensity scores. The comparator group can be an active-comparator or a never-exposed comparator. Patients in exposed and comparator cohorts may be matched in 1:1 or variable 1:*n* (*n*≤10) ratios within a requester-defined caliper. A caliper specifies a maximum matching distance and is specified on the natural scale (e.g., 0.01, 0.025, 0.05) of the propensity score. A caliper can be any number between 0 and 1 (value can be specified to the eighth decimal place). The objective

[7] Bross IDJ. Spurious effects from an extraneous variable. J Chronic Dis. 1966 Jun;19(6):637-47.

of the matching algorithm is to minimize the global absolute difference between matched pairs across all matches. Section VII.C provides a detailed description of the matching algorithm and examples.

E. EFFECT ESTIMATION

Data returned to the SOC by participating Data Partners are aggregated and analyzed to produce effect estimates and p-values. How effect estimation is performed is dependent on the level of data requested from participating Data Partner sites.

1. Individual-level Data Return

The program may return individual-level, de-identified datasets to SOC for exposed and active-comparator cohorts. While the datasets contain a single row per patient for each specified analysis, patient identifiers such as PatID are not included in the output. Individual-level datasets are returned to the SOC, aggregated, and used to calculate effect estimates via Cox (proportional hazards) regression. Based on requester needs, the program can calculate an effect estimate for the base population (i.e., all patients eligible to be matched) adjusted by Data Partner, and two effect estimates for the matched population: a conditional and unconditional analysis.

- Unmatched analysis: a Cox model, stratified by Data Partner site, is run on the eligible population.
- Matched analysis (conditional): a Cox model, stratified by Data Partner site and matched set, is run on the matched population. This can be done for both the both 1:1 and 1:*n* matched cohorts.
- Matched analysis (unconditional): a Cox model, stratified by Data Partner site only, is run on the matched population. This can be done for the 1:1 matched cohort only.

2. Risk-set-level Data Return

An alternative to the patient-level data return approach is risk-set level data return. In this approach, the PSA tool will produce de-identified, risk-set level datasets instead of or in addition to individual-level output. Whereas each observation in the patient-level datasets represents one patient in the cohort, each observation in the risk set dataset represents one event. Risk sets are created at the Data Partner site, returned to the SOC, aggregated, and used to calculate effect estimates via case-centered logistic regression.[8]

Risk sets are created to support unmatched analyses, conditional matched analyses for the 1:1 and 1:*n* matched populations, and unconditional matched analyses for the 1:1 matched population.

[8] Fireman B, Lee J, Lewis N, et al. Influenza vaccination and mortality: differentiating vaccine effects from bias. Am J Epidemiol 2009;170(5):650-656.

a) Creating Risk-set Level Datasets for Unmatched Analyses

For unmatched analyses, risk sets are created within the entire eligible population. To convert individual-level datasets to risk set-level datasets, the following steps are taken:

1. The individual-level data is sorted by follow-up time. Follow-up time is the number of days a patient is followed post-exposure until they are either censored or have an event.
2. Each patient with an event is selected to contribute to a risk set.
3. The probability of exposure in the risk set is calculated using each patient with follow-up time greater than or equal to the follow-up time of the case.

Table 1 includes example output for an individual-level, de-identified dataset; Table 2 includes an example of a translating the individual-level dataset to a risk-set level dataset.

Table 1. Example individual-level output, unmatched analyses

Exposure (1=treated)	Follow-Up Time (days)	Study Class	Event (1=event)	*Add to Risk Level Data Set?*
1	**39**	**exposure**	**1**	***Yes***
1	**39**	**exposure**	**1**	***Yes***
0	**71**	**comparator**	**1**	***Yes***
0	72	comparator	0	*No*

Table 2. Example translation to risk set level output, unmatched analyses

Risk Set ID (event indicator)	Case Exposure (1=treated)	Exposure Probability (in risk set)	Follow-Up Time (days)	Number Exposed In Risk Set	Number In Risk Set
1	**1**	**0.5**	**39**	**2**	***4***
2	**1**	**0.5**	**39**	***2***	***4***
3	**0**	**0**	**71**	***0***	***2***
	0		72		

In Table 1 and Table 2, four patients in an unmatched analysis are converted to three risk sets.

There are three events, all of which are output to the risk set dataset with their corresponding follow-up time and exposure status. The exposure probability for the first risk set (Risk Set ID =1) is calculated by taking the number of exposed individuals in the eligible population with follow-up greater than or equal to the case (2) divided by the number of patients left in the risk set (4). The same process is repeated for the second event (Risk Set ID = 2) and third event (Risk Set ID = 3).

b) Creating Risk-set Level Datasets for Matched Analyses (Conditional)

For conditional analyses, risk sets are created within each matched set in the analysis. To convert individual-level datasets to risk set-level datasets, the following steps are taken:

1. The individual-level data is sorted by follow-up time within each matched set. Follow-up time is the number of days a patient is followed post-exposure until they are either censored or have an event.
2. Within each matched set, each patient with an event is selected to contribute to a risk set.
3. The probability of exposure in the risk set is calculated using each patient in the matched set with follow-up time greater than or equal to the follow-up time of the case.

Table 3 includes example output for an individual-level, de-identified dataset; Table 4 includes an example of a translating the individual-level dataset to a risk-set level dataset.

Table 3. Example individual-level output, conditional analysis

Match ID	Exposure (1=treated)	Follow-Up Time (days)	Study Class	Event (1=event)	*Add to Risk Level Data Set?*
1	1	39	exposure	0	*No*
1	**0**	**145**	**comparator**	**1**	***Yes***
1	**0**	**191**	**comparator**	**1**	***Yes***
2	1	55	exposure	0	*No*
2	0	61	comparator	0	*No*
2	0	99	comparator	0	*No*
3	**1**	**39**	**exposure**	**1**	***Yes***
3	**0**	**39**	**comparator**	**1**	***Yes***
3	0	72	comparator	0	*No*
4	**0**	**39**	**comparator**	**1**	***Yes***
4	0	39	comparator	0	*No*
4	**0**	**71**	**comparator**	**1**	***Yes***
4	**1**	**79**	**exposure**	**1**	***Yes***
4	**0**	**84**	**comparator**	**1**	***Yes***

Table 4. Example translation to risk set level output, conditional analysis

Match ID	Risk Set ID (event indicator)	Case Exposure (1=treated)	Exposure Probability (in risk set)	Follow-Up Time (days)	Number Exposed In Risk Set	Number In Risk Set
1		1		39		
1	**1**	**0**	**0**	**145**	***0***	***2***
1	**2**	**0**	**0**	**191**	***0***	***1***
2		1		55		
2		0		61		
2		0		99		
3	**3**	**1**	**0.33**	**39**	***1***	***3***
3	**4**	**0**	**0.33**	**39**	***1***	***3***
3		0		72		
4	**5**	**0**	**0.2**	**39**	***1***	***5***
4		0		39		
4	**6**	**0**	**0.33**	**71**	***1***	***3***
4	**7**	**1**	**0.5**	**79**	***1***	***2***
4	**8**	**0**	**0**	**84**	***0***	***1***

In Table 3 and Table 4, fifteen patients in a matched analysis are converted to 8 risk sets.

In the first matched set (Match ID = 1), there are two events; both are output to the risk set dataset with their corresponding follow-up time and exposure status. The exposure probability for the first risk set (Risk Set ID =1) is calculated by taking the number of exposed individuals in the matched set with follow-up greater than or equal to the case (0) divided by the number of patients left in the risk set (2). The same process is repeated for the second event (Risk Set ID = 2). For the second matched set (Match ID= 2), there are no events, so no risk sets are created. For the third matched set (Match ID= 3), two patients have an event, both with a follow-up time of 39 days. For both patients, there are three individuals with follow-up time equal to or greater than 39, of which one is exposed. Thus, for each risk set, the exposure probability is 0.33.

For the fourth matched set (Match ID= 4), there are five patients, of which four have events. The first event occurs at 39 days of follow-up. The exposure probability is calculated as 0.2 (one individual in the exposed group / five individuals in the matched set). A risk set is created for each of the other events.

Note that "uninformative" risk sets (i.e., risk sets with only exposed or only comparator patients) will be output with an exposure probability of 0 or 1. While included on the risk-set level dataset, these risk-sets will be discarded in a case-centered logistic regression.

c) Creating Risk-set Level Datasets for Matched Analyses (Unconditional)

For unconditional analyses, risk sets are created within the entire matched population. To convert individual-level datasets to risk set-level datasets, the following steps are taken:

1. The individual-level data for the matched population is sorted by follow-up time. Follow-up time is the number of days a patient is followed post-exposure until they are either censored or have an event.
2. Each patient with an event is selected to contribute to a risk set.
3. The probability of exposure in the risk set is calculated using each patient in the matched population with follow-up time greater than or equal to the follow-up time of the case.

The risk-set level creation process is similar to the one used for the unmatched analyses, except the base population is all matched patients instead of the entire eligible population.

3. Effect Estimation Summary

To summarize, the effect estimation method depends on the data returned to SOC and the proposed analysis (Table 5).

Table 5. Effect Estimation Process Summary

Analysis Type	Effect Estimation	
	Individual-level data return	*Risk-set-level data return*
Unmatched	Population: all exposed and comparator patients Method: Cox proportional hazards model stratified by DP	Population: all exposed and comparator patients Method: risk-sets created within the entire eligible population; case-centered logistic regression.
Matched, Conditional	Population: all matched patients Method: Cox proportional hazards model stratified by DP and matched set	Population: all matched patients Method: risk-sets created within each matched set; case-centered logistic regression.
Matched, Unconditional *(1:1 matching only)*	Population: all matched patients Method: Cox proportional hazards model stratified by DP	Population: all matched patients Method: risk-sets created within the matched population; case-centered logistic regression.

4. A Note on P-value Computation

Due to computational differences between the procedures used to calculate p-values, there may be slight differences in p-values between risk-set and individual-level data analyses. In theory, the procedure used for the Cox regression (PROC PHREG) and the procedure used for the case-centered logistic regression (PROC GENMOD) maximize the same likelihood function; however, since they use slightly different numeric algorithms, there may be minor differences in numeric results. Requesters comparing individual-level and risk-set-level results for the same analyses should expect minor p-value differences.

5. Subgroup Analyses

Subgroup analyses may be conducted using any requester-defined covariates. Subgroup analyses may be performed in the eligible population and the matched population. The method for effect estimation varies depending on the level of data returned to SOC:

- **Unmatched analysis:** all eligible patients are included. The exposure and comparator cohorts are subset based on the values of the subgroup variable and effect estimation is performed.
    - Individual level data return: a Cox model, stratified by Data Partner site, is run on each level of the subgroup variable
    - Risk-set level data return: risk-sets are created within levels of the subgroup variable (and by Data Partner) and estimation is performed using case-centered logistic regression
- **Matched analysis (conditional):** all matched patients are included (i.e., for 1:1 matched analyses, only exposed and comparator patients selected in the 1:1 match are considered; for 1:*n* matched analyses, only exposed and comparator patients selected in the 1:*n* match are considered). The matched cohort is subset based on the values of the subgroup variable and re-matched within values of the subgroup variable (using specified matching ratio and caliper).
    - Individual level data return: a Cox model, stratified by Data Partner site and matched set, is run on each level of the subgroup variable. This can be done for both the 1:1 and 1:*n* matched cohorts.
    - Risk-set level data return: risk-sets are created within levels of the subgroup variable (and by Data Partner) and within matched set; estimation is performed using case-centered logistic regression. This can be done for both the 1:1 and 1:*n* matched cohorts.
- **Matched analysis (unconditional)**: all matched patients are included (i.e., for 1:1 matched analyses, only exposed and comparator patients selected in the 1:1 match are considered). The matched cohort is subset based on the values of the subgroup variable and re-matched within values of the subgroup variable (using specified matching ratio and caliper).
    - Individual level data return: a Cox model, stratified by Data Partner site only, is run on each level of the subgroup variable. This can be done for the 1:1 matched cohort only.
    - Risk-set level data return: risk-sets are created within levels of the subgroup variable (and by Data Partner); estimation is performed using case-centered logistic regression. This can be done for the 1:1 matched cohort only.

F. PROPENSITY SCORE PERCENTILE STRATIFICATION

The PSA tool can also stratify propensity scores based on requester-defined percentiles. Note that all patients identified in exposure and comparator cohorts are used in the analysis (i.e., eligible patients that were not included in the matched analyses are included in this analysis).

The method for effect estimation varies depending on the level of data returned to SOC:

Individual-level data return: patients are subset based on requester-defined subsets of propensity score percentile. A Cox model, stratified by Data Partner site and propensity score percentile, is run on each percentile subset and on the overall population.

Risk set-level data return: risk-sets are created within each percentile subset (and by Data Partner); estimation is performed using case-centered logistic regression run on each percentile subset and on the overall population.

G. OUTPUT

The PSA tool automatically generates tables of patient characteristics, stratified by exposure group, for the unmatched cohort and matched cohort, separately for each Data Partner and each monitoring period. Tables include measures of covariate balance, including absolute and standardized differences, which indicate balance in specific variables, and the Mahalanobis distance[9,10], which provides a measure of balance across all variables while accounting for their correlation. The tables also include the number of patients in each exposure group, the number matched from each group (where appropriate), the number that experienced HOIs, and the mean person-time of follow-up.

The program also automatically generates histograms depicting the propensity score distributions for each exposure group, separately for each Data Partner and each monitoring period, and before and after matching. Figures include c-statistics for each propensity score model.

1. Kaplan-Meier Plots

The PSA tool will automatically produce de-identified, aggregated data sets summarizing follow-up days, number of exposed and unexposed on each day, and the number of events on each day for exposed and unexposed groups. This dataset is returned to the SOC, aggregated, and used to produce Kaplan-Meier plots.

V. MULTIPLE FACTOR MATCHING (MFM) TOOL

A. OVERVIEW

The MFM tool performs effect estimation by comparing exposure exact matched parallel new user cohorts or comparing a new user cohort to a never-exposed cohort. Multiple factor matching is conducted within each Sentinel Data Partner site via distributed programming code; data are returned to the Sentinel Operations Center (SOC), aggregated, and used to calculate effect estimates.

The MFM tool will find an exact match between patients in exposed and comparator cohorts based on any requester-defined combionation of sex, age group, and/or year of index date. Patients in exposed and comparator cohorts are matched in 1:1 or variable 1:*n* ($n \leq 10$) ratios.

B. CIDA TOOL REQUIREMENTS AND OUTPUT PRE-PROCESSING

Both the exposed and comparator cohorts for analysis are identified by the CIDA tool. These cohorts are created within the CIDA tool and are further processed by the MFM tool. If the MFM tool is executed, requesters must specify the following in the CIDA tool:

- Use the exposures and follow-up time cohort identification strategy
    - Can define exposed time using episodes created by outpatient pharmacy dispensings days supplied or requester-defined number of days after exposure initiation

[9] Mahalanobis PC. On the generalized distance in statistics. Proc Natl Inst Sci (India). 1936; 12: 49-55
[10] Rassen JA, Brookhart MA, Glynn RJ, Mittleman MA, Schneeweiss S. Instrumental variables II: in 25 variations, the physician prescribing preference generally was strong and reduced imbalance. J Clin Epidemiol. 2009; 62: 1233-41.

- Allow only one exposure period per patient for exposure and comparator cohorts
- Exclude same day initiators of the exposure and comparator product of interest
    - This requires specifying an exclusion criterion on the index date for both cohorts
    - For example, if exposure is Drug A and the comparator is Drug B and a member is dispensed both Drug A and Drug B on the same day, CIDA will exclude this member from the analysis

The MFM tool will also perform the following pre-processing steps on the CIDA tool output:

- If the same patient is identified in the exposure and comparator cohort by the CIDA tool, the patient is only retained in the cohort of earliest exposure
- If in the CIDA tool output a patient initiates treatment with the exposure and comparator product on the same day, the patient is discarded from analysis
    - This should be specified by the CIDA tool, but the MFM tool will automatically check to see if the exposure and comparators of interest are initiated on the same day and exclude the patient from analysis.

C. EFFECT ESTIMATION

Data returned to the SOC by participating Data Partners are aggregated and analyzed to produce effect estimates and p-values. How effect estimation is performed is dependent on the level of data requested from participating Data Partner sites.

1. Individual-level Data Return

The program may return individual-level, de-identified datasets to SOC for exposed and active-comparator cohorts. These individual-level datasets are not available for never-exposed cohorts. While the datasets contain a single row per patient for each specified analysis, patient identifiers such as PatID are not included in the output. Individual-level datasets are returned to the SOC, aggregated, and used to calculate effect estimates via Cox (proportional hazards) regression. Based on requester needs, the program can calculate an effect estimate for the base population (i.e., all patients eligible to be matched) adjusted by Data Partner, and two effect estimates for the matched population: a conditional and unconditional analysis.

- Unmatched analysis: a Cox model, stratified by Data Partner site, is run on the eligible population.
- Matched analysis (conditional): a Cox model, stratified by Data Partner site and matched set, is run on the matched population. This can be done for both the both 1:1 and 1:*n* matched cohorts.
- Matched analysis (unconditional): a Cox model, stratified by Data Partner site only, is run on the matched population. This can be done for the 1:1 matched cohort only.

2. Risk-set-level Data Return

An alternative to the patient-level data return approach is risk-set level data return. In this approach, the MFM tool will produce de-identified, risk-set level datasets instead of or in addition to individual-level output. Whereas each observation in the patient-level datasets represents one patient in the cohort, each observation in the risk set dataset represents one event. Risk sets are created at the Data Partner site, returned to the SOC, aggregated, and used to calculate effect estimates via case-centered logistic regression.[11]

Risk sets are created to support unmatched analyses, conditional matched analyses for the 1:1 and 1:*n* matched populations, and unconditional matched analyses for the 1:1 matched population.

a) Creating Risk-set Level Datasets for Unmatched Analyses

For unmatched analyses, risk sets are created within the entire eligible population. To convert individual-level datasets to risk set-level datasets, the following steps are taken:

1. The individual-level data is sorted by follow-up time. Follow-up time is the number of days a patient is followed post-exposure until they are either censored or have an event.
2. Each patient with an event is selected to contribute to a risk set.
3. The probability of exposure in the risk set is calculated using each patient with follow-up time greater than or equal to the follow-up time of the case.

Table 6 includes example output for an individual-level, de-identified dataset; Table 7 includes an example of a translating the individual-level dataset to a risk-set level dataset.

Table 6. Example individual-level output, unmatched analyses

Exposure (1=treated)	Follow-Up Time (days)	Study Class	Event (1=event)	*Add to Risk Level Data Set?*
1	**39**	**exposure**	**1**	***Yes***
1	**39**	**exposure**	**1**	***Yes***
0	**71**	**comparator**	**1**	***Yes***
0	72	comparator	0	*No*

Table 7. Example translation to risk set level output, unmatched analyses

Risk Set ID (event indicator)	Case Exposure (1=treated)	Exposure Probability (in risk set)	Follow-Up Time (days)	Number Exposed In Risk Set	Number In Risk Set
1	**1**	**0.5**	**39**	**2**	***4***
2	**1**	**0.5**	**39**	***2***	***4***
3	**0**	**0**	**71**	***0***	***2***
	0		72		

In Table 6 and Table 7, four patients in an unmatched analysis are converted to three risk sets.

[11] Fireman B, Lee J, Lewis N, et al. Influenza vaccination and mortality: differentiating vaccine effects from bias. Am J Epidemiol 2009;170(5):650-656.

There are three events, all of which are output to the risk set dataset with their corresponding follow-up time and exposure status. The exposure probability for the first risk set (Risk Set ID =1) is calculated by taking the number of exposed individuals in the eligible population with follow-up greater than or equal to the case (2) divided by the number of patients left in the risk set (4). The same process is repeated for the second event (Risk Set ID = 2) and third event (Risk Set ID = 3).

b) Creating Risk-set Level Datasets for Matched Analyses (Conditional)

For conditional analyses, risk sets are created within each matched set in the analysis. To convert individual-level datasets to risk set-level datasets, the following steps are taken:

1. The individual-level data is sorted by follow-up time within each matched set. Follow-up time is the number of days a patient is followed post-exposure until they are either censored or have an event.
2. Within each matched set, each patient with an event is selected to contribute to a risk set.
3. The probability of exposure in the risk set is calculated using each patient in the matched set with follow-up time greater than or equal to the follow-up time of the case.

Table 8 includes example output for an individual-level, de-identified dataset; Table 9 includes an example of a translating the individual-level dataset to a risk-set level dataset.

Table 8. Example individual-level output, conditional analysis

Match ID	Exposure (1=treated)	Follow-Up Time (days)	Study Class	Event (1=event)	*Add to Risk Level Data Set?*
1	1	39	exposure	0	*No*
1	**0**	**145**	**comparator**	**1**	***Yes***
1	**0**	**191**	**comparator**	**1**	***Yes***
2	1	55	exposure	0	*No*
2	0	61	comparator	0	*No*
2	0	99	comparator	0	*No*
3	**1**	**39**	**exposure**	**1**	***Yes***
3	**0**	**39**	**comparator**	**1**	***Yes***
3	0	72	comparator	0	*No*
4	**0**	**39**	**comparator**	**1**	***Yes***
4	0	39	comparator	0	*No*
4	**0**	**71**	**comparator**	**1**	***Yes***
4	**1**	**79**	**exposure**	**1**	***Yes***
4	**0**	**84**	**comparator**	**1**	***Yes***

Table 9. Example translation to risk set level output, conditional analysis

Match ID	Risk Set ID (event indicator)	Case Exposure (1=treated)	Exposure Probability (in risk set)	Follow-Up Time (days)	Number Exposed In Risk Set	Number In Risk Set
1		1		39		
1	**1**	**0**	**0**	**145**	***0***	***2***
1	**2**	**0**	**0**	**191**	***0***	***1***
2		1		55		
2		0		61		
2		0		99		
3	**3**	**1**	**0.33**	**39**	***1***	***3***
3	**4**	**0**	**0.33**	**39**	***1***	***3***
3		0		72		
4	**5**	**0**	**0.2**	**39**	***1***	***5***
4		0		39		
4	**6**	**0**	**0.33**	**71**	***1***	***3***
4	**7**	**1**	**0.5**	**79**	***1***	***2***
4	**8**	**0**	**0**	**84**	***0***	***1***

In Table 8 and Table 9, fifteen patients in a matched analysis are converted to 8 risk sets.

In the first matched set (Match ID = 1), there are two events; both are output to the risk set dataset with their corresponding follow-up time and exposure status. The exposure probability for the first risk set (Risk Set ID =1) is calculated by taking the number of exposed individuals in the matched set with follow-up greater than or equal to the case (0) divided by the number of patients left in the risk set (2). The same process is repeated for the second event (Risk Set ID = 2). For the second matched set (Match ID= 2), there are no events, so no risk sets are created. For the third matched set (Match ID= 3), two patients have an event, both with a follow-up time of 39 days. For both patients, there are three individuals with follow-up time equal to or greater than 39, of which one is exposed. Thus, for each risk set, the exposure probability is 0.33.

For the fourth matched set (Match ID= 4), there are five patients, of which four have events. The first event occurs at 39 days of follow-up. The exposure probability is calculated as 0.2 (one individual in the exposed group / five individuals in the matched set). A risk set is created for each of the other events.

Note that "uninformative" risk sets (i.e., risk sets with only exposed or only comparator patients) will be output with an exposure probability of 0 or 1. While included on the risk-set level dataset, these risk-sets will be discarded in a case-centered logistic regression.

c) Creating Risk-set Level Datasets for Matched Analyses (Unconditional)

For unconditional analyses, risk sets are created within the entire matched population. To convert individual-level datasets to risk set-level datasets, the following steps are taken:

1. The individual-level data for the matched population is sorted by follow-up time. Follow-up time is the number of days a patient is followed post-exposure until they are either censored or have an event.
2. Each patient with an event is selected to contribute to a risk set.
3. The probability of exposure in the risk set is calculated using each patient in the matched population with follow-up time greater than or equal to the follow-up time of the case.

The risk-set level creation process is similar to the one used for the unmatched analyses, except the base population is all matched patients instead of the entire eligible population.

3. Effect Estimation Summary

To summarize, the effect estimation method depends on the data returned to SOC and the proposed analysis (Table 10).

Table 10. Effect Estimation Process Summary

Analysis Type	Effect Estimation	
	Individual-level data return	*Risk-set-level data return*
Unmatched	Population: all exposed and comparator patients Method: Cox proportional hazards model stratified by DP	Population: all exposed and comparator patients Method: risk-sets created within the entire eligible population; case-centered logistic regression.
Matched, Conditional	Population: all matched patients Method: Cox proportional hazards model stratified by DP and matched set	Population: all matched patients Method: risk-sets created within each matched set; case-centered logistic regression.
Matched, Unconditional *(1:1 matching only)*	Population: all matched patients Method: Cox proportional hazards model stratified by DP	Population: all matched patients Method: risk-sets created within the matched population; case-centered logistic regression.

4. A Note on P-value Computation

Due to computational differences between the procedures used to calculate p-values, there may be slight differences in p-values between risk-set and individual-level data analyses. In theory, the procedure used for the Cox regression (PROC PHREG) and the procedure used for the case-centered logistic regression (PROC GENMOD) maximize the same likelihood function; however, since they use slightly different numeric algorithms, there may be minor differences in numeric results. Requesters comparing individual-level and risk-set-level results for the same analyses should expect minor p-value differences.

D. OUTPUT

The MFM tool automatically generates tables of patient characteristics, stratified by exposure group, for the unmatched cohort and matched cohort, separately for each Data Partner and each monitoring period. Tables include measures of covariate balance, including absolute and standardized differences, which indicate balance in specific variables. The tables also include the number of patients in each exposure group, the number matched from each group (where appropriate), the number that experienced HOIs, and the mean person-time of follow-up.

VI. PROSPECTIVE SURVEILLANCE WITH QUERYING TOOLS

Sentinel querying tools are designed to support prospective surveillance activities in addition to one-time comparative analyses. Prospective surveillance can be performed using either the self-controlled risk interval design or with a propensity score-matched new user parallel cohort design.

Surveillance requires multiple executions of one (e.g., CIDA) or multiple (e.g., CIDA + PSA) querying tools across dynamic databases over time. This section briefly describes the Data Partner database update process, implications for prospective surveillance, and considerations and options for both the propensity score matched design and self-controlled risk interval design.

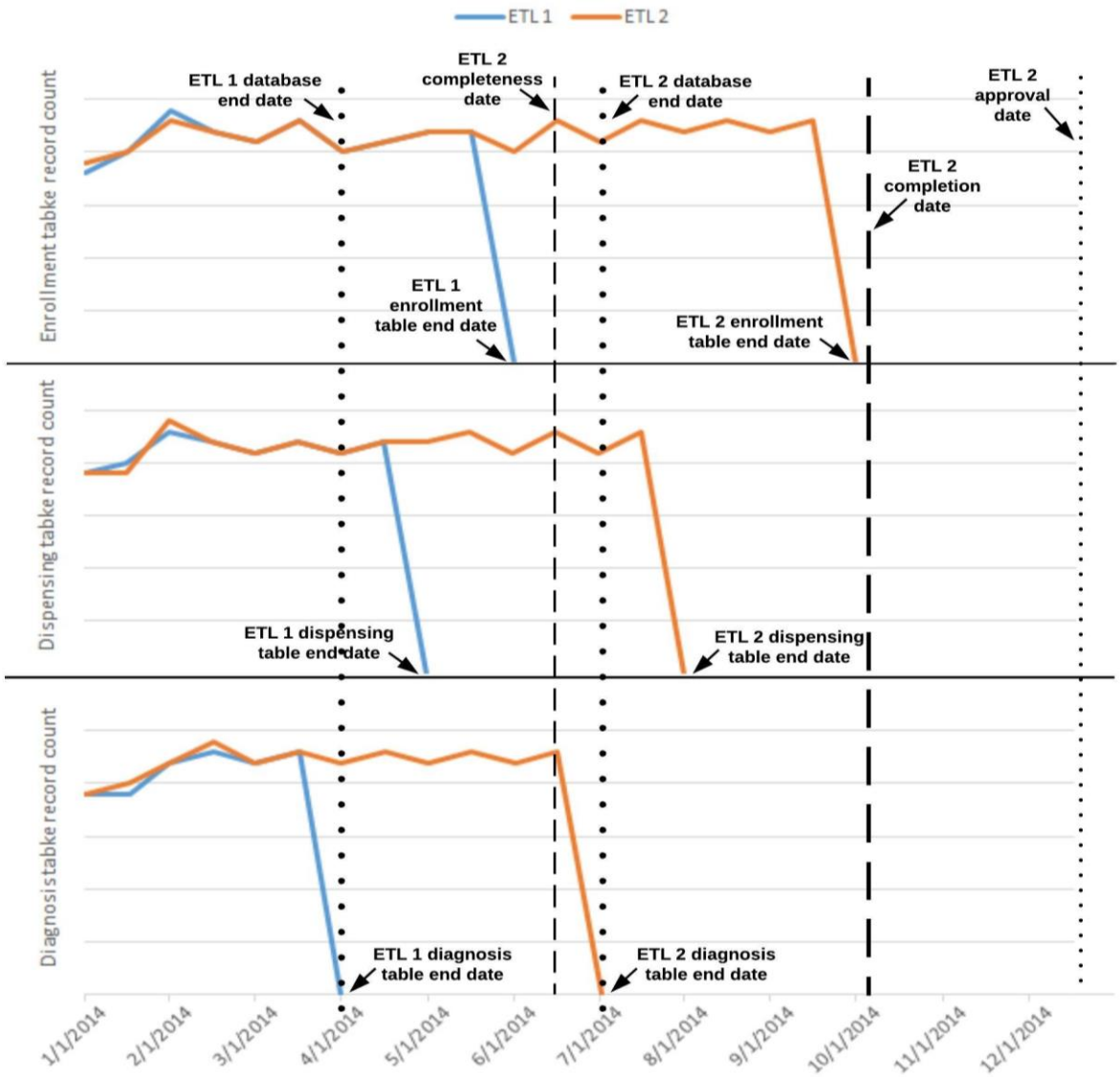
A. DATA PARTNER DATABASE UPDATE PROCESS

In order to create the Sentinel database, Data Partners extract data from multiple, local source systems, transform that data into the SCDM format, and load the information into a data warehouse to enable routine querying. This process, known as the extract, transform, load (ETL) process, is completed on a quarterly or annual basis by Data Partners to update and refresh data available to the Sentinel system. Each subsequent ETL process refreshes information from the previous version of the database, and updates the database with newly available information.

How quickly information becomes available and included in an ETL is variable by the type and source of data, and therefore may be variable by SCDM table. For example, Data Partners may have near real-time access to information about patient enrollment; however, information from claims-based systems to populate diagnosis and procedure tables may take longer to become available due to adjudication and other administrative processes. It is important, therefore, to understand the start and end dates of data availability for each table in the SCDM, to understand how recency and completeness of data for any given table may impact analyses.

Figure 32 displays trends in SCDM enrollment, dispensing, and diagnosis table record counts over time for two Data Partner ETLs. The figure demonstrates how data for newly available time periods are added with each successive ETL, and how underlying data may change for the same time period across ETLs.

Figure 32. ETL Update Process and SCDM Tables



There are several key concepts illustrated in Figure 32:

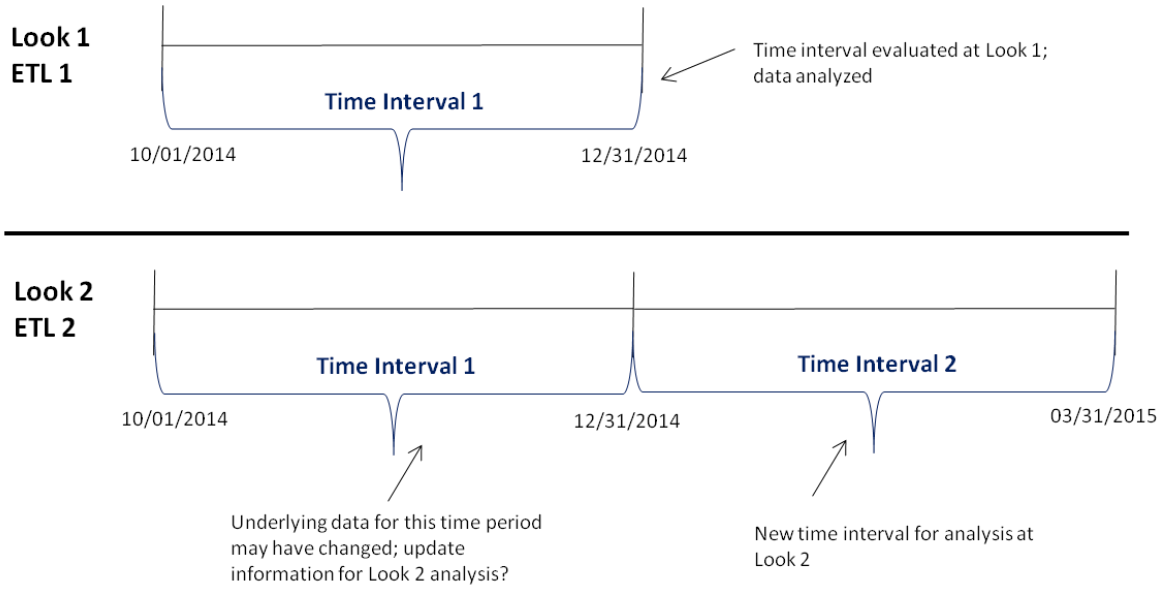
- **Table end date**: the last date for which there is a record in a database table. In Figure 32, the ETL 1 table end dates are 5/31/2014, 4/30/2014, and 3/31/2014 for the enrollment, dispensing, and diagnosis tables, respectively. The ETL 2 table end dates are 9/30/2014, 7/31/2014, and 6/30/2014 for the enrollment, dispensing, and diagnosis tables, respectively. Table end dates vary across SCDM tables (based on availability of source data) and each table is updated with 3-4 months of data from ETL 1 to ETL 2.
- **Database end date**: the earliest date across all SCDM table end dates, i.e., the last date on which all SCDM tables have at least one record.
- **Completeness date**: date on which data are deemed to be sufficiently complete for a particular query (i.e., the requester-defined completeness date determines the query end date for each program package execution). There is an inherent tradeoff between recency, or freshness, of data and data completeness; requesters must assess the cost/benefit of waiting for data to settle versus timeliness of analyses for every query. Figure 32 illustrates a potential data-driven completeness date based on evidence of record count stability. It evaluates a table completeness date (e.g., last date for which record counts appear to be stable) and then selects the earliest table completeness date to determine the ETL completeness date. There are several algorithms that could be employed to set a completeness date but, conceptually, the requester must determine the acceptable tradeoff between recency of data and data completeness for a particular request.
- **Data Partner ETL completion date**: this date coincides with the completion of the ETL process by the Data Partner. There will naturally be a lag between the database table end dates and the ETL completion date, as partners need time to create the ETL and perform quality assurance (QA) procedures.
- **ETL approval date:** this is the date that the database passes all QA checks evaluated by the Sentinel Operations Center (MSOC). This is also the date that the ETL becomes available for use in routine queries. In Figure 32, the ETL becomes available for use in queries on 12/15/2014, while the database completeness date is 6/15/2014. This lag in data availability is a function of the Data Partner's ETL process and MSOC's quality assurance processes.

Understanding Data Partner ETL processes and data availability is critical for requesters planning a surveillance activity; understanding how Data Partners' dynamic databases are changing over time is necessary to select appropriate routine querying tools and options.

1. Underlying Data Changes in Dynamic Databases

For prospective surveillance activities, typically, a program package is executed each time a Data Partner ETL is approved (i.e., each "look" at a Data Partner's data typically occurs after a database update). For example, "Look 1" at a Data Partner's data may execute on ETL 1; "Look 2," then, would execute on ETL 2. Given that each subsequent ETL may contain both new data (i.e., data for time intervals not previously included in the database) and refreshed data (i.e., modified data for time intervals previously included in the database), decisions must be made during prospective surveillance activities on whether information from time periods previously evaluated during a surveillance activity should be updated to reflect the most recent version of the database (Figure 33).

Figure 33. Look and Time Interval Concepts for Prospective Surveillance



An important factor for consideration in these decisions is whether the study design allows a fixed or variable risk window. In a fixed risk window design (e.g., SCRI design), requesters may require that the exposure and follow-up duration occur within a single time interval. This could eliminate the need to update previous time interval information across ETL versions, as a single ETL could provide complete information on exposure, follow-up time, and occurrence of health outcomes of interest. Once data has been analyzed for a specific time interval, information is never updated or analyzed again.

This approach is problematic in a variable risk window design, as it is not possible to ensure that exposures and complete follow-up time occur during a single time interval (e.g., for some exposures, follow-up time during active treatment could be several months, even years, in duration). In a variable risk window design, alternative options must be considered to determine how underlying data changes across ETLs should be addressed.

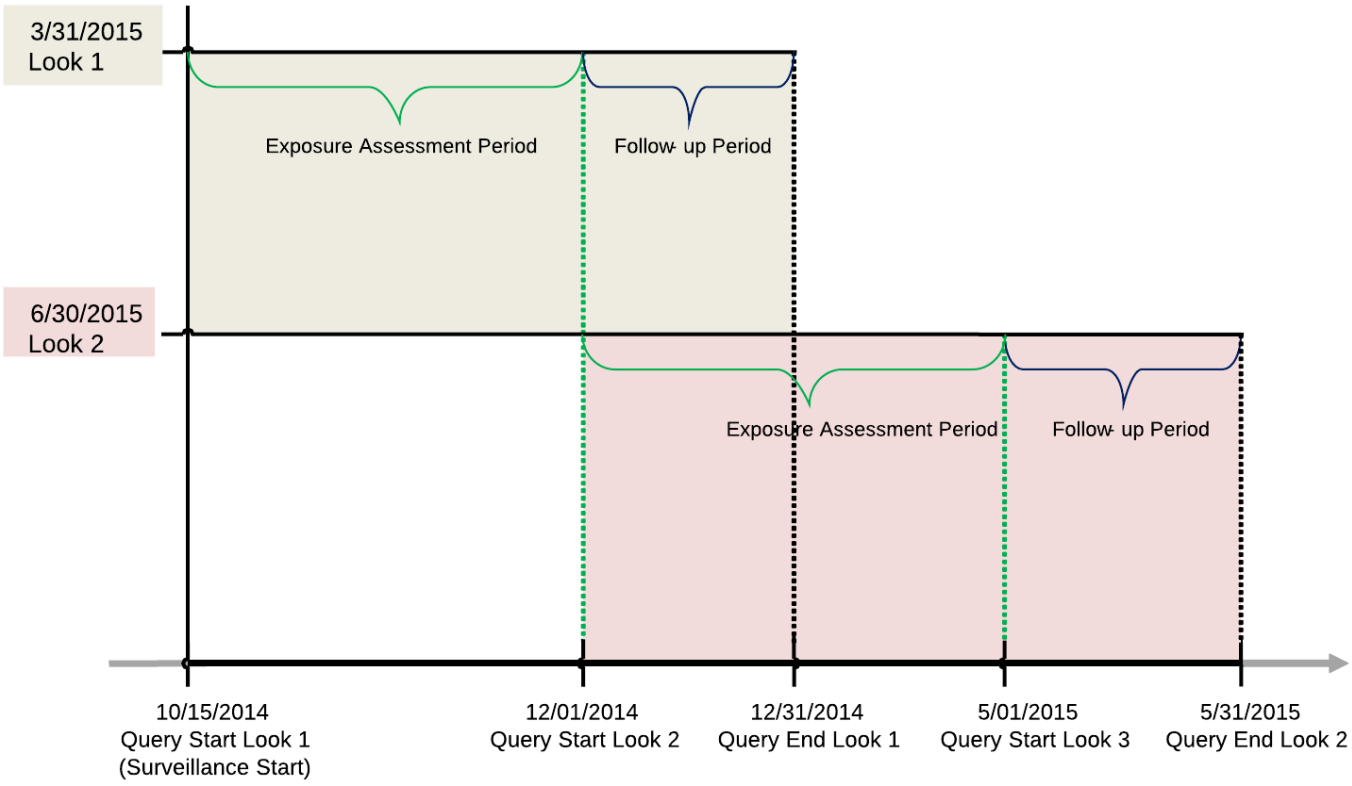
B. PROSPECTIVE SURVEILLANCE WITH THE SCRI DESIGN

Prospective surveillance using the SCRI design requires multiple executions of the CIDA tool over the course of a surveillance activity. In this fixed risk and control window design, only one option for prospective surveillance is currently available: the evaluation of mutually exclusive periods over time. Exposure and follow-up duration must occur within a single time interval (i.e., within the same version of a Data Partner's database) and, once data has been analyzed for a specific time interval, information is never updated or analyzed again. Prospective surveillance with this design, therefore, requires that surveillance teams carefully consider Data Partner database completeness dates, to improve confidence that data have settled and accurate information is captured.

For example, suppose a surveillance team wants to evaluate exposure to a new drug launched on October 15, 2014, and defines a risk window 1-15 days after exposure and control window 20-31 days after exposure. On March 31, 2015, surveillance starts. Data Partner X has data complete through December 31, 2014. The first evaluation of the data at Data Partner X (i.e., Look 1) has a query start date of October 15, 2014, an exposure assessment period of October 15, 2014 – November 30, 2014, and a follow-up period of December 1, 2014 – December 31, 2014.

Now, suppose Data Partner X updates their database again on July 31, 2015. Data are now complete through May 31, 2015.The second evaluation of the data at Data Partner X (i.e., Look 2) begins immediately after the Look 1 exposure assessment period (December 1, 2014), with an exposure assessment period of December 1, 2014 – April 30, 2015, and a follow-up period of May 1, 2015 – May 31, 2015. A visual representation of this example is included in Figure 34).

Figure 34. Prospective Surveillance with SCRI Design



For the surveillance activity shown in Figure 34, the Look 1 and Look 2 exposure assessment periods combined can be described as the cumulative exposure assessment period (at time of Look 2).

C. PROSPECTIVE SURVEILLANCE WITH PROPENSITY SCORE MATCHED DESIGN

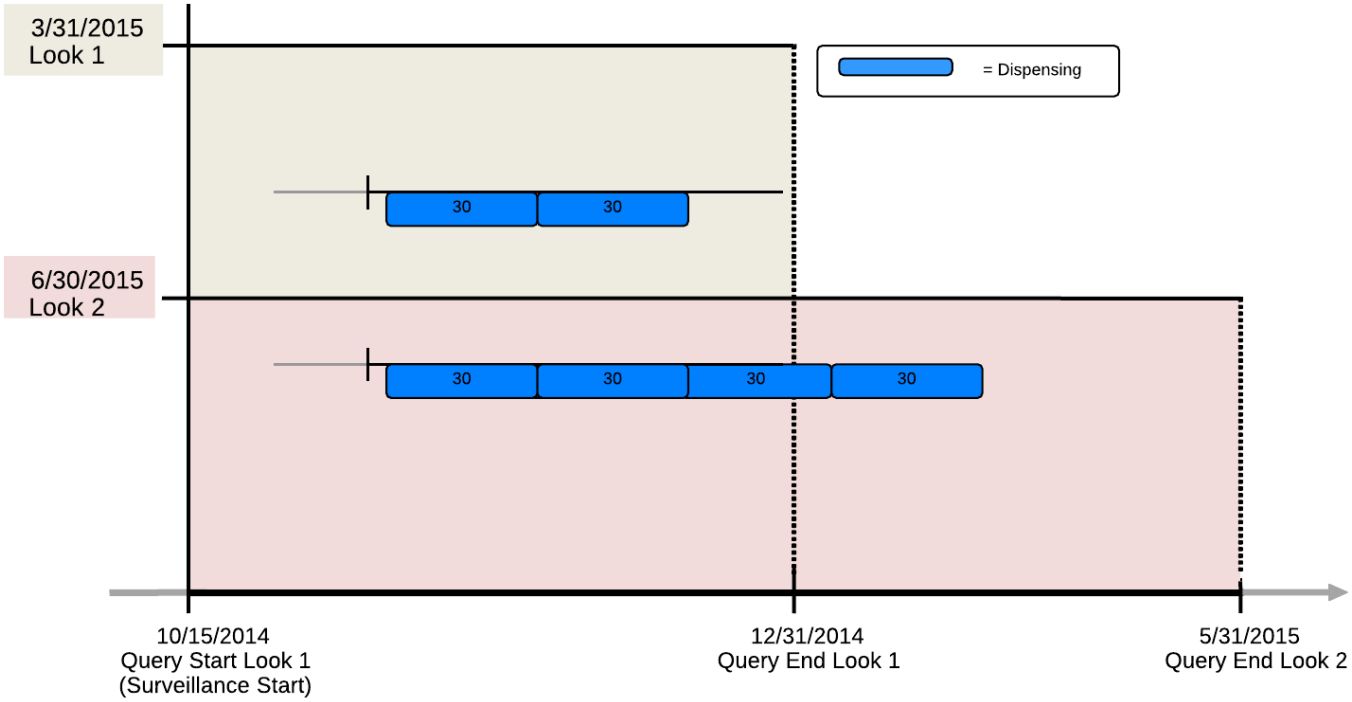
Prospective surveillance with a propensity score-matched new user parallel cohort design requires multiple executions of the CIDA and PSA tools over the course of a surveillance activity. Typically, a program package is executed each time a Data Partner updates, or refreshes, their database; therefore, each look at a Data Partner's data typically occurs after a database refresh.

Unlike the SCRI design, which sets fixed risk and control windows, the propensity score matched design allows follow-up time to be determined by each member's duration of treatment (i.e., days of exposure based on patterns of outpatient pharmacy dispensings and associated dispensing's days supply). This means that while the SCRI design can require that exposures and complete follow-up time occur in the same version of a Data Partner's database (as is done by differentiating an exposure assessment period and follow-up period), the propensity score matched approach must allow for patient follow-up to span

multiple database versions, or ETLs. This requires surveillance teams to determine how to handle changes in underlying data across database versions.

Consider an example where a patient is identified as exposed in a Data Partner's ETL 1. We observe two outpatient pharmacy dispensings, each with 30 days supply, for the exposure of interest before the ETL 1 end date of December 31, 2014 (Figure 35). For the purposes of this example, query end dates are set to the database end date for each ETL.

Figure 35. Exposure Assessment across Data Partner ETLs



After the Data Partner updates their database to ETL 2, we observe four outpatient pharmacy dispensings for the exposure of interest: three initiated before December 31, 2014, and one initiated after. The third dispensing in the series occurred during the ETL 1 time interval, but did not appear in the data until after the ETL 2 update.

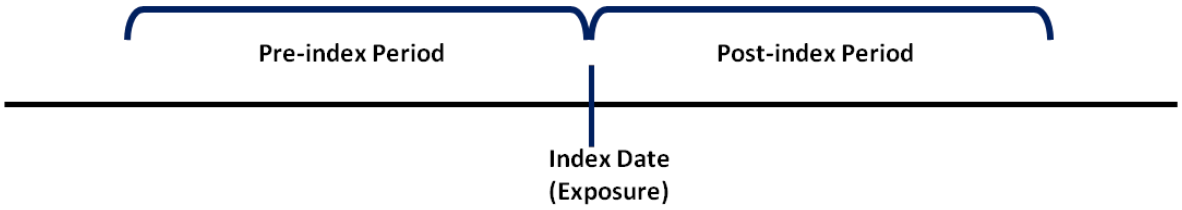
This change in data across ETLs is not confined to the duration of treatment with the exposures of interest; it is possible to see changes that may affect cohort inclusion criteria (e.g., pre-existing condition requirements no longer met after database update), event information (e.g., event is observed in ETL 1 time interval but only after ETL 2 update), and covariate assessment (e.g., presence/absence of condition of interest changes after a database update) across database versions.

While setting a time interval for each look that is before the database end date may reduce the potential for data changes (i.e., allowing more recent data to settle before initiating a query using a requester-defined data completeness date), surveillance teams must consider the possibility of data changes across databases and choose the manner in which they wish to address it in their surveillance activity.

1. Surveillance Options

Surveillance teams have three options for conducting prospective surveillance with the propensity score matched design, which differ on how underlying data changes and matches are handled during the course of the activity. In order to understand these options clearly, it is important to understand three relevant timeframes in any routine query request (Figure 36).

Figure 36. Routine Querying Tool Timeframes



The index date is the exposure initiation date, and is determined by cohort inclusion criteria. The pre-index period is the time period used to identify baseline covariates for propensity score estimation. The post-index period evaluates patient follow-up time and the occurrence of HOIs. Available surveillance options for propensity score matched analyses differ in how patient data changes are handled in each of these timeframes across multiple looks of a surveillance activity.

Surveillance Option 1:

Once a time interval is evaluated during a surveillance activity, it is never evaluated again.

If the index date or post-index at-risk time changes from one look to another, the information from the earlier look is retained and no new information is considered for the patient. If covariate information extracted from the pre-index period changes from one look to another, changes in covariate information are ignored and the patient continues to be followed.

With this option, once patients are matched the match and propensity score are retained for the duration of the surveillance activity.

Surveillance Option 2:

Once a time interval is evaluated during a surveillance activity, only information in the post-index at-risk period may be updated in subsequent looks.

If the index date changes from one look to another, the information from the earlier look is retained and the patient is lost to follow-up. If covariate information extracted from the pre-index period changes from one look to another, changes in covariate information are ignored and the patient continues to be followed. If post-index at-risk time changes from one look to another, information is updated and the patient continues to be followed.

With this option, once patients are matched the match and propensity score are retained for the duration of the surveillance activity.

Surveillance Option 3:

As underlying data change across looks, information is updated and used in analyses. Any changes in index date, pre-index period covariates and post-index period at-risk time are updated at each look. Because the pre-index period and index date information are allowed to change, matches may not be retained for the duration of a surveillance activity (as propensity scores are estimated using information in the pre-index and index periods, if information changes, propensity scores may change, and therefore

optimal matches may change). Note that even minimal underlying data changes can lead to significant cohort re-matching.

Table 11 provides a summary of the three available surveillance options.

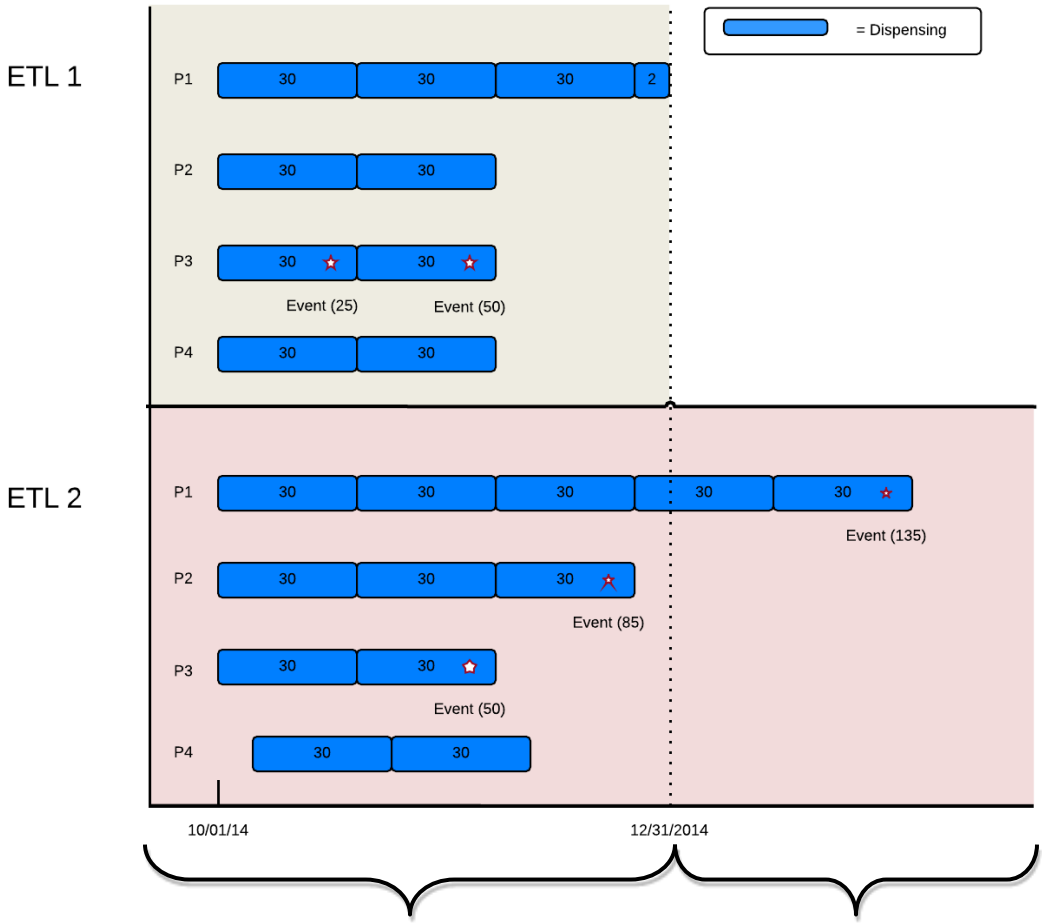
Table 11. Summary of Surveillance Option Differences: Addressing Underlying Data Changes

Approach	Pre-index period covariate changes	Index date changes[1]	Post-index at-risk period changes	New patients identified
Option 1	Information is not updated.	Information is not updated; information from the earlier look is carried forward for the duration of the surveillance activity.	Information is not updated; information from the earlier look is carried forward for the duration of the surveillance activity.	New patients ignored.
Option 2	Information is not updated.	Information is not updated; information from the earlier look is carried forward for the duration of the surveillance activity.	Information is updated.	New patients ignored.
Option 3	Information is updated.	Information is updated.	Information is updated.	New patients included.

[1] Index date changes are due to either data changes affecting cohort inclusion (e.g., enrollment requirements, inclusion/exclusion criteria, new use criteria, etc.) or changes to the patient identifier.

Figure 37 displays data for four patients (P1-P4) with and without underlying index date and post-index at-risk period changes across an ETL 1 and ETL 2. How the index date, event date, and follow-up time are determined at Look 2 for the four patients is described by surveillance option.

Figure 37. Surveillance Option Performance Examples



	Index Date	At-risk time	Event Date
Patient 1			
Option 1	10/01/2014	135	02/13/2015
Option 2	10/01/2014	135	02/13/2015
Option 3	10/01/2014	135	02/13/2015
Patient 2			
Option 1	10/01/2014	**60**	**N/A**
Option 2	10/01/2014	85	12/25/2014
Option 3	10/01/2014	85	12/25/2014
Patient 3			
Option 1	10/01/2014	**25**	**10/26/2014**
Option 2	10/01/2014	50	11/20/2014
Option 3	10/01/2014	50	11/20/2014
Patient 4			
Option 1	10/01/2014	60	N/A
Option 2	10/01/2014	60	N/A
Option 3	**10/15/2014**	60	N/A

Patient 1 has four dispensings in ETL 1, and five dispensings and an event in ETL 2. At Look 2:

- Option 1 evaluates if the patient's index date and at-risk time changed in time interval 1 at Look 2. Since the index date and at-risk time are unchanged, at-risk time and event status are updated in Look 2.
- Option 2 evaluates if the patient's index date changed in time interval 1 at Look 2. Since the index date is unchanged, at-risk time and event status are updated in Look 2.
- Option 3 only includes information in ETL 2.

Patient 2 has two dispensings and no events in ETL 1. In ETL 2, the patient has three dispensings and an event. At Look 2:

- Option 1 evaluates if the patient's index date and at-risk time have changed in time interval 1 at Look 2. Option 1 sees that the at-risk time in time interval 1 changed from 60 days to 85 days. Since there has been a change in at-risk time, the patient's information from ETL 1 is retained and no additional information is updated for the duration of the surveillance activity. The at-risk time for the duration of the surveillance activity is 60 days and the patient will not contribute an event.
- Option 2 evaluates if the patient's index date changed in time interval 1 at Look 2. Since the patient's index date is unchanged, at-risk time and event status are updated in Look 2.
- Option 3 only includes information in ETL 2.

Patient 3 has two dispensings and two events in ETL 1; in ETL 2, the patient has the same two dispensings but only one event. At Look 2:

- Option 1 evaluates if the patient's index date and at-risk time have changed in time interval 1 at Look 2. Option 1 sees that the at-risk time in time interval 1 changed from 25 days to 50 days, as the event on day 25 is not present in ETL2. Since there is a change in at-risk time, the patient's information from ETL 1 is retained and no additional information is updated for the duration of the surveillance activity. Therefore, the at-risk time for the duration of the surveillance activity is 25 days and the patient will contribute the event on day 25.
- Option 2 evaluates if the patient's index date changed in time interval 1 at Look 2. Since the patient's index date is unchanged, at-risk time and event status are updated in Look 2. The patient's at-risk time will be 50 days and they will contribute the event on day 50.
- Option 3 only includes information in ETL 2.

Patient 4 has two dispensings in ETL 1 and ETL 2, but the index date has changed. At Look 2:

- Option 1 evaluates if the patient's index date and at-risk time have changed in time interval 1 at Look 2. Option 1 sees that the index date in time interval 1 changed from 10/01/2014 to 10/15/2014. Since there is a change in index date, the patient's information from ETL 1 is retained and no additional information is updated for the duration of the surveillance activity. The exposure initiation date will remain 10/1/2014 and the at-risk period will remain 60 days for the duration of the surveillance activity.
- Option 2 evaluates if the patient's index date changed in time interval 1 at Look 2. Since there has been a change in index date, the patient's information from ETL 1 is retained and no additional information is updated for the duration of the surveillance activity. Therefore, the exposure initiation date will remain 10/1/2014 and the at-risk period will remain 60 days for the duration of the surveillance activity.
- Option 3 only includes information in ETL 2. Therefore, the patient's index date will be 10/15/2014, and the patient is eligible to further accumulate at-risk time in later ETLs.

VII. REPORTING TOOLS

After execution of the CIDA tool, a PDF report can be produced to visually summarize the results of the query. This section describes the available reports.

A. TYPE 1 AND TYPE 2 REPORT

At the completion of a query utilizing either the 1) background rate calculation cohort identification strategy or 2) exposures and follow-up time cohort identification strategy a Type 1/Type 2 report can be produced with the following elements:

- Baseline covariate table (i.e. "Table 1"). If requested, the report will include a table containing the baseline prevalence of covariates of interest, the distribution of drug and medical utilization, and the distribution of Charlson/Elixhauser combined comorbidity score for each cohort of interest.
- Summary table. The report will contain a table with aggregate counts of users, episodes, events, follow-up time, and eligible members for each cohort of interest. This table can be produced overall, and for any available stratification (i.e. by sex or age-group).
- Cumulative Density Function (CDF) and Kaplan-Meier (KM) plots. If requested, CDF plots for time to censor for each cohort of interest stratified by reason for censor will be produced. Additionally, for an exposures and follow-up time cohort identification strategy, KM plots for time to event can be produced.

VIII. APPENDIX A: PROGRAM PACKAGE AND EXECUTION

When implementing modular programs within the SDD, the SOC uses a uniform folder structure to facilitate communications between SOC and Data Partners and to streamline file management. This appendix describes the program package structure and requirements for package execution.

A. PROGRAM PACKAGE

Each request package distributed by SOC is assigned a unique Request Identifier, and contains several folders to organize program inputs and outputs:

- *sasprograms*: folder contains the master SAS program that must be edited and then executed by the Data Partner.
- *inputfiles*: folder contains input files and lookup tables needed to execute a request. Input files contain parameter values specific to a particular request (*e.g.,* medical product exposures and outcomes of interest, continuous enrollment requirements, and incidence criteria). Input files are created for each request by the SOC query fulfillment team; the contents of this folder are not edited by the Data Partner. The folder also contains one subfolder:
    - *macros*: folder contains the macros that comprise the modular program. The contents of this folder are not edited by the Data Partner.
- *msoc*: folder contains output generated by the request that should be sent to SOC.
- *dplocal*: folder contains output generated by the request that should remain with the Data Partner (and may be used to facilitate follow-up queries).

1. Common Components

Prior to executing the request package, a set of SAS programs known as common components must be initialized. In this context, common components refer to a set of SAS programs that provide appropriate site-specific attributes (e.g., data partner description variables, SCDM table names, folder paths, data completeness dates, etc.) to distributed SAS program packages at the time of code execution. More specifically, when an executing SAS program package accesses the file ms_common_components.sas, global macro variable definitions for key site-specific attributes are made available to the calling program. In this context, common components support two important goals: 1) streamline the setup for the distributed SAS program packages, 2) improve the accuracy of results.

Users must specify the location of their common components file path in the master SAS program in the *sasprograms* folder in order for the package to execute. For more information about common components installation, and to download the SAS programs, visit the common components page on the Sentinel website.

2. Master Program Parameters

In the master SAS program, there are several parameters that must be specified. These include the common components include file, project, work plan, and Data Partner identifiers, and a run identifier. Note that all main program parameters specified are fixed for a single execution of the program. Table 12 contains detailed specifications for master program parameters.

Table 12. CIDA Tool Master Program Parameter Specifications

Parameter	Field Name	Description
Common Components Include file	MSCC	**Details**: location for user's common components file path **Defined by**: User programmer **Input type**: Required **Format:** Alphanumeric **Example**: MSCC = //Sentinel/common-components/ms_cc.sas
Prior DPLOCAL	DPLPRIOR	**Details**: location of the DPLOCAL file path for the prior look when using CIDA for prospective surveillance with propensity score matched design **Note 1**: Should be used only when utilizing surveillance options 1 and 2 and should be left blank for option 3 **Note 2**: Should be left blank for Look 1. Will point to the DPLOCAL file path of Look 1 when running Look 2, and so on. **Defined by:** User programmer **Input type:** Optional **Format:** Alphanumeric **Example**: DPLPRIOR=//to16_cap_mpl2r_wp01_nsdp_v01/DPLOCAL/

Parameter	Field Name	Description
Patients to Exclude List File Path	PTSTOEXCLUDE	**Details**: optional. Location of user's patients to exclude list. Allows Data Partners to exclude patients from a particular request. The file must contain on variable, PatID, and list all PatID values to exclude from the request. **Defined by**: User programmer **Input type**: Required **Format:** Alphanumeric **Example**: PTSTOEXCLUDE = indata.ptstoexclude
Project Identifier	MSPROJID	**Details**: project identifier for internal SOC identification and tracking. **Defined by:** Request programmer **Input type:** Required **Format:** Alphanumeric **Example**: MSPROJID=to16_cap
Work Plan Type	MSWPTYPE	**Details**: work plan type for internal SOC identification and tracking. **Defined by:** Request programmer **Input type:** Required **Format:** Alphanumeric **Example**: MSWPTYPE=mpl2r
Work Plan Identifier	MSWPID	**Details**: work plan identifier for internal SOC identification and tracking. **Note 1:** should follow the format [wp###]. **Note 2:** should be used to uniquely identify a modular program request. **Defined by:** Request programmer **Input type:** Required **Format:** Alphanumeric **Example**: MSWPID= wp01
Data Partner Identifier	MSDPID	**Details**: Data Partner identifier for internal SOC identification and tracking. **Note 1:** if a package is not Data Partner specific, MSDPID should equal "nsdp". **Defined by:** Request programmer **Input type:** Required **Format:** Alphanumeric **Example**: MSDPID =nsdp

Parameter	Field Name	Description
Version Identifier	MSVERID	**Details**: version identifier for internal SOC identification and tracking. Should track each re-distribution of the package (if multiple distributions are required). **Note 1:** should follow the format [v##]. **Defined by:** Request programmer **Input type:** Required **Format:** Alphanumeric **Example**: MSVERID =v01

IX. APPENDIX B: CIDA TOOL TECHNICAL DOCUMENTATION

The CIDA tool is designed to be executed both as a standalone tool and in combination with compatible analytic tools. This technical specification document details the lookup tables, program parameters and input files that must be specified to execute the CIDA tool. Where applicable, selections that have implications for subsequent processing with an analytic tool (*e.g.,* the PSA tool) are noted.

A. LOOKUP TABLES, PROGRAM PARAMETERS, AND INPUT FILES

1. Lookup Tables

There are several lookup tables that may be required for the execution of the CIDA tool depending on the nature of the request. These lookup tables are created and maintained by the SOC.

a) Lab Code Lookup Table

The Lab Code Lookup File is required in the *inputfiles* folder if a request queries the SCDM Laboratory Result table. The Lab Code Lookup File is a master lookup file of SOC-defined "lab codes" denoting allowable combinations of lab test name, lab test subcategory, specimen source, result type, fasting indicator, patient location, and result unit.

Lab codes are 14-digit identifiers developed by the SOC to represent a unique laboratory test result value for querying. The first digit of this code is an "L" indicative of a lab code, digits 2-4 indicate a unique lab test name, digit 5 indicates a unique result type value (numeric or character), digits 6-7 indicate a unique lab test subcategory, digit 8 represents a unique fasting indicator value, digits 9-10 indicate a unique specimen source, digits 11-12 indicate a unique patient location, and digits 13-14 indicate the result unit.

As SOC continues to develop and expand the SCDM Laboratory Result table and allowable values, this master lookup file may be modified to ensure that the program always queries the SCDM Laboratory Result table based on current specifications.

The Lab Codes defined in this lookup table are used to query the lab test and results of interest just as any NDC, diagnosis, or procedure code is queried in the SDD.

Table 13 defines the variables included in this lookup table.

Table 13. Lab Code Lookup File

Parameter	Field Name	Description
Lab Test Name	MS_TEST_NAME	MS_Test_Name value in the SCDM Laboratory Results table.
Result Type	RESULT_TYPE	Result_Type value in the SCDM Laboratory Results table.
Lab Test Subcategory	MS_TEST_SUB_CATEGORY	MS_Test_Sub_Category value in the SCDM Laboratory Results table.
Fasting Indicator	FAST_IND	Fast_Ind value in the SCDM Laboratory Results table.
Lab Specimen Source	SPECIMEN_SOURCE	Specimen_Source value in the SCDM Laboratory Results table.
Patient Location	PT_LOC	Pt_Loc value in the SCDM Laboratory Results table.
Lab Result Unit	MS_RESULT_UNIT	MS_Result_Unit value in the SCDM Laboratory Results table.
Lab Code	CODE	SOC-defined code indicative of the MS_Test_Name, Result_Type, MS_Test_Sub_Category, Fast_Ind, Specimen_Source, Pt_Loc and MS_Result_Unit combination. **CODE values in the lookup table should contain an exhaustive list of all combinations of these variable values.** **Note 1**: CODE values can be listed in program Input Files to query the desired laboratory result values, just as any other NDC, diagnosis and/or procedure code is queried.

b) Comorbidity Score Code Lookup Table

The Comorbidity Score Code Lookup table is required in the *inputfiles* folder if a request is calculating the Charlson/Elixhauser combined comorbidity score. The file contains the comprehensive set of codes used to define the medical conditions contributing to the calculation of the score. Table 14 defines the variables included in this lookup table.

Table 14. Comorbidity Score Codes Lookup Table

Parameter	Field Name	Description
Code	CODE	Code to define medical condition.
Code Type	CODETYPE	=09. Parameter included allowing for potential future expansion to other code types.
Code Group	CODE_GRP	=DIAG. Parameter included to allow for potential future expansion to other code types.
Group	GROUP	Numeric indicator identifying individual conditions used to calculate the comorbidity score.
Group Description	GROUP_DESCR	Description of the condition defined by the GROUP value.
Group Weight	WEIGHT	Weight that each GROUP contributes to the comorbidity score calculation.
Wildcard Indicator	WILDCARD	Y/N indicator if the CODE value should be processed as "starts with," to include both parent and child codes.

c) Drug Class Lookup Table

The Drug Class Lookup table is required in the *inputfiles* folder if a request requires the use of drug utilization metrics. The table is used as reference to calculate the number of unique dispensings, unique generics, and unique drug classes dispensed per cohort member during the covariate evaluation window.

The Drug Class Lookup table includes a list of NDCs by unique generic name indicator and unique drug class indicator. Table 15 contains specifications for this lookup table.

Table 15. Drug Class Lookup File Specification

Parameter	Field Name	Description
National Drug Code	NDC	11-digit NDC.
Generic Name Identifier	GENERIC	SOC-defined character string indicative of a unique generic name. SOC maintains the mapping key to actual generic name locally.
Class Name Identifier	CLASSNAME	SOC-defined character string indicative of a unique generic name. SOC maintains the mapping key to actual generic name locally.

d) Geography Lookup Table

The Geography Lookup table is required in the *inputfiles* folder if a request requires stratification of results by geographic location. The table is used as reference to map 5-digit ZIP code to State, Health and Human Services (HHS) Region, and Census Bureau region.

Table 16 contains the specifications for this lookup table.

Table 16. Geography Lookup File Specification

Parameter	Field Name	Description
ZIP Code	ZIP	5-digit ZIP code.
State Code	STATECODE	2-digit state code.
HHS Region	HHS_REGION	01 = Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont. 02 = New Jersey, New York, Puerto Rico, Virgin Islands. 03 = Delaware, Maryland, Pennsylvania, Virginia, West Virginia, District of Columbia. 04 = Alabama, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, Tennessee. 05 = Illinois, Indiana, Michigan, Minnesota, Ohio, Wisconsin. 06 = Arkansas, Louisiana, New Mexico, Oklahoma, Texas. 07 = Iowa, Kansas, Missouri, Nebraska. 08 = Colorado, Montana, North Dakota, South Dakota, Utah, Wyoming. 09 = Arizona, California, Hawaii, Nevada, American Samoa, Federated States of Micronesia, Guam, Palau. 10 = Alaska, Idaho, Oregon, Washington. 11 = Northern Mariana Islands, Marshall Islands. Missing = Missing.

Parameter	Field Name	Description
Census Bureau Region	CB_REGION	NE = Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont, New Jersey, New York, Pennsylvania. MW = Illinois, Indiana, Michigan, Ohio, Wisconsin, Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota. S = Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia, Alabama, Kentucky, Mississippi, Tennessee, Arkansas, Louisiana, Oklahoma, Texas. W = Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming, Alaska, California, Hawaii, Washington. Other = Northern Mariana Islands, Marshall Islands, Puerto Rico, US Virgin Islands, American Samoa, Micronesia, Guam, Palau. Missing = Missing.

2. Main Program Parameters

There are several main program parameters that must be specified (Table 17). These include a run identifier, indicators for the start and end dates for the query period, an indicator if and what additional analyses are being performed after CIDA tool execution, and the names of all input files. These parameter values should be set in a program called run_programs.sas, located in the *inputfiles* folder. Note that all main program parameters specified are fixed for a single execution of the program. Table 17 contains detailed specifications for main program parameters.

Table 17. CIDA Tool Main Program Parameter Specifications

Parameter	Field Name	Description
Run Identifier	RUNID	**Details**: run identifier for internal SOC identification and tracking. Should uniquely identify each execution of a modular program within the same work plan. **Note 1:** should follow the format [r##]. **Defined by:** Request programmer **Input type:** Required **Format:** Alphanumeric **Example**: RUNID =r01
Query Period Start Identifier	PERIODIDSTART	**Details:** identifies at what time period the modular program should begin execution. **Note 1:** for TYPE 1, TYPE 2, TYPE 5, and TYPE 6, PERIODIDSTART should correspond to the PERIODID value in the input MONITORINGFILE to identify at what time period the modular program should begin execution.

Parameter	Field Name	Description
		Note 2: for TYPE 3 analyses, PERIODIDSTART should correspond to the "look" number; i.e., indicate what data extraction an analysis is being performed for a sequential analysis activity (PERIODIDSTART value for the third iteration of a request should = 3). For Type 3 analyses, PERIODIDSTART should always = PERIODIDEND. **Note 3:** for Type 4 analysis, PERIODSTART binds the delivery date and not the index date (calculated start of pregnancy). **Note 4**: will now require multiple executions of QRP if different study start dates are required. **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example**: PERIODIDSTART= 1
Query Period End Identifier	PERIODIDEND	**Details:** identifies at what time period the modular program should end execution. **Note 1:** for TYPE 1, TYPE 2, TYPE 5, and TYPE 6, PERIODIDEND should correspond to the PERIODID value in the input MONITORINGFILE to identify at what time period the modular program should end execution. A PERIODIDEND value >1 allows the modular program to generate and output information based on different requester-defined time periods. **Note 2**: for Type 3 analyses, PERIODIDEND should correspond to the "look" number; i.e., indicate what data extraction an analysis is being performed for a sequential analysis activity (PERIODIDEND value for the third iteration of a request should = 3). For Type 3 analyses, PERIODIDSTART should always = PERIODIDEND. **Note 3**: for requests that are not part of a sequential analysis activity, PERIODIDEND will often be set to "1" as a single query period is needed. **Note 4:** for Type 4 analysis, PERIODEND binds the delivery date and not the index date (calculated start of pregnancy). **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example**: PERIODIDEND= 3

Parameter	Field Name	Description
Further Analysis Indicator	ANALYSIS	**Details:** Indicates what, if any, additional analyses will be performed on the cohort after extraction. • **Note 1:** baseline covariate table will be output for Type 1, Type 2, Type 3, Type 4, and Type 5 analyses, independent of the value of "ANALYSIS" parameter. The program will generate an output tables ([RUNID]_baseline_[PERIODID].sas7bdat) containing the baseline prevalence of covariates of interest. Requester may include following valid values: • **PS:** indicates that, the user would like to execute the propensity score estimation and matching modules. Relevant for Type 2 analyses only. • **MS:** indicates that, the user would like to execute the multi-factor matching algorithm. Relevant for Type 2 analyses only. <blank> no additional analyses requested **Note 1:** If ANALYSIS= "PS" or ANALYSIS = "MS", a TYPE2FILE must be specified. **Defined by:** Requester **Input type:** Optional **Format:** Alphanumeric **Example**: ANALYSIS = PS
Save all SCDM Data Indicator	FREEZEDATA	**Details:** indicates if all SCDM data for patients selected in the cohort(s) of interest will be saved in the *dplocal* folder for further processing. Allowable values are: • **Y:** Yes (excludes never-exposed cohort) • **A:** All (includes never-exposed cohort) • **N:** No **Note 1**: should be set to "Y" for prospective surveillance using the self-controlled risk interval design (i.e., a "Type 3" analysis). **Note 2**: due to storage space concerns regarding saving all SCDM data for patients selected in the never-exposed cohort, the default will be to only save active-exposure cohorts unless otherwise specified. **Defined by:** Requester **Input type:** Required **Format:** Alphanumeric

Parameter	Field Name	Description
		Example: FREEZEDATA = Y
Monitoring File	MONITORINGFILE	**Details:** name of the SAS dataset defining the time period(s) for each data extraction. **Note 1**: not required for "Type 3" analyses. **Named by:** Request programmer **Input type:** Required for Type 1, Type2, Type 4,Type 5, and Type 6 analyses **Format: .**sas7bdat file format **Example:** MONITORINGFILE=drugname_monitoring
Cohort File	COHORTFILE	**Details:** name of the SAS dataset defining the cohort identification strategy used and continuous enrollment requirements. **Named by:** Request programmer **Input type:** Required **Format: .**sas7bdat file format **Example:** COHORTFILE=drugname_cohort
Type 1 File	TYPE1FILE	**Details:** name of the SAS dataset defining parameters required for a background rate calculation cohort identification strategy. **Note 1:** a single execution of the program can only process one cohort identification strategy (i.e., only TYPE1FILE or TYPE2FILE or TYPE3FILE or TYPE4FILE or TYPE5FILE or TYPE6FILE may be specified). **Named by:** Request programmer **Input type:** Optional; TYPE1FILE or TYPE2FILE or TYPE3FILE or TYPE4FILE or TYPE5FILE or TYPE6FILE must be specified **Format: .**sas7bdat file format **Example:** TYPE1FILE=drugname_type1
Type 2 File	TYPE2FILE	**Details:** name of the SAS dataset defining parameters required for an exposures and follow-up time cohort identification strategy. **Note 1:** a single execution of the program can only process one cohort identification strategy (i.e., only TYPE1FILE or TYPE2FILE or TYPE3FILE or TYPE4FILE or TYPE5FILE or TYPE6FILE may be specified). **Named by:** Request programmer **Input type:** Optional; TYPE1FILE or TYPE2FILE or TYPE3FILE or TYPE4FILE or TYPE5FILE or TYPE6FILE must be specified **Format: .**sas7bdat file format **Example:** TYPE2FILE=drugname_type2

Parameter	Field Name	Description
Type 3 File	TYPE3FILE	**Details:** name of the SAS dataset defining parameters required for a self-controlled risk interval design cohort identification strategy. **Note 1:** a single execution of the program can only process one cohort identification strategy (i.e., only TYPE1FILE or TYPE2FILE or TYPE3FILE or TYPE4FILE or TYPE5FILE or TYPE6FILE may be specified). **Named by:** Request programmer **Input type:** Optional; TYPE1FILE or TYPE2FILE or TYPE3FILE or TYPE4FILE or TYPE5FILE or TYPE6FILE must be specified **Format: .**sas7bdat file format **Example:** TYPE3FILE=drugname_type3
Type 3 Metadata File	T3METADATA	**Details:** name of the SAS dataset defining parameters required for prospective surveillance with a self-controlled risk interval design cohort identification strategy. **Note 1:** for PERIODIDSTART and PERIODIDEND = 1, no T3METADATA file should b e specified. For PERIODIDSTART and PERIODIDEND >1, a T3METADATA file must be specified. **Note 2**: T3METADATA files are generated by the modular program, for use during a subsequent execution of the modular program during a sequential analysis activity. For example, when PERIODIDSTART and PERIODIDEND=1, the program package will output to the *msoc* folder a dataset called Metadata_for_time_period_1.sas7bdat. In the subsequent package where PERIODIDSTART and PERIODIDEND=2, this Metadata_for_time_period_1.sas7bdat file should be included in the *inputfiles* folder of the package and T3METADATA= Metadata_for_time_period_1. This allows the package for look 2 to access metadata generated by look 1. **Named by:** Request programmer **Input type:** Required for Type 3 analyses with PERIODIDSTART>1 **Format: .**sas7bdat file format **Example:** T3METADATA= Metadata_for_time_period_1

Parameter	Field Name	Description
Type 4 File	TYPE4FILE	**Details:** name of the SAS dataset defining parameters required for a pregnancy episodes cohort identification strategy. **Note 1:** a single execution of the program can only process one cohort identification strategy (i.e., only TYPE1FILE or TYPE2FILE or TYPE3FILE or TYPE4FILE or TYPE5FILE or TYPE6FILE may be specified). **Named by:** Request programmer **Input type:** Optional; TYPE1FILE or TYPE2FILE or TYPE3FILE or TYPE4FILE or TYPE5FILE or TYPE6FILE must be specified **Format: .**sas7bdat file format **Example:** TYPE4FILE=drugname_type4
Type 5 File	TYPE5FILE	**Details:** name of the SAS dataset defining parameters required for a medical product utilization cohort identification strategy. **Note 1:** a single execution of the program can only process one cohort identification strategy (i.e., only TYPE1FILE or TYPE2FILE or TYPE3FILE or TYPE4FILE or TYPE5FILE or TYPE6FILE may be specified). **Named by:** Request programmer **Input type:** Optional; TYPE1FILE or TYPE2FILE or TYPE3FILE or TYPE4FILE or TYPE5FILE or TYPE6FILE must be specified **Format: .**sas7bdat file format **Example:** TYPE5FILE=drugname_type5
Type 6 File	TYPE6FILE	**Details:** name of the SAS dataset defining parameters required for a manufacture level product utilization and switching patterns cohort identification strategy. **Note 1:** a single execution of the program can only process one cohort identification strategy (i.e., only TYPE1FILE or TYPE2FILE or TYPE3FILE or TYPE4FILE or TYPE5FILE or TYPE6FILE may be specified). **Named by:** Request programmer **Input type:** Optional; TYPE1FILE or TYPE2FILE or TYPE3FILE or TYPE4FILE or TYPE5FILE or TYPE6FILE must be specified **Format: .**sas7bdat file format **Example:** TYPE6FILE=drugname_type6
Cohort Codes File	COHORTCODES	**Details:** name of the SAS dataset listing codes used to define the cohort.

Parameter	Field Name	Description
		Note 1: for Type 1, Type 2, Type 3, Type 5 and Type 6 analyses this file defines the index date **Note 2:** for Type 4 analysis this file defines the delivery date **Named by:** Request programmer **Input type:** Required **Format: .**sas7bdat file format **Example:** COHORTCODES=drugname_cohortcodes
Inclusion/Exclusion Codes File	INCLUSIONCODES	**Details:** name of the SAS dataset listing codes used to define additional cohort inclusion and exclusion criteria. **Note 1**: Cohort inclusion/exclusion criteria are assessed relative to the index date. In Type 5, a patient can have only one index date (first) index date. **Named by:** Request programmer **Input type:** Optional **Format: .**sas7bdat file format **Example:** INCLUSIONCODES=drugname_inclusioncodes
Covariate Codes File	COVARIATECODES	**Details:** name of the SAS dataset listing codes used to define covariates. **Note 1:** must be specified if ANALYSIS=PS. May be specified for requests with ANALYSIS=ADS. **Named by:** Request programmer **Input type:** Optional **Format: .**sas7bdat file format **Example:** INCLUSIONCODES=drugname_covariatecodes
Profile Output Generation Indicator	PROFILE	Valid values include: • **Y:** generate and output [*RUNID*]_profile_[*PERIODID]* • **N:** do not generate or output [*RUNID*]_profile_[*PERIODID]* • **<blank>:** do not generate or output [*RUNID*]_profile_[*PERIODID]*
Most Frequent Utilization File	MFUFILE	**Details:** name of the SAS dataset to request most frequent utilization assessment. **Note 1**: this file is available to specify for all Types of analysis and reference a previously defined index date. **Named by:** Request programmer **Input type:** Optional

Parameter	Field Name	Description
		Format: .sas7bdat file format **Example:** MFUFILE = most_frequent_use
Stockpiling File	STOCKPILINGFILE	**Details:** name of the SAS dataset defining how dispensings, days supplied, and amount supplied are handled by the MP (if defaults should be modified). **Note 1**: this file needs to be specified only if program defaults must be changed. **Named by:** Request programmer **Input type:** Optional **Format: .**sas7bdat file format **Example:** STOCKPILINGFILE = drugname_stockpiling
Utilization File	UTILFILE	**Details:** name of the SAS dataset defining the drug and medical utilization evaluation windows. **Note 1:** may be specified if ANALYSIS=PS orANALYSIS=ADS. **Named by:** Request programmer **Input type:** Optional **Format: .**sas7bdat file format **Example:** UTILFILE =drugname_util
Combo File	COMBOFILE	**Details:** name of the SAS dataset used to specify the algorithms used to define complex events (*i.e.,* any event that cannot be defined as a simple list of codes, which requires temporal relationships or multiple criteria to be true to define an event). **Note 1:** specifications for this file and detailed documentation of Combo tool functionality can be found in a separate document: 'Sentinel Toolkit Combo Tool Documentation'. **Named by:** Request programmer **Input type:** Optional **Format: .**sas7bdat file format **Example:** COMBOFILE=drugname_combofile
Comorbidity Score File	COMORBFILE	**Details:** name of the SAS dataset defining the Charlson/Elixhauser combined comorbidity score calculation and stratification parameters. **Note 1:** may be specified if ANALYSIS=PS orANALYSIS=ADS. **Note 2:** requires inclusion of the Comorbidity Score Code Lookup Table in the *inputfiles* folder. **Named by:** Request programmer **Input type:** Optional **Format: .**sas7bdat file format

Parameter	Field Name	Description
		Example: COMORBFILE =drugname_comorbfile
Drug Class Lookup Table	DRUGCLASSFILE	**Details:** name of the SAS lookup table containing a list of NDCs by unique generic name indicator and unique drug class indicator. **Named by:** Request programmer **Input type:** Required **Format: .**sas7bdat file format **Example:** DRUGCLASSFILE =drugclass
Pregnancy Duration File	PREGDUR	**Details:** name of the SAS dataset defining the codes used to define pregnancy episode duration. Required for pregnancy episodes cohort identification strategy, Type 4 analysis (not applicable for other Types). **Named by:** Request programmer **Input type:** Optional **Format: .**sas7bdat file format **Example:** PREGDUR =t4_pregdur
Mother-Infant cohort file	MICOHORTFILE	**Details:** contains parameters to query the Mother-Infant Linkage Table and to specify criteria to build new pregnant exposure, pregnant comparator, and preganant unexposed cohorts. **Note 1:** Only applicable for Type 4 analysis creating cohorts for further analysis with the PSA tool. **Defined by:** Request programmer **Input type:** Optional **Format:** .sas7bdat file **Example:** MICOHORTFILE = w014_micohort
Surveillance Strategy for Propensity Score Analyses	SURVEILLANCEMODE	**Details:** specifies the method for performing prospective surveillance when using the propensity score matching tool (ANALYSIS=PS). Leave parameter value blank for analyses not using the propensity score matching tool. Valid values include: • **F:** Option 1 (Full lock). • **P:** Option 2 (Partial lock). • **<blank>:** Option 3 (No lock). **Defined by:** Requester **Input type:** Optional **Format:** Alphanumeric **Example**: SURVEILLANCEMODE = F
Lab Code Lookup Table	LABSCODEMAP	**Details:** name of the SAS lookup table defining lab codes used to query the SCDM Laboratory Result table.

Parameter	Field Name	Description
		Note 1: must be included in the *inputfiles* folder if laboratory result values are queried. **Named by:** Request programmer **Input type:** Optional **Format: .**sas7bdat file format **Example:** LABSCODEMAP =lab_lookup
Geography Lookup Table	ZIPFILE	**Details:** name of the SAS lookup table defining ZIP codes, state codes, Health and Human Services region codes and Census Bureau region codes. **Note 1**: must be included in the *inputfiles* folder if results are output by geographic location **Named by:** Request programmer **Input type:** Optional **Format: .**sas7bdat file format **Example:** ZIPFILE =Ziplkp
Turn off envelope macro	RUN_ENVELOPE	**Details:** specifies the method for data cleaning with the envelope macro. Valid values include: • **0:** Run envelope macro. Reclassify outpatient (AV), emergency department (ED), and other ambulatory (OA) encounters that occur during an inpatient stay as inpatient (IP) encounters. • **1:** Do not run envelope on IP ADate. Reclassify outpatient (AV), emergency department (ED), and other ambulatory (OA) encounters that occur during an inpatient stay as inpatient (IP) encounters. Do not reclassify if encounter occurs on day of admission (ADate). • **2:** Turn off envelope. No reclassification. **Defined by:** Requester **Input type:** Required **Format:** Alphanumeric **Example**: RUN_ENVELOPE = 0
Report Type	CREATEREPORT_TYPE	**Details:** specifies which report to produce following the execution of the CIDA tool. • Specify "1" to produce a background rate calculation cohort identification strategy report. • Specify "2" to produce a exposures and follow-up time cohort identification strategy report. **Note 3:** Leave blank if no report is produced.

Parameter	Field Name	Description
		Named by: Request programmer **Input type:** Optional **Format:** Alphanumeric **Example:** CREATEREPORT_TYPE=1
Create Report File	CREATEREPORT_FILE	**Details:** name of the SAS dataset used to specify the elements of the reports and to customize the report. **Named by:** Request programmer **Input type:** Optional **Format: .**sas7bdat file format **Example:** CREATEREPORT_FILE=report_parameters
Distribution of Index Codes Output Indicator	DISTINDEX	**Details:** indicates if an optional output file with the distribution of index defining codes will be output for this request. Valid values include: • **Y:** generate and output distribution files [RUNID]_distindex.sas7bdat and RUNID]_distindexmax.sas7bdat • **N:** do not generate or output distribution files [RUNID]_distindex.sas7bdat and RUNID]_distindexmax.sas7bdat **Named by:** Request programmer **Input type:** Optional **Format:** SAS character $1 **Example:** DISTINDEX = Y
Treatment Pathways	TREATMENTPATHWAYS	**Details:** name of the SAS dataset used to evaluate and characterize switch pattern. **Note 1:** Relevant for Type 6 only **Named by:** Request Programmer **Input type:** Required **Format:** .sas7bdat file format **Example:** TREATMENTPATHWAYS = treatmentpathways
User-defined Strata Levels File	USERSTRATA	**Details:** name of the SAS dataset listing user-defined strata levels to include in output. **Named by:** Request programmer **Input type:** Required **Format: .**sas7bdat file format **Example:** USERSTRATA =userdefstrata
Overlap file	OVERLAPFILE	**Details:** name of the SAS dataset defining parameters to assess overlap analysis. **Note 1:** Relevant only for Type 2 analysis.

Parameter	Field Name	Description
		Named by: Request programmer **Input type:** Optional **Format: .**sas7bdat file format **Example:** OVERLAPFILE= overlap_wp006
Overlap Adherence File	OVERLAPFILE_ADHERE	**Details:** name of the SAS dataset defining treatment adherence for an overlap analysis. **Note 1:** Relevant only for Type 2 analysis. **Named by:** Request programmer **Input type:** Optional **Format: .**sas7bdat file format **Example:** OVERLAPFILE_ADHERE = ADHERENCEFILE
Concomitant file	CONCFILE	**Details:** name of the SAS dataset defining parameters to assess concomitant use. **Note 1:** Relevant only for Type 2 analysis. **Named by:** Request programmer **Input type:** Optional **Format: .**sas7bdat file format **Example:** CONCFILE= conc_wp006
Mutiple events file	MULTEVENTFILE	**Details:** name of the SAS dataset defining parameters to assess multiple events during requestor defined observation window. **Note 1:** Relevant only for Type 2 analysis. **Named by:** Request programmer **Input type:** Optional **Format: .**sas7bdat file format **Example:** MULTEVENTFILE = mevfile_wp006
Multiple Events Adherence File	MULTEVENTFILE_ADHERE	**Details:** name of the SAS dataset defining treatment adherence for a multiple events analysis. **Note 1:** Relevant only for Type 2 analysis. **Named by:** Request programmer **Input type:** Optional **Format: .**sas7bdat file format **Example:** MULTEVENTFILE_ADHERE = ADHERENCEFILE

3. Input Files

The CIDA tool allows requesters to specify multiple scenarios (or, in other words, define multiple cohorts) within a *single execution* of the program. Each cohort is assigned a unique GROUP value in input files to differentiate cohorts.

There are some parameters that are allowed to vary within a single execution of the program, and some that are not. As noted above, main program parameters are fixed for a single execution of the program.

In addition, there are several input file parameters that may not vary within a single execution of the program. Where applicable this is noted for each input file described in this section.

a) Cohort File

The Cohort File is required. It is used to define enrollment and demographic requirements, select the type of cohort identification strategy for the request, and indicate if extraction should be restricted to individuals for whom medical records may be requested.

There are five cohort identification strategies that can be employed with the CIDA tool:

1. Extract information to calculate background rates: program identifies an event (exposure, outcome, condition) and calculates the rate of that event in the SDD.
2. Extract information on exposures and follow-up time: program identifies an exposure of interest, determines exposed time (either requester-defined number of days after treatment initiation or based on drug dispensing' days supply), and looks for the occurrence of an HOI during exposed time.
3. Extract information for a self-controlled risk interval design: program identifies an exposure of interest, identifies a risk and control window relative to the exposure date, and examines the occurrence of HOIs during the risk and control windows.
4. Extract information to define pregnancy episodes and concurrent medical product use: program identifies live birth deliveries, calculates pregnancy duration, identifies comparator episodes with no live births, and examines the use of medical products by trimester.
5. Extract information for medical product utilization: program identifies the "first valid" exposure episode (i.e., the first episode during the query period that meets cohort entry criteria) as the index date, and then includes all subsequent exposure episodes.

To extract information to calculate background rates, a "Type 1" analysis must be performed. This means that the TYPE1FILE must be created and included in the program package.

To extract information on exposures and follow-up time, a "Type 2" analysis must be performed. This means that the TYPE2FILE must be created and included in the program package.

To extract information for a self-controlled risk interval design, a "Type 3" analysis must be performed. This means that the TYPE3FILE must be created and included in the program package.

To extract information on pregnancy episodes and medical product use, a "Type 4" analysis must be performed. This means that the TYPE4FILE must be created and included in the program package.

To extract information on medical product utilization, a "Type 5" analysis must be performed. This means that the TYPE5FILE must be created and included in the program package.

To extract information on product utilization and switching patterns, a "Type 6" analysis must be performed. This means that the TYPE6FILE must be created and included in the program package.

Note that in a single execution of the CIDA tool, only a Type 1 (extract information to calculate background rates) *or* Type 2 (extract information on exposures and follow-up time) *or* Type 3 (extract information for a self-controlled risk interval design) *or* Type 4 (extract information to define pregnancy episodes and concurrent medical product use) *or* Type 5 (medical product utilization) *or* Type 6 (product utilization and switching) cohort identification strategy can be specified.

Table 18 contains detailed specifications for this file.

Table 18. COHORTFILE Specifications

Parameter	Field Name	Description
Name of Cohort (Scenario)	COHORTGRP	**Details:** standardized name used to differentiate cohorts. **Note 1**: multiple cohorts can be defined within the same Cohort File. In this case all cohorts are queried independently and results are reported separately and labeled using each COHORTGRP name specified. **Note 2:** COHORTGRP is the primary key linking cohorts across input files; COHORTGRP values must match (including case) GROUP in other input files. **Note 3**: GROUP values must remain consistent during the course of a surveillance activity. **Named by:** Request programmer **Input type:** Required **Format:** Alphanumeric; SAS character $40; no special characters (*e.g.,* commas, periods, hyphens, etc.) allowed, and underscores must be used to mark spaces. **Example:** Insulin
Coverage Type Requirement	COVERAGE	**Details:** indicates medical and drug coverage type requirements for the cohort. Valid values are: • **M:** only enrollment spans with at least medical coverage should be considered by the MP algorithm • **D:** only enrollment spans with at least drug coverage should be considered by the MP algorithm • **MD:** only enrollment spans with both medical and drug coverage should be considered by the MP algorithm (**default value**) **Note 1**: Users must specify multiple groups if different COVERAGE requirements are needed. **Note 2:** the type of coverage specified is used when creating continuous enrollment periods and assessing cohort eligibility requirements. **Note 3:** if the COVERAGE value is left blank, or contains invalid values (*i.e.,* values other than "M", "D", or "MD"), the MP algorithm will consider only enrollment spans with both medical and drug coverage by default.

Parameter	Field Name	Description
		Defined by: Requester **Input type:** Optional (default value is **MD**) **Format:** SAS character $2 **Example**: MD
Enrollment Gap	ENROLGAP	**Details:** sets the number of days that will be bridged between two consecutive enrollment periods to create a "continuously enrolled" period. For example, if ENROLGAP=30 and a member is eligible for medical and drug coverage in periods 1/1/2007-3/27/2007 and 4/1/2007-12/21/2007 (*i.e.,* a 4-day gap between two consecutive enrollment episodes), the member will be considered continuously enrolled from 1/1/2007 to 12/21/2007. Any gaps in enrollment greater than 30 days will result in a new enrollment period, and all the days in the gap will be considered un-enrolled. **Note 1**: a gap of 45-days is recommended for most uses. **Note 2:** multiple continuous enrollment periods per member may be assessed. **Defined by:** Requester **Input type:** Required **Format**: Numeric **Example**: ENROLGAP=45 (gaps less than or equal to 45 days will be "bridged" to form one "continuously enrolled" sequence)
Minimum Pre-Index Enrollment Days	ENRDAYS	**Details:** optional parameter to specify the number of days of continuous enrollment required before the index date. **Note 1**: if not specified, a default value of 0 days is used. **Note 2:** this parameter allows requesters to specify enrollment criteria that *is greater* in duration than any washout periods or exclusion criteria specified. The value of ENRDAYS is only binding if: • ENRDAYS> T1WASHPER (in Type 1 File, if specified) and • ENRDAYS > T2WASHPER (in Type 2 File, if specified) and • ENRDAYS > T2FUPWASHPER (in Type 2 File, if specified) and • ENRDAYS> T3WASHPER (in Type 3 File, if specified) and

Parameter	Field Name	Description
		• ENRDAYS> T3FUPWASHPER (in Type 3 File, if specified and control window is after exposure) and • ENRDAYS> |T3CTRLFROM- T3FUPWASHPER| (in Type 3 File, if specified and control window is before exposure) and ENRDAYS > CONDFROM| (in Inclusion/Exclusion Codes File when INCLUSION=0, if specified) • ENRDAYS>T4WASHPER (in Type 4 File, if specified) and • ENRDAYS>T4FUPWASHPER (in Type 4 File, if specified) and • ENRDAYS>T5WASHPER (in Type 5 File, if specified) • ENRDAYS>T6WASHPER (in Type 6 File, if specified) The program will automatically use the longest duration before index date specified in the above parameters to assess continuous enrollment requirements. **Note 3:** in the pregnancy episodes cohort identification strategy (Type 4 analysis), ENRDAYS is assessed in relation to delivery date. In order to capture the entire pregnancy period, ENRDAYS should be at least 294. **Defined by:** Requester **Input type:** Optional (default value is 0) **Format:** Numeric **Example:** 365
Minimum Post-Index Enrollment Days	ENRDAYSAFTIND	**Details:** optional parameter to specify the number of days of continuous enrollment required after the index date. Note that the program does not, by default, require post-index date enrollment if the assessment period for covariates, exclusion criteria, most frequent utilization analyses, or high dimensional propensity score calculation extend beyond the index date. If enrollment is required, ENRDAYSAFT must be specified for the appropriate duration. **Note 1:** may be left blank if no post-index enrollment is required. **Defined by:** Requester **Input type:** Optional

Parameter	Field Name	Description
		Format: Numeric **Example:** 183
Type 1 Cohort Identification Strategy Indicator	TYPE1	**Details:** indicates if background rate cohort identification should be performed (see TYPE1FILE specifications). Allowable values are "Y" and "N". **Note 1:** If TYPE1=Y, a TYPE1FILE must be specified and included in the program package. **Note 2:** only TYPE1 *or* TYPE2 *or* TYPE3 *or* TYPE4 *or* TYPE5 *or* Type6 can have a value of "Y". **Defined by:** Requester **Input type:** Required **Format:** Char (1) **Example:** N
Type 2 Cohort Identification Strategy Indicator	TYPE2	**Details:** indicates if an exposures and follow-up time cohort identification should be performed (see TYPE2FILE specifications). Allowable values are "Y" and "N". **Note 1:** If TYPE2=Y, a TYPE2FILE must be specified and included in the program package. **Note 2:** only TYPE1 *or* TYPE2 *or* TYPE3 *or* TYPE4 *or* TYPE5 *or* TYPE6 can have a value of "Y". **Defined by:** Requester **Input type:** Required **Format:** Char (1) **Example:** Y
Type 3 Cohort Identification Strategy Indicator	TYPE3	**Details:** indicates if self-controlled risk interval design cohort identification should be performed (see TYPE3FILE specifications). Allowable values are "Y" and "N". **Note 1:** If TYPE3=Y, a TYPE3FILE must be specified and included in the program package. **Note 2:** only TYPE1 *or* TYPE2 *or* TYPE3 *or* TYPE4 *or* TYPE5 *or* TYPE6 can have a value of "Y". **Defined by:** Requester **Input type:** Required **Format:** Char (1) **Example:** Y
Type 4 Cohort Identification Strategy Indicator	TYPE4	**Details:** indicates if pregnancy episodes cohort identification strategy should be performed (see TYPE4FILE specifications). Allowable values are "Y" and "N".

Parameter	Field Name	Description
		Note 1: If TYPE4=Y, a TYPE4FILE must be specified and included in the program package. **Note 2:** only TYPE1 *or* TYPE2 *or* TYPE3 *or* TYPE4 *or* TYPE5 *or* TYPE6 can have a value of "Y". **Defined by:** Requester **Input type:** Required **Format:** Char (1) **Example:** Y
Type 5 Cohort Identification Strategy Indicator	TYPE5	**Details:** indicates if a medical product utilization cohort identification strategy should be performed. Allowable values are "Y" and "N". **Note 1:** If TYPE5=Y, a TYPE5FILE must be specified and included in the program package. **Note 2:** only TYPE1 *or* TYPE2 *or* TYPE3 *or* TYPE4 *or* TYPE5 *or* TYPE6 can have a value of "Y". **Defined by:** Requester **Input type:** Required **Format:** Char (1) **Example:** Y
Type 6 Cohort Identification Strategy Indicator	TYPE6	**Details:** indicates if switching analysis should be performed (see TYPE6FILE specifications). Allowable values are "Y" and "N". **Note 1:** If TYPE6=Y, a TYPE6FILE must be specified and included in the program package. **Note 2:** only TYPE1 *or* TYPE2 *or* TYPE3 *or* TYPE4 *or* TYPE5 *or* TYPE6 can have a value of "Y". **Defined by:** Requester **Input type:** Required **Format:** Char (1) **Example:** Y
Chart Availability Restriction Indicator	CHARTRES	**Details:** indicates if extraction should *exclude* members for whom medical charts cannot be requested for the entire study period. Allowable values are "Y" and "N". **Note 1**: If CHARTRES= "Y" the program will *exclude* individuals with at least one enrollment span with the SCDM variable Chart=N during the study period. **Defined by:** Requester **Input type:** Required **Format:** Char (1) **Example:** N (default)

Parameter	Field Name	Description
Sex criteria to apply to cohort	Sex	**Details:** optional parameter to restrict cohort to only specified Sex values. Blank will ensure that all Sex values are included in analyses. **Note 1:** valid values will be in single quotes and separated by a space. Valid values are: • **A**: ambiguous • **F**: female • **M**: male • **U**: unknown **Note 2:** restriction by Sex values does not ensure that matching is performed within values of Sex **Named by:** Requester **Input type:** Required **Format:** Char (3) **Example:** 'F' 'M' 'A' 'U'
Race criteria to apply to cohort	Race	**Details:** optional parameter to restrict cohort to only specified Race values. Blank will ensure that all Race values are included in analyses. **Note 1:** valid values will be in single quotes and separated by a space. Valid values are: • **0**: Unknown • **1**: American indian or Alaska Native • **2**: Asian • **3**: Black or African American • **4**: Native Hawaiian or Other Pacific Islander • **5**: White **Note 2:** restriction by Race values does not ensure that matching is performed within values of Race **Named by:** Requester **Input type:** Required **Format:** Char (1) **Example**: '3'
Hispanic criteria to apply to cohort	Hispanic	**Details:** optional parameter to restrict cohort to only specified Hispanic values. Blank will ensure that all Hispanic values are included in analyses. **Note 1:** valid values will be in single quotes and separated by a space. Valid values are: • **N**: no • **U**: unknown • **Y**: yes

Parameter	Field Name	Description
		Note 2: restriction by Hispanic values does not ensure that matching is performed within values of Hispanic **Named by:** Requester **Input type:** Required **Format:** Char (1) **Example:** 'N'
Age Groups	AGESTRAT	**Details:** age group categories for reporting. Specifying this parameter will (1) restrict to certain age groups and (2) specify how age groups will be stratified in result tables. For example, to have results stratified by 20 year increments for members 40-99 years of age, enter AGESTRAT=40-59 60-79 80-99. **Note 1:** For Type 1, Type 2, Type 3, Type 5, and Type 6 analyses, age is calculated at index date. **Note 2:** For Type 4 analysis, age is calculated at delivery date. **Note 3:** various units of time can be used. Valid values are: • **D**: days • **W**: weeks • **Q**: quarters • **M**: months • **Y**: years (default value) **Note 4:** lower value is binding. If AGESTRAT=0-5 5-10, then all 5 year olds will be placed in the second age group. If AGESTRAT=0-5 6-10, then all 5 year olds will be placed in the first age group. For example, to have results stratified by 6 month increments for the first two years of life and then by 2 year increments until the age of 6, AGESTRAT = 00M-05M 06M-11M 12M-17M 18M-23M 02Y-03Y 04Y-05Y needs to be entered. **Note 5**: using an open ended age category (e.g., 85+) imposes an age ceiling of 110 years. If age >110 is desired, the final age category ceiling must be specified (e.g., 85-125). **Note 6:** age groups must be mutually exclusive (i.e., non overlapping). **Note 7:** When constructing age categories that only include one age, the lower and upper values are equal. For example, 00M-<01M, 01M-<02M, 02M-

Parameter	Field Name	Description
		<03M, should be specified as 00M-00M 01M-01M 02M-02M **Note 8:** For PSA, age groups should be the same for both the exposure and control groups. **Defined by:** Requester **Input type:** Optional (default value is **00-01 02-04 05-09 10-14 15-18 19-21 22-44 45-64 65-74 75+ in years**) **Format:** Char (100) **Example:** AGESTRAT=**40-59 60-79 80-99**

b) Type 1 File

The Type 1 File is optional and its specification is only required for a background rate calculation cohort identification strategy. Options include selecting the number of events an individual can contribute to the request, the number of days before index date to assess incidence criteria, whether to truncate enrollment at death date, and whether to output a table characterizing reason for censoring eligibility Table 19 contains detailed specifications for this file.

Table 19. TYPE1FILE Specification

Parameter	Field Name	Description
Name of Cohort	GROUP	**Details:** standardized name used to differentiate cohorts. **Note 1**: multiple cohorts can be defined within the same Type 1 File. In this case all cohorts are queried independently and results are reported separately and labeled using each GROUP name specified. **Note 2:** GROUP is the primary key linking cohorts across input files; GROUP values must match (including case) between the TYPE1FILE and other input files. **Named by:** Request programmer **Input type:** Required **Format:** Alphanumeric; SAS character $40; no special characters (*e.g.*, commas, periods, hyphens, etc.) allowed, and underscores must be used to mark spaces. **Example:** Insulin
Allowed Number of Index Dates per Individual	T1COHORTDEF	**Details:** indicates how many index dates an individual can contribute. Options include: • **01:** Cohort includes only the first valid index date per individual during the query period. • **02:** Cohort includes all valid index dates per individual during the query period.

Parameter	Field Name	Description
		Note 1: T1COHORTDEF parameter is used in conjunction with the T1WASHPER variable (below) to define valid index date(s). **Named by:** Requester **Input type:** Required **Format:** Alphanumeric; SAS character $2 **Example:** 01
Type 1 Index Washout Period	T1WASHPER	**Details:** length of washout period in days. The washout period is a period before an index date during which an individual cannot have evidence of incidence-defining criteria (see Cohort Codes File specification for additional details on incidence-defining criteria). **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** 365
Censor Enrollment at Evidence of Death	CENSOR_DTH	**Details:** indicates if enrollment should be censored based on death date. Allowable values are "Y" and "N". Date of death can be determined two ways: 1. Using discharge status = expired in the SDD Encounter table. Death date is set to discharge date. OR 2. Using death date in the SDD Death table for records with Confidence=Excellent. **Note 1:** censoring is implemented by restricting enrollment eligibility. Member eligibility is truncated at death date. Once a death date is observed, a member can no longer contribute eligible periods (even if they are observed in the data). **Defined by:** Requester **Input type:** Required **Format:** SAS character $1 **Example:** Y
Censor Enrollment at DP Data End Date	CENSOR_ DPEND	**Details:** indicates if enrollment should be censored based on DP data end date. Allowable values are "Y" and "N". **Note 1:** when CENSOR_DPEND = "Y", the program will adjust the query period end date (QUERYTO) to reflect DP_MaxDate.

Parameter	Field Name	Description
		Defined by: Requester **Input type:** Required **Format:** SAS character $1 **Example:** Y
Censor Enrollment at Query End Date	CENSOR_QRYEND	**Details:** indicates if enrollment should be censored based on query data end date. Allowable values are "Y" and "N". **Defined by:** Requester **Input type:** Required **Format:** SAS character $1 **Example:** Y
Categories for Follow-up Time	CENSOR_ OUTPUT_CAT	**Details:** indicates ranges (in days) for stratification variable CENSDAYS_VALUE in [RUNID]_censor_CIDA.sas7bdat output. **Note 1:** leave blank if only continuous values of CENSDAYS_VALUE are desired. If this field is left blank, output stratified by CENSDAYS_VALUE in [RUNID]_censor_CIDA will have one category that includes all values of CENSDAYS_VALUE. **Defined by:** Requester **Input type:** Optional **Format:** Alphanumeric **Example:** 0-364 365-729 730-1094 1095+

c) Type 2 File

The Type 2 File is optional and its specification is only required for an exposures and follow-up time cohort identification strategy.

In an exposures and follow-up time strategy, requesters can create exposure episodes based on outpatient pharmacy dispensing days supplied. The exposure episode in this case would be defined as a dispensing sequence that has no interruption in days supplied greater than a requester-defined allowable "gap". The allowable gap is the number of days used to bridge dispensings to create a continuous exposure episode.

Alternatively, exposure episodes can be created based on a requester-defined number of days after the exposure. The parameters that need to be defined in the Type 2 File are dependent on the exposure episode creation method.

Note that if exposure episodes are created based on outpatient pharmacy dispensing days supplied and a procedure, diagnosis, or laboratory result code is included in the definition of exposure, the CIDA tool will assign the code a default value of 1 day of supply. Table 20 contains detailed specifications for this file.

Table 20. TYPE2FILE Specification

Parameter	Field Name	Description
Name of Cohort	GROUP	**Details:** standardized name used to differentiate cohorts. **Note 1**: multiple cohorts can be defined within the same Type 2 File. In this case all cohorts are queried independently and results are reported separately and labeled using each GROUP name specified. **Note 2:** GROUP is the primary key linking cohorts across input files; GROUP values must match (including case) between the TYPE2FILE and other input files. **Note 3**: GROUP values must remain consistent during the course of a surveillance activity. **Named by:** Request programmer **Input type:** Required **Format:** Alphanumeric; SAS character $40; no special characters (*e.g.,* commas, periods, hyphens, etc.) allowed, and underscores must be used to mark spaces. **Example:** Insulin
Allowed Number of Exposure Episodes per Individual	T2COHORTDEF	**Details:** indicates how many exposure periods an individual can contribute. Options include: • **01:** Cohort includes only the first valid exposure episode during the query period • **02:** Cohort includes all valid exposure episodes during the query period

Parameter	Field Name	Description
		• **03:** Cohort includes all valid exposure episodes during the query period until an outcome of interest occurs **Note 1:** T2COHORTDEF parameter is used in conjunction with the T2WASHPER parameter (below) to define valid exposure episode(s). **Note 2:** T2COHORTDEF must equal "01" for requests that will use the propensity score matching module. **Defined by:** Requester **Input type:** Required **Format:** Alphanumeric; SAS character $2 **Example:** 01
Type 2 Exposure Washout Period	T2WASHPER	**Details:** length of washout period in days. The washout period is a period before an exposure episode during which an individual cannot have evidence of incidence-defining criteria (see Cohort Codes File specification for additional details on incidence-defining criteria). **Note 1:** the MP algorithm may use days before the query start date to determine if continuous enrollment and incidence criteria are met. **Note 2**: **special case**: when T2WASHPER = missing the program requires ENRDAYS of continuous enrollment but only considers an exposure episode valid if, at index date, the member has no evidence of the exposure in their entire available enrollment history. **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** 365
Requester-defined Exposure Episode Length	ITTDAYS	**Details:** number of days after exposure initiation that is considered "exposed time." **Note 1:** if exposure episode will be determined by outpatient pharmacy dispensing days supplied, leave this field blank. **Note 2:** exposure episodes will be censored at the first occurrence of the following: 1) end of enrollment; 2) occurrence of HOI; 3) occurrence of any additional requester-defined censoring criteria. **Note 3:** if this field is populated, the following fields should be left blank in this file: EPISODEGAP,

Parameter	Field Name	Description
		EXPEXTPER, and MINEPISDUR (these parameters are for exposed time that is determined by outpatient pharmacy dispensing days supplied). **Defined by:** Requester **Input type:** Optional **Format:** Numeric **Example:** 30
Treatment Episode Gap Type	EPISODEGAPTYPE	**Details:** specifies the type of algorithm to use for the calculation of episode gaps. Valid values are: • **F**: Fixed episode gap. The value specified in EPISODEGAP will be used to determine if two consecutive claims are in the same episode. • **P**: Percentage episode gap. The value specified in EPISODEGAP will represent a percentage of the previous dispensing's days of supply to determine if two consecutive claims are in the same episode. **Note 1:** this value must be the same within a given query GROUP. **Note 2:** default value is "P." **Defined by:** Request Programmer **Input type:** Required (default value is P) **Format:** Alphanumeric; SAS character $1. **Example:** P
Exposure Episode Gap	EPISODEGAP	**Details:** used in conjunction with EPISODEGAPTYPE; sets the number of days allowed between two consecutive claims to consider them as part of the same treatment episode. For a given claim, if EPISODEGAPTYPE is Fixed (F), a gap of more than EPISODEGAP days between the claim date and the date of last day of supply of the previous dispensing triggers a new exposure episode. If EPISODEGAPTYPE is Percentage (P), a gap of more than EPISODEGAP percent of the previous claim days of supply days between the claim date and the date of the last day of supply of the previous claim triggers a new treatment episode. For example, if EPISODEGAP=10 and EPISODEGAPTYPE=F, claim 1's last day of supply is on 1/31/2012 and claim 2's start date is 2/12/2012, the MP algorithm starts a new treatment episode on

Parameter	Field Name	Description
		2/12/2012 because there are more than 10 days between the two claims. **Note 1:** this value must be the same within a given query GROUP. **Note 2:** only relevant for requests creating exposure episodes based on dispensing days supplied. If exposure episode length is defined by the requester using the ITTDAYS parameter (*i.e.,* using a requester-defined number of days after exposure initiation), leave this field blank. **Note 3**: gap days bridged are included in the days at risk metrics. **Note 4:** gaps are assessed and bridged before the application of exposure episode extensions (EXPEXTPER; below). **Defined by:** Requester **Input type:** Required (0 must be entered if no EPISODEGAP is required; default value is 0 for Fixed EPISODEGAPTYPE and 30 for Percentage EPISODEGAPTYPE) **Format:** Numeric **Example:** 0
Exposure Episode Extension Period	EXPEXTPER	**Details:** extends the length of an exposure episode by specified number of days. An exposure episode can be extended EXPEXTPER days after the last day of supply of the treatment episode's last dispensing. **Note 1:** only relevant for requests creating exposure episodes based on dispensing days supplied. If exposure episode length is defined by the requester using the ITTDAYS parameter (*i.e.,* using a requester-defined number of days after exposure initiation), leave this field blank. **Note 2:** extension days are added after the stockpiling algorithm has been applied and exposure episodes are created. **Note 3:** extensions days are added after any episode gaps have been bridged (see EPISODEGAP parameter). **Note 4:** extension days are included in days at-risk metrics. **Defined by:** Requester

Parameter	Field Name	Description
		Input type: Required for requests that are creating exposure episodes using dispensing days supply **Format:** Numeric **Example:** 0
Minimum Exposure Episode Duration	MINEPISDUR	**Details:** rejects exposure episodes of fewer than MINEPISDUR days. **Note 1:** only relevant for requests creating exposure episodes based on dispensing days supplied. If exposure episode length is defined by the requester using the ITTDAYS parameter (*i.e.,* using a requester-defined number of days after exposure initiation), leave this field blank. **Note 2:** criterion applied after any gaps are bridged and extension days added to the length of the exposure episode. **Note 3:** MINEPISDUR will count any gap days that bridge episodes. This is how the parameter differs from MINDAYSUPP. **Note 4**: minimum episode duration criteria are assessed prior to censoring due to the occurrence of an HOI (e.g., if an episode is 45 days, and an HOI occurs on day 15, a MINEPISDUR value of 30 days will capture both the exposure episode and the HOI [the episode will *not* be excluded due to days at risk censoring on day 15]). **Defined by:** Requester **Input type:** Required for requests that are creating exposure episodes using dispensing days supply **Format:** Numeric **Example:** 0
Maximum Exposure Episode Duration	MAXEPISDUR	**Details:** censors/truncates exposure episodes after a requester-specified number of exposed days. **Note 1:** only relevant for requests creating exposure episodes based on dispensing days supplied. If exposure episode length is defined by the requester using the ITTDAYS parameter (*i.e.,* using a requester-defined number of days after exposure initiation), leave this field blank. **Note 2:** criterion applied after any gaps are bridged and extension days added to the length of the exposure episode.

Parameter	Field Name	Description
		Note 3: criterion applied before any specified blackout period is considered. For example, if there is a 1-day blackout, and MaxEpisDur = 20, then maximum TTE would be 19. **Defined by:** Requester **Input type:** Optional **Format:** Numeric **Example:** 60
Minimum Days Supplied	MINDAYSUPP	**Details:** rejects exposure episodes where less than MINDAYSUPP days supplied were used to create the exposure episode. **Note 1:** MINDAYSUPP evaluates dispensings days supply in treatment episode, and does not count any gap days that bridge episodes. This is how the parameter differs from MINEPISDUR. **Note 2:** when a Type 2 analysis is run, and the cohort is defined using age anniversary or calendar date, then mindaysupp should be set to 0. **Defined by:** Requester **Input type:** Required (0 must be entered if no MINDAYSUPP is required) **Format:** Numeric **Example:** 0
Censor Episodes at Evidence of Death	CENSOR_DTH	**Details:** indicates if a treatment episode should be censored based on death date. Allowable values are "Y" and "N". Date of death can be determined two ways: 1. Using discharge status = expired in the SDD Encounter table. Death date is set to discharge date. OR 2. Using death date in the SDD Death table for records with Confidence=Excellent. **Note 1:** censoring is implemented by restricting enrollment eligibility. Member eligibility is truncated at death date. Once a death date is observed, a member can no longer contribute eligible periods (even if they are observed in the data). **Defined by:** Requester **Input type:** Required **Format:** SAS character $1 **Example:** Y

Parameter	Field Name	Description
Type 2 HOI Washout Period	T2FUPWASHPER	**Details:** The washout period is a period before an exposure episode during which an individual cannot have evidence of the HOI. **Note 1**: the HOI washout period looks back from the *exposure episode* index date. **Note 2:** the MP algorithm may use days before the query start date to determine if continuous enrollment and incidence criteria are met. **Note 3**: **special case**: when T2FUPWASHPER = missing the program requires ENRDAYS of continuous enrollment but only considers an exposure episode valid if, at index date, the member has no evidence of an HOI in their entire available enrollment history. **Named by:** Requester **Input type:** Required (0 must be entered if no WASHPER is required) **Format:** Numeric **Example:** 365
HOI Characterization De-duplication Process	EVENTCOUNT	**Details:** by design, individuals stop contributing days at risk during an exposure episode when an HOI occurs. HOIs/Days at-risk metrics reported allow individuals to contribute, at most, one HOI per episode. However, the MP algorithm is able to characterize the number of total HOIs observed during valid treatment episodes. Requesters can use this field to determine how this characterization should count the number of HOIs. Again, this is for characterization only, and will not affect HOI/Days at-risk metrics. Valid values are: • **0:** counts all occurrences of an HOI during an exposure episode. • **1:** de-duplicates occurrences of the *same HOI code and code type* on the same day (*i.e.,* de-duplicates at the exact match code level). Note: a patient may have the same HOI code and code type on the same day if they were recorded by different providers and/or occurred in different care settings.

Parameter	Field Name	Description
		• **2:** de-duplicates occurrences of the *same HOI GROUP* on the same day (*e.g.,* de-duplicates at the GROUP level). Consider the example where the HOI is defined with ICD-9-CM diagnosis codes 250.01 and 250.11 in any care setting. A member has an occurrence of code=250.01 on two separate AV records and of code=250.11 on another AV record on the same date during his/her incident treatment episode. EVENTCOUNT=0 will identify three HOIs. EVENTCOUNT=1 will identify two HOIs. EVENTCOUNT=2 will identify one HOI. **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** 2
HOI Blackout Period	BLACKOUTPER	**Details:** the HOI blackout period in days. The requester can specify a period at the start of an exposure episode during which HOIs found by the MP algorithm are ignored. That is, the at-risk period starts at the end of the blackout period. Moreover, if an HOI occurs during the blackout period, the exposure episode will not be considered incident with respect to the HOI (and thus excluded from output metrics). **Note 1**: this allows a requester to exclude "same-day" HOIs by setting the BLACKOUTPER to 1. **Defined by:** Requester **Input type:** Required (0 must be entered if no BLACKOUTPER is required) **Format:** Numeric **Example:** 7
Categories for Follow-up Time	CENSOR_ OUTPUT_CAT	**Details:** indicates ranges (in days) for stratification variable CENSDAYS_VALUE in [RUNID]_censor_CIDA.sas7bdat output. **Note 1:** This should be populated if stratification by censdays_value_cat is requested. Leave blank if continuous values of CENSDAYS_VALUE are desired. **Defined by:** Requester **Input type:** Optional **Format:** Alphanumeric **Example:** 0-364 365-729 730-1094 1095+

Parameter	Field Name	Description
Censor Episodes at DP Data End Date	CENSOR_ DPEND	**Details:** indicates if a treatment episode should be censored based on DP data end date. Allowable values are "Y" and "N". **Note 1:** when CENSOR_DPEND = "Y", the program will adjust the query period end date (QUERYTO) to reflect DP_MaxDate. **Note 2:** CENSOR_DPEND must = "N" for requests that will use the prospective surveillance with propensity score matched design. **Defined by:** Requester **Input type:** Required **Format:** SAS character $1 **Example:** Y
Censor Episodes at Query End Date	CENSOR_QRYEND	**Details:** indicates if a treatment episode should be censored based on query data end date. Allowable values are "Y" and "N". **Defined by:** Requester **Input type:** Required **Format:** SAS character $1 **Example:** Y
Instructions to create never cohort	NEVEREXPOSEDCOHORT	**Details:** Determines whether to create a never-exposed cohort. Valid values are: • **Y:** Create never-exposed cohort • **N:** Do not create never-exposed cohort **Note 1:** Never-exposed cohort creation is not available when conducting multi-look surveillance **Named by:** Requester **Input type:** Required **Format:** SAS character $1 **Example:** Y
Minimum Post-Episode Enrollment Days For Type 2 Analyses	ENRDAYSAFTEPI	**Details:** optional parameter to specify the number of days of continuous enrollment required after the episode end date for Type 2 analyses. **Note 1:** may be left blank if no post-episode enrollment is required. **Named by:** Requester **Input type:** Optional **Format:** Numeric **Example:** 90

d) Type 3 File

The Type 3 File is optional and its specification is only required for a self-controlled risk interval design cohort identification strategy. In a self-controlled risk interval design strategy, requesters define an exposure of interest, specify a risk and control window relative to the exposure date, and examine the occurrence of HOIs during the risk and control windows. Table 21 contains detailed specifications for this file.

Table 21. TYPE3FILE Specification

Parameter	Field Name	Description
Name of Cohort	GROUP	**Details:** standardized name used to differentiate cohorts. **Note 1**: multiple cohorts can be defined within the same Type 3 File. In this case all cohorts are queried independently and results are reported separately and labeled using each GROUP name specified. **Note 2:** GROUP is the primary key linking cohorts across input files; GROUP values must match (including case) between the TYPE3FILE and other input files. **Note 3**: GROUP values must remain consistent during the course of a surveillance activity. **Named by:** Request programmer **Input type:** Required **Format:** SAS character $40; no special characters (*e.g.,* commas, periods, hyphens, etc.) allowed, and underscores must be used to mark spaces. **Example:** Insulin
Risk Window Interval Start	T3RISKFROM	**Details:** indicates the risk window interval start, as days relative to the exposure start date (i.e., day zero). **Note 1:** a T3RISKFROM value of 3 indicates that the risk window should start three days after the exposure date. **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** 3
Risk Window Interval End	T3RISKTO	**Details:** indicates the risk window end, as days relative to the exposure start date (i.e., day zero). **Note 1:** A T3RISKTO value of 18 indicates that the risk window should end 18 days after the exposure date. **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** 18
Control Window Interval Start	T3CTRLFROM	**Details:** indicates the control window start, as days relative to the exposure start date (i.e., day zero).

Parameter	Field Name	Description
		Note 1: A T3CTRLFROM value of 19 indicates that the risk window should start 19 days after the exposure date. **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** 19
Control Window Interval End	T3CTRLTO	**Details:** indicates the control window end, as days relative to the exposure start date (i.e., day zero). **Note 1:** A T3CTRLTO value of 25 indicates that the risk window should end 25 days after the exposure date. **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** 25
Allowed Number of Exposure Episodes per Individual for the Surveillance Activity	T3COHORTDEF	**Details:** indicates how many exposures ("day zeros") an individual can contribute to the query or the surveillance activity (and, by extension, the number of risk/control interval pairs an individual can contribute). Options include: **01:** Cohort includes only the first exposure **02:** Cohort includes all exposures **Note 1:** for one time assessments, the T3COHORTDEF parameter determines the number of exposures that are identified during the query period. For surveillance activities, the T3COHORTDEF parameter determines the number of exposures that are identified during the duration of the surveillance activity. This has implications when T3COHORTDEF=01. For any evaluations after Look 1, the program will ensure that any exposures identified are the first observed in all prior exposure assessment periods. For example, if the Look 1 exposure assessment period is 10/15/2014-11/30/2014, and the Look 2 exposure assessment period is 12/1/2014-1/31/2014, in Look 2 the program will ensure that any exposures identified were the first that occurred since the surveillance start date (T3SURVSTARTDATE; 10/15/2014). **Note 2:** T3COHORTDEF parameter is used in conjunction with the T3WASHPER parameter (below) to define valid exposures. In order to ensure that multiple exposure risk and control windows do not overlap (when T3COHORTDEF = 02), the T3WASHPER value must be set accordingly:

Parameter	Field Name	Description
		Control window is after risk window: T3WASHPER must be *at least* T3CTRLTO + 1 days in duration. Control window is before exposure: T3WASHPER must be at least |T3CTRLFROM | + T3RISKTO + 1 days in duration. **Defined by:** Requester **Input type:** Required **Format:** Alphanumeric; SAS character $2 **Example:** 01
Type 3 Exposure Washout Period	T3WASHPER	**Details:** length of washout period in days. The washout period is a period before exposure during which an individual cannot have evidence of incidence-defining criteria (see Cohort Codes File specification for additional details on incidence-defining criteria). **Note 1:** the MP algorithm may use days before the surveillance start date and exposure assessment period to determine if continuous enrollment and incidence criteria are met. **Note 2**: **special case**: when T3WASHPER = missing the program requires the default pre-exposure continuous enrollment requirements, but only considers an exposure episode valid if, at index date, the member has no evidence of the exposure in their entire available enrollment history. **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** 365
Type 3 HOI Incidence Assessment Period	T3FUPWASHPER	**Details:** length of HOI incidence assessment period in days. The incidence assessment period is a period before an HOI during which an individual cannot have evidence of HOI incidence-defining criteria (see Cohort Codes File specification for additional details on incidence-defining criteria). **Note 1:** a Type 3 cohort identification strategy requires that a patient only contribute an HOI to the risk *or* the control window in the analytic cohort; a patient is not allowed to contribute an HOI to both windows in the analytic cohort. This requirement is ensured by enforcing a minimum HOI incidence assessment period. This minimum duration is calculated as: [Maximum (Risk interval end date, Control interval end date, Exposure date)] – [minimum (Control interval start date, Risk interval start date, Exposure date)] + 1 days in duration.

Parameter	Field Name	Description
		Requesters should ensure that T3FUPWASHPER is greater than or equal to the minimum HOI incidence assessment period. However, should T3FUPWASHPER be set less than the minimum HOI incidence assessment period (or set to zero), the program will override the value and ensure the minimum requirements. **Note 2**: **special case**: when T3FUPWASHPER = missing the program requires the default pre- and post-exposure continuous enrollment (see Enrollment Requirements section for additional details) but only considers an HOI valid if, at HOI date, the member has no evidence of the exposure in their entire available enrollment history. **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** 365
Censor Evaluation Windows at Evidence of Death	T3CENSOR_DTH	**Details:** indicates if risk and evaluation windows should be censored based on death date. Allowable values are "Y" and "N". Date of death can be determined two ways: 1. Using discharge status = expired in the SDD Encounter table. Death date is set to discharge date. OR 2. Using death date in the SDD Death table for records with Confidence=Excellent. **Note 1:** censoring is implemented by restricting enrollment eligibility. Member eligibility is truncated at death date. Once a death date is observed, a member can no longer contribute eligible periods (even if they are observed in the data). **Note 2**: output will differentiate censoring based on end of enrollment versus evidence of death. **Defined by:** Requester **Input type:** Required **Format:** Char (1) **Example:** Y
NDC same day exclusion	T3EXCLONSAMEDAY	**Details:** indicates if an exposure defined using NDCs should be excluded from consideration if more than one of the NDCs used to define the exposure is observed on the same day (i.e., the patient cannot have evidence of more than one of the NDCs used to define the exposure on day 0). Allowable values are 'Y' and 'N'.

Parameter	Field Name	Description
		Note 1: applies to NDCs only. For codes other than NDCs, set to 'N'. **Defined by:** Requester **Input type:** Required **Format:** Char (1) **Example:** Y
	T3SURVSTARTDATE	**Details:** surveillance start date/exposure identification period start date for time period 1. Can be unique per value of GROUP. This parameter's value is used to identify the start of the exposure identification period for time period 1 (but must be specified for all time periods). For subsequent time periods, the exposure identification period start date is determined by the prior time period's exposure identification period stop date + 1. **Note 1:** MONITORINGFILE is not required for TYPE3 analyses. **Defined by:** Requester **Input type:** Required **Format:** mm/dd/yyyy **Example**: 03/01/2012
	T3ENDOFUPDATE	**Details**: Data Partner data completeness date. Should be determined by the surveillance team with information available from the SOC Data Management and Quality Assurance (DMQA) team. **Note 1:** MONITORINGFILE is not required for TYPE3 analyses. **Note 2:** field cannot be left blank. **Defined by:** Requester **Input type:** Required **Format:** mm/dd/yyyy **Example**: 06/30/2012

e) Type 4 File

The Type 4 File is optional and its specification is only required for a pregnancy episodes identification strategy. Options include the number of days before index date to assess incidence criteria, the length of pregnancy in absence of preterm/postterm codes, and the allowable gap between delivery code and preterm/postterm code. Parameters without a suffix are used to define pregnancy episodes. Parameters with the suffix "2" are used to define medical product use.
Table 22 contains detailed specifications for this file.

Table 22. TYPE4FILE Specification

Parameter	Field Name	Description
Name of Cohort	GROUP	**Details:** standardized name used to differentiate cohorts. **Note 1**: multiple cohorts can be defined within the same Type 4 File. In this case all cohorts are queried independently and results are reported separately and labeled using each GROUP name specified. **Note 2:** GROUP is the primary key linking cohorts across input files; GROUP values must match (including case) between the TYPE2FILE and other input files. **Named by:** Request programmer **Input type:** Required **Format:** Alphanumeric; SAS character $40; no special characters (*e.g.,* commas, periods, hyphens, etc.) allowed, and underscores must be used to mark spaces. **Example:** Preg1
Allowed Number of Pregnancy Episodes per Individual	T4COHORTDEF	**Details:** indicates how many pregnancy episodes an individual can contribute. Options include: • **02:** Cohort includes all valid pregnancy episodes during the query period **Note 1:** T4COHORTDEF parameter is used in conjunction with the T4WASHPER parameter (below) to define valid exposure episode(s). **Defined by:** Requester **Input type:** Required **Format:** Alphanumeric; SAS character $2 **Example:** 02
Type 4 Exposure Washout Period	T4WASHPER	**Details:** length of washout period in days. The washout period is a period before a delivery during which an individual cannot have evidence of a prior delivery. **Note 1:** the washout is applied to delivery code and not index date (start of pregnancy episode). **Note 2: special case:** if T4WashPer < length of pregnancy, and the index date of a pregnancy episode

Parameter	Field Name	Description
		falls during a prior episode, the second episode is truncated to start 1 day after the prior episode. **Note 3:** the MP algorithm may use days before the query start date to determine if continuous enrollment and incidence criteria are met. **Note 4**: **special case**: when T4WASHPER = missing the program requires ENRDAYS of continuous enrollment but only considers a pregnancy episode valid if, at delivery date, the member has no evidence of pregnancy in their entire available enrollment history. **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** 365
Requester-defined Pregnancy Episode Length	ITTDAYS	**Details:** pregnancy duration in the absence of preterm/postterm codes within EPISODEGAP of delivery. **Defined by:** Requester **Input type:** Optional **Format:** Numeric **Example:** 273
Preterm/Postterm Code Evaluation Window	EPISODEGAP	**Details:** sets the number of days around the delivery date where the program looks for evidence of preterm/postterm codes to calculate the length of the pregnancy episode. For example, if EPISODEGAP=10, and a delivery is on 1/31/2012, the program will look in the 10 days before and 10 days after 1/31/2012 to evaluate presence of preterm/postterm codes. **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** 14
Type 4 HOI Washout Period	T4FUPWASHPER	**Details:** The washout period is a period before a pregnancy episode during which an individual cannot have evidence of the HOI. **Note 1**: Accessing HOI is not currently functional, so this should be left blank. Note that the parameter must be included (with a blank value) in the input file for the program to execute. **Named by:** Requester **Input type:** Optional **Format:** Numeric

Parameter	Field Name	Description
		Example: .
Minimum Pre-Delivery Enrollment Days for Type4 Analysis	ENRDAYSFLOOR	**Details:** indicates the minimum number of days of continuous enrollment required prior to the delivery date for Type 4 analysis only. **Note 1:** Enrollment requirements are assessed via the ENRDAYS parameter with respect to the delivery date and not the index date. To ensure medical product use is captured during the entire pregnancy period, this parameter triggers a custom warning to the log if ENRDAYS < ENRDAYSFLOOR. This parameter is only used to generate a warning message; it is not used to assess enrollment requirements. It should be set to 294 if the default MEPREP algorithm to determine pregnancy duration is used. **Note 2:** Must be left blank for Type 1, Type 2, Type 3, Type 5, and Type 6 analyses. Required for Type 4 analysis. **Named by:** Requester **Input type:** Optional **Format:** Numeric **Example:** 294
Type 4 Medical Product Cohort Definition	T4COHORTDEF2	**Details:** indicates how many medical product exposure periods an individual can contribute. Options include: • **01:** Cohort includes first valid medical product exposure episodes during a pregnancy episode • **02:** Cohort includes all valid medical product exposure episodes during a pregnancy episode • **99:** Cohort includes last valid medical product exposure episodes during a pregnancy episode **Note 1:** T4COHORTDEF2 parameter is used in conjunction with the T4WASHPER2 parameter (below) to define valid medical product exposure episode(s). **Defined by:** Requester **Input type:** Required **Format:** Alphanumeric; SAS character $2 **Example:** 02
Type 4 Medical Product Washout Period	T4WASHPER2	**Details:** length of washout period in days for medical products. The washout period is a period before start of medical product episode during which an individual cannot have evidence of a prior episode. **Note:** MOI episodes are incident to itself and any MOI incidence codes.

Parameter	Field Name	Description
		Defined by: Requester **Input type:** Required **Format:** Numeric **Example:** 0
Requester-defined Medical Product of Interest Episode Length	ITTDAYS2	**Details:** number of days after medical product use episode initiation that is considered "exposed time." **Note 1:** if exposure episode will be determined by outpatient pharmacy dispensing days supplied and forced code supply, leave this field blank. **Defined by:** Requester **Input type:** Optional **Format:** Numeric **Example:** .
Treatment Episode Gap Type	EPISODEGAPTYPE2	**Details:** specifies the type of algorithm to use for the calculation of episode gaps for medical product use. Valid values are: • **F**: Fixed episode gap. The value specified in EPISODEGAP2 will be used to determine if two consecutive claims are in the same episode. • **P**: Percentage episode gap. The value specified in EPISODEGAP2 will represent a percentage of the previous dispensing's days of supply to determine if two consecutive claims are in the same episode. **Note 1:** this value must be the same within a given query GROUP. **Note 2:** default value is "P." **Defined by:** Request Programmer **Input type:** Required (default value is P) **Format:** Alphanumeric; SAS character $1. **Example:** P
Exposure Episode Gap	EPISODEGAP2	**Details:** used in conjunction with EPISODEGAPTYPE2; sets the number of days allowed between two consecutive dispensings to consider them as part of the same medical product episode. For a given claim, if EPISODEGAPTYPE2 is Fixed (F), a gap of more than EPISODEGAP2 days between the dispensing date and the date of the last day of supply of the previous dispensing triggers a new medical product episode. If EPISODEGAPTYPE2 is Percentage (P), a gap of more than EPISODEGAP2 percent of the previous dispensing days of supply days between the dispensing date and the date of

Parameter	Field Name	Description
		the last day of supply of the previous dispensing triggers a new treatment episode. For example, if EPISODEGAP2=10 and EPISODEGAPTYPE2=F, claim1's last day of supply is on 1/31/2012 and claim2's start date is 2/12/2012, the MP algorithm starts a new medical product episode on 2/12/2012 because there are more than 10 days between the two dispensings. **Note 1:** this value must be the same within a given query GROUP. **Note 2:** only relevant for requests creating exposure episodes based on dispensing days supplied. If exposure episode length is defined by the requester using the ITTDAYS2 parameter (*i.e.,* using a requester-defined number of days after exposure initiation), leave this field blank. **Defined by:** Requester **Input type:** Required (0 must be entered if no EPISODEGAP2 is required; default value is 0 for Fixed EPISODEGAPTYPE2 and 30 for Percentage EPISODEGAPTYPE2) **Format:** Numeric **Example:** 10
Type 4 Medical Product HOI Washout Period	T4FUPWASHPER2	**Details:** The washout period is a period before a medical product episode during which an individual cannot have evidence of the HOI. **Note 1**: Accessing HOI is not currently functional, so this should be left blank. Note that the parameter must be included (with a blank value) in the input file for the program to execute. **Named by:** Requester **Input type:** Optional **Format:** Numeric **Example:** .
Concurrent Washout Period	ConcWashPer	**Details:** the number of days prior to a pregnancy episode that an individual must be free of a medical product episode in order for the medical product to be included in the "trimester only" statistics. **Note 1**: Accessing concurrent washout is not currently functional, so this should be left blank. Note that the parameter must be included (with a blank value) in the input file for the program to execute.

Parameter	Field Name	Description
		Defined by: Requester **Input type:** Required **Format:** Numeric **Example:** .
Type 4 Medical Product Removal Start Day	REMOVEMOISTART	**Details:** number of days, relative to pregnancy start date, to start interval where MOIs will not be counted. **Note:** concept does not apply to as-treated episodes or episodes where ITTDays2 > 1 **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** -90
Type 4 Medical Product Removal End Day	REMOVEMOIEND	**Details:** number of days, relative to pregnancy start date, to end interval where MOIs will not be counted. **Note:** concept does not apply to as-treated episodes or episodes where ITTDays2 > 1 **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** 0
Type 4 Medical Product Pregnancy Start Date Removal Indicator	REMOVEMOIADATE	**Details:** indicator variable to provide option of excluding MOIs with ADate equivalent to delivery date. Valid values are: • **Y**: Exclude MOIs with ADate equivalent to pregnancy delivery date • **N**: Do not exclude MOIs with ADate equivalent to pregnancy delivery date **Defined by:** Requester **Input type:** Required **Format:** Character **Example:** Y
Type 4 Pre-pregnancy MOI evaluation period	PREPREGDAYS	**Details:** Number of days prior to pregnancy start to start counting MOI episodes. **Note 1:** If left blank, no pre-pregnancy period will be evaluated. **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** 90

Parameter	Field Name	Description
Maximum infant grace period in days	GRACEDAYS	**Details:** optional parameter to specify the maximum number of days to allow between an infant's birth and enrollment start. **Note 1:** Only applicable to analyses involving a linked mother-infant cohort **Note 2:** If the number of days between an infant's birth and enrollment start exceeds GRACEDAYS, the pregnancy episode will be excluded. **Named by:** Requester **Input type:** Optional (default value is 0) **Format:** Numeric **Example:** 30
Allowed number of days between Infant Birth date and Delivery date	MIDAYSDIFF	**Details:** indicates the allowable number of absolute difference between Infant Birth Date and delivery date used when selecting pregnancy episodes from MIL table. **Note 1:** Values must be continuous positive integer number. **Note 2:** Only applicable to analyses involving a linked mother-infant cohort. **Note 3:** If the number of days between the infant birth date and delivery date exceeds MIDAYSDIFF, the pregnancy episode will be excluded. **Defined by:** Requester **Input type:** Required **Format:** numeric; SAS numeric 4. **Example:** 3

f) Type 5 File

The Type 5 file is optional and its specification is only required for medical product utilization cohort identification strategy. Table 23 contains detailed specifications for this file.

Table 23. TYPE5FILE Specification

Parameter	Field Name	Description
Name of Cohort	GROUP	**Details:** standardized name used to differentiate cohorts. **Note 1**: multiple cohorts can be defined within the same Type 5 File. In this case all cohorts are queried independently and results are reported separately and labeled using each GROUP name specified. **Note 2:** GROUP is the primary key linking cohorts across input files; GROUP values must match (including case) between the TYPE5FILE and other input files. **Note 3**: GROUP values must remain consistent during the course of a surveillance activity. **Named by:** Request programmer **Input type:** Required **Format:** Alphanumeric; SAS character $40; no special characters (*e.g.,* commas, periods, hyphens, etc.) allowed, and underscores must be used to mark spaces. **Example:** Insulin
Allowed Number of Exposure Episodes per Individual	T5COHORTDEF	Details: indicates how many exposure periods an individual can contribute. Options include: 04: cohort includes all valid exposure episodes during the query period. Only the first valid episode's incidence is assessed using T5_WASHPER. **Defined by**: Requester **Input type**: Required **Format**: Alphanumeric; SAS character $2 **Example**: 04
Type 5 Exposure Washout Period	T5WASHPER	Details: length of washout period in days. The washout period is a period before an exposure episode during which an individual cannot have evidence of incidence-defining criteria (see Cohort Codes File specification for additional details on incidence-defining criteria). **Note 1**: for a Type 5 Analysis, only the first valid episode's incidence is assessed using specified washout criteria

Parameter	Field Name	Description
		Note 2: the MP algorithm may use days before the query start date to determine if continuous enrollment and incidence criteria are met. **Note 3**: special case: when T5WASHPER = missing the program requires ENRDAYS of continuous enrollment but only considers an exposure episode valid if, at index date, the member has no evidence of the exposure in their entire available enrollment history. **Defined by**: Requester **Input type**: Required **Format**: Numeric **Example**: 365
Requester-defined Exposure Episode Length	ITTDAYS	**Details:** number of days after exposure initiation that is considered "exposed time." **Note 1:** if exposure episode will be determined by outpatient pharmacy dispensing days supplied, leave this field blank. **Note 2:** exposure episodes will be censored at the first occurrence of the following: 1) end of enrollment; 2) occurrence of any additional requester-defined censoring criteria. **Note 3:** if this field is populated, the following fields should be left blank in this file: EPISODEGAP and EPISODEGAPTYPE (these parameters are for exposed time that is determined by outpatient pharmacy dispensing days supplied). **Defined by:** Requester **Input type:** Optional **Format:** Numeric **Example:** 30
Treatment Episode Gap Type	EPISODEGAPTYPE	**Details:** specifies the type of algorithm to use for the calculation of episode gaps. Valid values are: • **F**: Fixed episode gap. The value specified in EPISODEGAP will be used to determine if two consecutive claims are in the same episode. • **P**: Percentage episode gap. The value specified in EPISODEGAP will represent a percentage of the previous dispensing's days of supply to determine if two

Parameter	Field Name	Description
		consecutive claims are in the same episode (e.g., if EpisodeGap = 50, and day supply = 30, the gap would be 15). **Defined by:** Request Programmer **Input type:** Required (default value is P) **Format:** Alphanumeric; SAS character $1. **Example:** P
Exposure Episode Gap	EPISODEGAP	**Details:** used in conjunction with EPISODEGAPTYPE; sets the number of days allowed between two consecutive claims to consider them as part of the same treatment episode. For a given claim, if EPISODEGAPTYPE is Fixed (F), a gap of more than EPISODEGAP days between the claim date and the date of the last day of supply of the previous claim triggers a new treatment episode. If EPISODEGAPTYPE is Percentage (P), a gap of more than EPISODEGAP percent of the previous claim days of supply days between the claim date and the date of the last day of supply of the previous claim triggers a new treatment episode. For example, if EPISODEGAP=10 and EPISODEGAPTYPE=F, claim1's last day of supply is on 1/31/2012 and claim2's start date is 2/12/2012, the MP algorithm starts a new treatment episode on 2/12/2012 because there are more than 10 days between the two claims. **Note 1:** this value must be the same within a given query GROUP. **Defined by:** Requester **Input type:** Required (0 must be entered if no EPISODEGAP is required; default value is 0 for Fixed EPISODEGAPTYPE and 30 for Percentage EPISODEGAPTYPE) **Format:** Numeric **Example:** 10
Censor Episodes at Evidence of Death	CENSOR_DTH	**Details:** indicates if a treatment episode should be censored based on death date. Allowable values are "Y" and "N". Date of death can be determined two ways: 1. Using discharge status = expired in the SDD Encounter table. Death date is set to discharge date. OR

Parameter	Field Name	Description
		2. Using death date in the SDD Death table for records with Confidence=Excellent. **Note 1:** censoring is implemented by restricting enrollment eligibility. Member eligibility is truncated at death date. Once a death date is observed, a member can no longer contribute eligible periods (even if they are observed in the data). **Defined by:** Requester **Input type:** Required **Format:** SAS character $1 **Example:** Y
Censor Episodes at DP Data End Date	CENSOR_ DPEND	**Details:** indicates if a treatment episode should be censored based on DP data end date. Allowable values are "Y" and "N". **Note 1:** when CENSOR_DPEND = "Y", the program will adjust the query period end date (QUERYTO) to reflect DP_MaxDate. This will have implications on the attrition table and calculation of denominators, since episodes with an index date before QUERYTO but after DP_MaxDate cannot contribute to the cohort. **Defined by:** Requester **Input type:** Required **Format:** SAS character $1 **Example:** Y
Censor Episodes at Query End Date	CENSOR_QRYEND	**Details:** indicates if a treatment episode should be censored based on query data end date. Allowable values are "Y" and "N". **Defined by:** Requester **Input type:** Required **Format:** SAS character $1 **Example:** Y

g) Type 6 file

The Type 6 file is required for evaluating manufacture level product utilization and switching patterns. Table below contains detailed specifications for this file.

Table 24. TYPE6FILE Specification

Parameter	Field Name	Description
Name of Cohort	GROUP	**Details:** standardized name used to differentiate cohorts. **Note 1**: multiple cohorts can be defined within the same Type 6 File. In this case all cohorts are queried independently and results are reported separately and labeled using each GROUP name specified. **Note 2:** GROUP is the primary key linking cohorts across input files; GROUP values must match (including case) between the TYPE6FILE and other input files. **Note 3:** GROUP is the variable that indicates groupings of NDCs to represent manufacturer-level products other other requested-defined product groups. **Named by:** Request programmer **Input type:** Required **Format:** Alphanumeric; SAS character $40; no special characters (*e.g.*, commas, periods, hyphens, etc.) allowed, and underscores must be used to mark spaces. **Example:** Insulin
Allowed Number of Exposure Episodes per Individual	T6COHORTDEF	**Details:** indicates how many exposure periods an individual can contribute. Options include: • **01:** Cohort includes only the first valid exposure episode during the query period • **02:** Cohort includes all valid exposure episodes during the query period **Defined by:** Requester **Input type:** Required **Format:** Alphanumeric; SAS character $2 **Example:** 01
Type 6 Exposure Washout Period	T6WASHPER	**Details:** length of washout period in days. The washout period is a period before an exposure episode during which an individual cannot have evidence of incidence-defining criteria (see Cohort Codes File specification for additional details on incidence-defining criteria). **Note 1:** for a Type 6 Analysis, T6WASHPER is only used to calculate number of incident episodes. All other metrics are based on prevalent (0 day washout) episodes. This parameter will only be used to identify incident episodes for utilization

Parameter	Field Name	Description
		reporting purposes only. When episodes are assessed for product switching, all episodes (not just incident) will be used. **Note 2:** the MP algorithm may use days before the query start date to determine if continuous enrollment and incidence criteria are met. **Note 3:** special case: when T6WASHPER is missing the program requires ENRDAYS of continuous enrollment but only considers an exposure episode valid if, at index date, the member has no evidence of the exposure in their entire available enrollment history. **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** 365
Treatment Episode Gap Type	EPISODEGAPTYPE	**Details:** specifies the type of algorithm to use for the calculation of episode gaps. Valid values are: • **F**: Fixed episode gap. The value specified in EPISODEGAP will be used to determine if two consecutive claims are in the same episode. • **P**: Percentage episode gap. The value specified in EPISODEGAP will represent a percentage of the previous dispensing's days of supply to determine if two consecutive claims are in the same episode (e.g., if EpisodeGap = 50, and day supply = 30, the gap would be 15). **Defined by:** Requester **Input type:** Required (default value is P) **Format:** Alphanumeric; SAS character $1. **Example:** P
Exposure Episode Gap	EPISODEGAP	**Details:** used in conjunction with EPISODEGAPTYPE; sets the number of days allowed between two consecutive claims to consider them as part of the same treatment episode. For a given claim, if EPISODEGAPTYPE is Fixed (F), a gap of more than EPISODEGAP days between the claim date and the date of the last day of supply of the previous claim triggers a new treatment episode. If EPISODEGAPTYPE is Percentage (P), a gap of more than EPISODEGAP percent of the previous claim days of supply days between the claim date and the date of the last day of supply of the previous claim triggers a new treatment episode.

Parameter	Field Name	Description
		For example, if EPISODEGAP=10 and EPISODEGAPTYPE=F, claim1's last day of supply is on 1/31/2012 and claim2's start date is 2/12/2012, the MP algorithm starts a new treatment episode on 2/12/2012 because there are more than 10 days between the two claims. **Note 1:** this value must be the same within a given query GROUP. **Defined by:** Requester **Input type:** Required (0 must be entered if no EPISODEGAP is required; default value is 0 for Fixed EPISODEGAPTYPE and 30 for Percentage EPISODEGAPTYPE) **Format:** Numeric **Example:** 10
Censor Episodes at Evidence of Death	CENSOR_DTH	**Details:** indicates if a treatment episode should be censored based on death date. Allowable values are "Y" and "N". Date of death can be determined two ways: 1. Using discharge status = expired in the SDD Encounter table. Death date is set to discharge date. OR 2. Using death date in the SDD Death table for records with Confidence=Excellent. **Note 1:** censoring is implemented by restricting enrollment eligibility. Member eligibility is truncated at death date. Once a death date is observed, a member can no longer contribute eligible periods (even if they are observed in the data). **Defined by:** Requester **Input type:** Required **Format:** SAS character $1 **Example:** Y
Censor Episodes at DP Data End Date	CENSOR_ DPEND	**Details:** indicates if a treatment episode should be censored based on DP data end date. Allowable values are "Y" and "N". **Note 1:** when CENSOR_DPEND = "Y", the program will adjust the query period end date (QUERYTO) to reflect DP_MaxDate. This will have implications on attrition and switching calculations since episodes with an index date before QUERYTO but after DP_MaxDate cannot contribute to the cohort. **Defined by:** Requester **Input type:** Required **Format:** SAS character $1 **Example:** Y

Parameter	Field Name	Description
Censor Episodes at Query End Date	CENSOR_QRYEND	**Details:** indicates if a treatment episode should be censored based on query data end date. Allowable values are "Y" and "N". **Defined by:** Requester **Input type:** Required **Format:** SAS character $1 **Example:** Y
Date for product uptake computations	UPTAKEDATE	**Details:** which date field to use for product uptake duration computations (e.g. computation of time, in days, from [user-specified date]) to first observed valid dispensing date. **Valid values are:** PRODUCTAPPROVALDATE PRODUCTMARKETINGDATE OTHERPRODUCTDATE COMPUTEDSTARTMARKETINGDATE **Defined by:** Requester **Input type:** Required **Format:** SAS character $30 **Example:** PRODUCTAPPROVALDATE

h) Monitoring File

The Monitoring File is required for Type 1, Type 2, Type 4, Type 5, Type 6 analyses only. The file allows requesters to define specific time periods, or cumulative "looks" at data as part of sequential monitoring activities. Each time period is assigned a unique PERIODID value in the file. The main program parameters PERIODIDSTART and PERIODIDEND allow the request programmer to selectively execute time periods of interest using the Monitoring File.

For example, a requester may a priori specify the following time periods for evaluation:

PERIODID=1: January 1, 2015 – March 31, 2015
PERIODID=2: January 1, 2015 – June 30, 2015
PERIODID=3: January 1, 2015 – September 30, 2015
PERIODID=4: January 1, 2015 – December 31, 2015

These four periods are included in the Monitoring File with the corresponding PERIODID values. When data are complete through March 31, 2015, SOC can distribute a program package with the above Monitoring File contents and macro parameters PERIODIDSTART=1 and PERIODIDEND=1. When data are complete through June 30, 2015, SOC can distribute the same package with macro parameters PERIODIDSTART=1 and PERIODIDEND=2 (if the requester wants to execute a query starting in PERIODID 1 and ending in PERIODID 2).

The CIDA tool, to support sequential monitoring activities, will generate output by PERIODID.

Table 25 contains detailed specifications for this file.

Table 25. MONITORINGFILE Specification

Parameter	Field Name	Description
Time Period Indicator	PERIODID	**Details:** identifier for each STARTFOLLOWUP/ ENDDATE combination. **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** 1
Query Period Start	STARTFOLLOWUP	**Details:** start date for the query period. Should be identical across all PERIODIDs (*i.e.,* the start date for analysis should always be the same). **Defined by:** Requester **Input type:** Required **Format:** Numeric; Date9. **Example:** 01JAN2015
Query Period End	ENDDATE	**Details:** end date for the query period. **Defined by:** Requester **Input type:** Required **Format:** Numeric; Date9. **Example:** 31DEC2015

i) Cohort Codes File

The Cohort Codes File is required. It is the primary file for specifying codes used to define exposures, exposure incidence criteria, outcomes, outcome incidence criteria, and live births (for the pregnancy episodes and medical product use cohort identification strategy). NDCs, ICD procedure and diagnosis codes, HCPCS codes, and/or laboratory result values can be used in any combination and can be restricted to specific care settings and diagnosis code positions (*e.g.,* principal discharge diagnoses only). Table 26 contains detailed specifications for this file.

Table 26. COHORTCODES Specification

Parameter	Field Name	Description
Name of Cohort	GROUP	**Details:** standardized name used to differentiate cohorts. **Note 1**: multiple cohorts can be defined within the same Cohort Codes File. In this case all cohorts are queried independently and results are reported separately and labeled using each GROUP name specified. **Note 2:** GROUP is the primary key linking cohorts across input files; GROUP values must match (including case) between the COHORTCODES file and other input files.

Parameter	Field Name	Description
		Note 3: GROUP values must remain consistent during the course of a surveillance activity. **Named by:** Request programmer **Input type:** Required **Format:** Alphanumeric; SAS character $40; no special characters (*e.g.,* commas, periods, hyphens, etc.) allowed, and underscores must be used to mark spaces. **Example:** Insulin
Name of Stockpiling Group within the Cohort	STOCKGROUP	**Details:** standardized name used to refer to a specific exposure/HOI within a given GROUP. **Note 1:** the STOCKGROUP field is used by the stockpiling_algorithm as group categories to adjust service dates. **Note 2:** useful when a GROUP contains multiple exposures of interest. For example, if GROUP= "Insulin" STOCKGROUP could take values of "Insulin_Oral" and "Insulin_Injectable". **Note 3:** no output will be presented by STOCKGROUP. All output is presented at the GROUP level. **Named by:** Request programmer **Input type:** Required **Format:** Alphanumeric; SAS character $30; special characters (*e.g.,* commas, periods, hyphens, etc.) allowed and underscores must be used to mark spaces. **Example:** Insulin_Oral
Code Category	CODECAT	**Details:** type of each code category value included in the CODETYPE field (below) of this file. Valid values include: • **RX:** NDC • **DX:** Diagnosis code • **PX:** Procedure code • **LB:** Lab code • **AN:** Target age anniversary • **DT:** Fixed calendar date • **MI:** Mother-Infant Linkage table code **Note 1:** values AN and DT are only relevant for Types 1, 2 and 3. **Note 2**: value MI is only relevant for Type 4. **Note 3:** for CODECAT=AN and DT, CARESETTINGPRINCIPAL, CODESUPPLY,

Parameter	Field Name	Description
		RAWLABDATETYPE, RAWLABRESULT should be left blank. **Defined by:** Requester **Input type:** Required **Format:** SAS character $2. **Example:** DX

Code Type	CODETYPE	**Details:** type of each code value included in the CODE field (below) of this file. Valid values include:
		
		If CODECAT = RX:
		• **09**: 9-digit NDC
		• **11**: 11-digit NDC
		
		If CODECAT = DX:
		• **09**: ICD-9-CM
		• **10**: ICD-10-CM
		• **11:** ICD-11-CM
		• **OT:** Other
		
		If CODECAT = PX:
		• **09**: ICD-9-CM
		• **10**: ICD-10-CM
		• **11**: ICD-11-CM
		• **C4**: CPT-4 (*i.e.,* HCPCS Level I)
		• **HC**: HCPCS (*i.e.,* HCPCS Level II)
		• **H3**: HCPCS Level III
		• **C2**: CPT Category II
		• **C3**: CPT Category III
		• **ND**: 11-digit NDC
		• **RE**: Revenue
		• **LO**: Local homegrown
		• **OT**: Other
		
		If CODECAT = LB:
		• **01N**: extract quantitative lab test result using SOC-defined lab code (see Lab Code Lookup Table)
		• **02N**: extract quantitative lab test result using LOINC
		• **'px'N**: extract quantitative lab test result using the following codes
		  • **09N**: ICD-9-CM
		  • **10N**: ICD-10-CM
		  • **11N**: ICD-11-CM
		  • **C4N**: CPT-4 (*i.e.,* HCPCS Level I)
		  • **HCN**: HCPCS (*i.e.,* HCPCS Level II)

Parameter	Field Name	Description
		• **H3N:** HCPCS Level III • **C2N:** CPT Category II • **C3N:** CPT Category III • **NDN:** 11-digit NDC • **REN:** Revenue • **LON:** Local homegrown • **01C:** extract qualitative lab test result using SOC-defined lab code (see Lab Code Lookup Table) • **02C:** extract qualitative lab test result using LOINC • **'px'C:** extract qualitative lab test result using the following codes • **09C**: ICD-9-CM • **10C**: ICD-10-CM • **11C**: ICD-11-CM • **C4C**: CPT-4 (*i.e.,* HCPCS Level I) • **HCC**: HCPCS (*i.e.,* HCPCS Level II) • **H3C:** HCPCS Level III • **C2C:** CPT Category II • **C3C:** CPT Category III • **NDC:** 11-digit NDC • **REC:** Revenue • **LOC:** Local homegrown If CODECAT = AN: • **Y:** age anniversary specified in years • **M:** age anniversary specified in months • **W:** age anniversary specified in weeks • **D:** age anniversary specified in days If CODECAT = MI: • **M:** cohort selection based on mother (i.e., extract a list of mother delivery dates and PatIDs from the SCDM MIL table. If users wish to restrict to mothers linked to babies in the SCDM MIL table, use CODECAT = MI and CODETYPE = M to required match method values). **Note 1:** this parameter is not used when CODECAT=DT and should be left blank. **Note 2:** as the LOINC field is not populated by all Data Partners in the SCDM Laboratory Result table and the CPT code may not be specific to a particular lab test, it

Parameter	Field Name	Description
		is strongly recommended that the Laboratory Result table be queried using SOC-defined lab codes. **Defined by:** Requester **Input type:** Required **Format:** Alphanumeric; SAS character $3. **Example:** 09
Code	CODE	**Details:** NDC, procedure, diagnosis, and/or lab code of interest. For CODECAT=RX, DX, PX, or LB, the requester should input the NDC, diagnosis, procedure and/or lab code values used to determine the index date. For CODECAT=AN, the requester should input the desired age in the unit specified in CODETYPE. For CODECAT=DT, the requester should input the calendar date in the SAS format Date9. (e.g., 01Jan2015) For CODECAT=MI, the requester should input the concatenation of valid values of MatchMethod and BirthType (e.g., SI1). **Valid MatchMethod values are:** • BC = Birth Certificate • RE = DP maintained birth registry • SI = health plan subscriber or family number • LA = exact or probabilistic last name and address match based upon health plan administrative data • N1 = No subscriber/family IDs available for linkage • N2 = No name/address available for linkage • N3 = Neither subscriber/family IDs nor name/address available for linkage • NA = no linkage made; any other reasons • OT = other **Valid Birth_Type values are:** • 1 = 1 live birth **Note 1:** Codes are matched using exact values (*i.e.,* 3-digit code lookup requires an exact 3-digit code match). Wildcard match (*) functionality is also available for ICD-9 diagnosis codes (*e.g.,* querying "250*0" would be used to find any ICD-9-CM diagnosis codes for diabetes type II, or "250**" to find ICD-9-CM

Parameter	Field Name	Description
		diagnosis codes for all diabetes codes in the range "250.00 - 250.99"). To get "starts with" codes, the user will have to specify 250, 250*, 250**. **Note 2**: For NDCs, either 9 or 11 digit codes can be entered. **Note 3:** remove decimal points in the code value. **Note 4**: CODETYPE/CODECAT must be consistent with the expected format of the CODE value (*e.g.,* the program will not find any valid matches in the data for CODECAT=RX, CODETYPE=11 and a 9-digit NDC value). **Note 5**: Duplicate CODECAT-CODETYPE-CODE-CARESETTING-PRINCIPAL combinations are removed by the MP algorithm. **Note 6**: 'V' and 'E' ICD-9-CM diagnosis codes must be specified using uppercase 'V' and 'E'. **Note 7:** For CODECAT = MI, multiple match methods can be selected by including multiple rows. Only Singleton births are considered. **Defined by:** Requester, with support from the SOC as needed **Input type:** Required **Format:** Alphanumeric; SAS character $11. **Example: (CODECAT=RX; CODETYPE=11):** 12345678911
Care Setting and Diagnosis Position Requirements	CARESETTINGPRINCIPAL	**Details:** defines the care setting and principal diagnosis position requirements for each code. This field uses combination(s) of the SCDM variables care setting (ENCTYPE) and principal discharge diagnosis flag (PDX) to restrict the observance of codes to those in the requested care settings and with the requested diagnosis position. If no restrictions are required (*e.g.,* requester wants all care settings and any value of PDX), leave the field blank. The following are valid entries; all entries must be in single quotes and separated by a space: • **IPP:** inpatient hospital stays, principal diagnoses • **IPS:** inpatient hospital stays, secondary diagnoses • **IPX:** inpatient hospital stays, unclassified diagnoses

Parameter	Field Name	Description
		• **ISP:** non-acute institutional stays, principal diagnoses • **ISS:** non-acute institutional stays, secondary diagnoses • **ISX:** non-acute institutional stays, unclassified diagnoses • **ED*:** emergency department encounters • **AV*:** ambulatory visits • **OA*:** other ambulatory visits **Request Programmer Note 1:** the wildcard symbol (*) can be used to represent "any" values of either care setting or principal discharge diagnosis flag. For example, CARESETTINGPRINCIPAL = 'IP*' will restrict codes to those observed in the inpatient setting irrespective of the principal diagnosis flag value. CARESETTINGPRINCIPAL = '**P' will restrict diagnosis codes to those in the principal position, irrespective of the care setting. **Request Programmer Note 2:** the principal discharge diagnosis flag is only relevant for diagnosis codes. All other codes should use the * wildcard for the third digit of the CARESETTINGPRINCIPAL value. **Note 3:** CARESETTINGPRINCIPAL is allowed to vary between CODEs within the same GROUP. For example, CARESETTINGPRINCIPAL is allowed to equal 'IPP' for one diagnosis code and 'IPP' 'EDP' for another diagnosis code *in the same GROUP*. **Defined by:** Requester **Input type:** Optional; Default: blank (*i.e.,* no restrictions) **Format:** Alphanumeric **Example:** 'IPX' 'ED*' '**P'
Code Relevance to Type 1 Cohort Index Date Definition	T1_INDEX	**Details:** indicates, for each code listed in the file, what role the code will play in defining a Type 1 (background rate calculation) index date. Valid values are: • **DEF**: code should be used to identify the TYPE1 cohort index date • **IOC:** code should be used to assess TYPE1 cohort index incidence criteria only • **NOT**: code should not be used to define TYPE1 cohort index date or incidence criteria

Parameter	Field Name	Description
		Note 1: for each GROUP, a CODE and CARESETTINGPRINCIPAL combination should have only one T1_INDEX value. If the same code must be used with two different CARESETTINGPRINCIPAL values (e.g., the requester wants to define the index date using code 410 in the IP setting, but define incidence using code 410 in any setting), the request programmer must 1) list code 410 with CARESETTINGPRINCIPAL value = 'IP*' and T1_INDEX= DEF; **and** 2) list code 410 with CARESETTINGPRINCIPAL value = 'ED*' 'IS*' 'OA*' 'AV*' and T1_INDEX= IOC. **Note 2:** if a Type 1 cohort is not being created, all codes in the file should have T1_INDEX=NOT. **Note 3:** when T1_INDEX= "DEF", the MP automatically assesses index incidence with respect to these codes. **Note 4:** the value IOC is reserved for codes that are used to define index date incidence only. Consider the example where a requester wants to look at new use of Drug A, but wants new use of Drug A to be defined as no use of *any drug in Drug A's class*. Drug A codes would be listed in the file with T1_Index=DEF, and all other codes in Drug A's class would be listed with T1_Index=IOC. **Defined by:** Requester **Input type:** Required **Format:** SAS character $3. **Example:** DEF
Code Relevance to Type 2 Cohort Index Date Definition	T2_INDEX	**Details:** indicates, for each code listed in the file, what role the code will play in defining a Type 2 (exposures and follow-up time) index date/exposure episode. Valid values are: • **DEF**: code should be used to identify the TYPE2 cohort index date • **IOT:** code should be used to assess TYPE2 cohort index incidence only; exposure episode should be truncated at the occurrence of the code • **IOD:** code should be used to assess TYPE2 cohort index incidence only; exposure episode should NOT be truncated at the occurrence of the code

Parameter	Field Name	Description
		• **FUT:** code should be used to truncate exposure episodes (truncate on the day of first occurrence) • **NOT**: code should not be used to define TYPE2 cohort index date or incidence criteria **Note 1:** if a Type 2 cohort is not being created, all codes in the file should have T2_INDEX=NOT. **Note 2:** when T2_INDEX= "DEF", the MP automatically assesses index incidence with respect to these codes. **Note 3:** the values IOT and IOD are reserved for codes that are used to define exposure episode incidence only. Consider the example where a requester wants to look at new use of Drug A, but wants new use of Drug A to be defined as no use of *any drug in Drug A's class*. Drug A codes would be listed in the file with T2_Index=DEF, and all other codes in Drug A's class would be listed with either T2_Index=IOT or T2_Index=IOD. **Note 4:** the difference between IOT and IOD is whether exposed time is censored at the occurrence of the code. If the exposure episode should be censored select IOT; if not, select IOD. **Note 5:** the value FUT is reserved for codes that are not used to define exposure episodes or incidence criteria, but are used to censor exposed time. **Note 6:** if a code is designated as FUT or IOT, and the code occurs on the same day as the index date, the patient will be included in the cohort and contribute one day at-risk. **Defined by:** Requester **Input type:** Required **Format:** SAS character $3. **Example:** DEF
Code Relevance to Type 2 Cohort HOI Definition	T2_FUP	**Details:** indicates, for each code listed in the file, what role the code will play in defining a Type 2 (exposures and follow-up time) HOI. Valid values are: • **DEF**: code should be used to identify the TYPE2 cohort HOI • **IOC:** code should be used to assess TYPE2 cohort HOI incidence only

Parameter	Field Name	Description
		• **NOT**: code should not be used to define TYPE2 cohort HOI or incidence **Note 1:** for each GROUP, a CODE and CARESETTINGPRINCIPAL combination should have only one T2_FUP value. If the same code must be used with two different CARESETTINGPRINCIPAL values (e.g., the requester wants to define the HOI using code 410 in the IP setting, but define incidence using code 410 in any setting), the request programmer must 1) list code 410 with CARESETTINGPRINCIPAL value = 'IP*' and T2_FUP = DEF; **and** 2) list code 410 with CARESETTINGPRINCIPAL value = 'ED*' 'IS*' 'OA*' 'AV*' and T2_FUP = IOC. **Note 2:** if a Type 2 cohort is not being created, all codes in the file should have T2_FUP=NOT. **Note 3:** when T2_FUP="DEF", the MP automatically assesses HOI incidence with respect to these codes. **Note 4:** the value IOC is reserved for codes that are used to define index date incidence only. Consider the example where a requester wants to look at new occurrence of Outcome Y, but wants new occurrence of Outcome Y to be defined as no codes for either Outcome Y or Outcome Z during the washout period. Outcome Y codes would be listed in the file with T2_FUP=DEF, and Outcome Z codes would be listed with T2_FUP =IOC. **Defined by:** Requester **Input type:** Required **Format:** SAS character $3. **Example:** IOC
Code Relevance to Type 2 Concomitant event Definition	CONC_FUP	**Details:** indicates, for each code listed in the file, what role the code will play in defining a concomitant event. Valid values are: • **DEF**: code should be used to identify the concomitant episodeevent • **IOC:** code should be used to assess concomitant event incidence only • **NOT**: code should not be used to define concomitant event or incidence **Note 1:** for each GROUP, a CODE and CARESETTINGPRINCIPAL combination should have only one CONC_FUP value. If the same code must be used

Parameter	Field Name	Description
		with two different CARESETTINGPRINCIPAL values (e.g., the requester wants to define the HOI using code 410 in the IP setting, but define incidence using code 410 in any setting), the request programmer must 1) list code 410 with CARESETTINGPRINCIPAL value = 'IP*' and CONC_FUP = DEF; **and** 2) list code 410 with CARESETTINGPRINCIPAL value = 'ED*' 'IS*' 'OA*' 'AV*' and CONC_FUP = IOC. **Note 2:** if a concomitant episodes are not being created, all codes in the file should have CONC_FUP =NOT. **Note 3:** when CONC_FUP ="DEF", the MP automatically assesses HOI incidence with respect to these codes. **Note 4:** the value IOC is reserved for codes that are used to define index date incidence only. Consider the example where a requester wants to look at new occurrence of Outcome Y, but wants new occurrence of Outcome Y to be defined as no codes for either Outcome Y or Outcome Z during the washout period. Outcome Y codes would be listed in the file with CONC_FUP =DEF, and Outcome Z codes would be listed with CONC_FUP =IOC. **Defined by:** Requester **Input type:** Required **Format:** SAS character $3. **Example:** DEF
Code Relevance to Type 3 Cohort Index Date Definition	T3_INDEX	**Details:** indicates, for each code listed in the file, what role the code will play in defining a Type 3 exposure. Valid values are: • **DEF**: code should be used to identify the TYPE3 cohort exposure • **IOC:** code should be used to assess TYPE3 cohort exposure incidence criteria only • **NOT**: code should not be used to define TYPE3 cohort exposure or exposure incidence criteria **Note 1:** for each GROUP, a CODE and CARESETTINGPRINCIPAL combination should have only one T3_INDEX value. If the same code must be used with two different CARESETTINGPRINCIPAL values (e.g., the requester wants to define the index date using code 410 in the IP setting, but define incidence using code 410 in any setting), the request programmer must

Parameter	Field Name	Description
		1) list code 410 with CARESETTINGPRINCIPAL value = 'IP*' and T3_INDEX= DEF; **and** 2) list code 410 with CARESETTINGPRINCIPAL value = 'ED*' 'IS*' 'OA*' 'AV*' and T3_INDEX= IOC. **Defined by:** Requester **Input type:** Required **Format:** Char (3) **Example:** DEF
Code Relevance to Type 3 Cohort HOI Definition	T3_FUP	**Details:** indicates, for each code listed in the file, what role the code will play in defining a Type 3 HOI. Valid values are: • **DEF**: code should be used to identify the TYPE3 cohort HOI • **IOC:** code should be used to assess TYPE3 cohort HOI incidence criteria only • **NOT**: code should not be used to define TYPE3 cohort HOI or HOI incidence criteria **Note 1:** for each GROUP, a CODE and CARESETTINGPRINCIPAL combination should have only one T3_FUP value. If the same code must be used with two different CARESETTINGPRINCIPAL values (e.g., the requester wants to define the HOI using code 410 in the IP setting, but define incidence using code 410 in any setting), the request programmer must 1) list code 410 with CARESETTINGPRINCIPAL value = 'IP*' and T3_FUP= DEF; **and** 2) list code 410 with CARESETTINGPRINCIPAL value = 'ED*' 'IS*' 'OA*' 'AV*' and T3_FUP= IOC. **Defined by:** Requester **Input type:** Required **Format:** Char (3) **Example:** DEF
Code Relevance to Type 4 Pregnancy Cohort Delivery DateDefinition	T4_INDEX	**Details:** indicates, for each code listed in the file, what role the code will play in defining a Type 4 (pregnancy) delivery date/exposure episode. Valid values are: • **DEF**: code should be used to identify the TYPE4 pregnancy cohort delivery date. • **IOC:** code should be used to assess TYPE4 pregnancy cohort delivery date incidence only • **MP*n*:** medical product exposure, where *n* identifies a unique medical product of interest

Parameter	Field Name	Description
		• **IC*n*:** incident medical product exposure, where *n* identifies a unique medical product of interest. The same n value for MPn and IMPn indicates exposure and incidence defining criteria for that medical product. • **NOT**: code should not be used to define TYPE4 cohort delivery date or incidence criteria **Note 1:** for each GROUP, CODE and CARESETTINGPRINCIPAL combination should have only one T4_INDEX value. If the same code must be used with two different CARESETTINGPRINCIPAL values (e.g., the requester wants to define the delivery date using code 670 in the IP setting, but define incidence using code 670 in any setting), the request programmer must 1) list code 670 with CARESETTINGPRINCIPAL value = 'IP*' and T4_INDEX= DEF; **and** 2) list code 670 with CARESETTINGPRINCIPAL value = 'ED*' 'IS*' 'OA*' 'AV*' and T2_INDEX= IOC. **Note 2:** if a Type 4 cohort is not being created, all codes in the file should have T4_INDEX=NOT. **Note 3:** when T4_INDEX= "DEF", the MP automatically assesses delivery incidence with respect to these codes. **Defined by:** Requester **Input type:** Required **Format:** SAS character >3. **Example:** DEF
Code Relevance to Type 4 Cohort HOI Definition	T4_FUP	**Details:** indicates, for each code listed in the file, what role the code will play in defining a Type 4 (pregnancy) HOI. Valid values are: • **DEF**: code should be used to identify the TYPE4 cohort HOI • **NOT**: code should not be used to define TYPE4 cohort HOI **Note 1:** if a Type 4 cohort is not being created, all codes in the file should have T4_FUP=NOT. **Defined by:** Requester **Input type:** Required **Format:** SAS character $3. **Example:** DEF

Parameter	Field Name	Description
Code Relevance to Type 5 Cohort Index Date Definition	T5_INDEX	**Details:** indicates, for each code listed in the file, what role the code will play in defining a Type 5 index date/exposure episode. Valid values are: • **DEF**: code should be used to identify the TYPE5 cohort index date • **IOT:** code should be used to assess TYPE5 cohort index incidence only; exposure episode should be truncated at the occurrence of the code • **IOD:** code should be used to assess TYPE5 cohort index incidence only; exposure episode should NOT be truncated at the occurrence of the code • **FUT:** code should be used to truncate exposure episodes (truncate on the day of first occurrence) • **NOT**: code should not be used to define TYPE5 cohort index date or incidence criteria **Note 1:** for each GROUP, a CODE and CARESETTINGPRINCIPAL combination should have only one T5_INDEX value. If the same code must be used with two different CARESETTINGPRINCIPAL values (e.g., the requester wants to define the index date using code 410 in the IP setting, but define incidence using code 410 in any setting), the request programmer must 1) list code 410 with CARESETTINGPRINCIPAL value = 'IP*' and T5_INDEX= DEF; **and** 2) list code 410 with CARESETTINGPRINCIPAL value = 'ED*' 'IS*' 'OA*' 'AV*' and T5_INDEX= IOD. **Note 2:** if a Type 5 cohort is not being created, all codes in the file should have T5_INDEX=NOT. **Note 3:** when T5_INDEX= "DEF", the MP automatically assesses index incidence with respect to these codes. **Note 4:** the values IOT and IOD are reserved for codes that are used to define exposure episode incidence only. Consider the example where a requester wants to look at new use of Drug A, but wants new use of Drug A to be defined as no use of *any drug in Drug A's class*. Drug A codes would be listed in the file with T5_Index=DEF, and all other codes in Drug A's class

Parameter	Field Name	Description
		would be listed with either T5_Index=IOT or T5_Index= IOD. **Note 5:** the difference between IOT and IOD is whether exposed time is censored at the occurrence of the code. If the exposure episode should be censored select IOT; if not, select IOD. **Note 6:** the value FUT is reserved for codes that are not used to define exposure episodes or incidence criteria, but are used to censor exposed time. **Note 7:** if a code is designated as FUT or IOT, and the code occurs on the same day as the index date, the patient will be included in the cohort and contribute one day at-risk. **Defined by:** Requester **Input type:** Required **Format:** SAS character $3. **Example:** DEF
Code Relevance to Type 6 Cohort Index Date Definition	T6_Index	**Details:** indicates, for each code listed in the file, what role the code will play in defining a Type 6 (utilization and switching) index date/exposure episode. Valid values are: • **DEF**: code should be used to identify the TYPE6 cohort index date • **IOD:** code should be used to assess TYPE6 cohort index incidence only; exposure episode should NOT be truncated at the occurrence of the code • **NOT**: code should not be used to define TYPE6 cohort index date or incidence criteria **Note 1:** for each GROUP, a CODE and CARESETTINGPRINCIPAL combination should have only one T6_INDEX value. If the same code must be used with two different CARESETTINGPRINCIPAL values (e.g., the requester wants to define the index date using code 410 in the IP setting, but define incidence using code 410 in any setting), the request programmer must 1) list code 410 with CARESETTINGPRINCIPAL value = 'IP*' and T6_INDEX= DEF; **and** 2) list code 410 with CARESETTINGPRINCIPAL value = 'ED*' 'IS*' 'OA*' 'AV*' and T6_INDEX= IOD. **Note 2:** if a Type 6 cohort is not being created, all codes in the file should have T6_INDEX=NOT.

Parameter	Field Name	Description
		Note 3: when T6_INDEX= "DEF", the MP automatically assesses index incidence with respect to these codes. **Note 4:** the value IOD is reserved for codes that are used to define exposure episode incidence only. Consider the example where a requester wants to look at new use of Drug A, but wants new use of Drug A to be defined as no use of *any drug in Drug A's class*. Drug A codes would be listed in the file with T6_Index=DEF, and all other codes in Drug A's class would be listed with T6_Index=IOD. **Defined by:** Requester **Input type:** Required **Format:** SAS character $3. **Example:** DEF
Forced supply to attach to code	CODESUPPLY	**Details:** indicates, for each code listed in the file, a forced supply that should be attached to the code. The specified code supply will replace the RxSup for RX codes. **Note** Non-Rx codes are not stockpiled **Defined by:** Requester **Input type:** Optional **Format:** Numeric **Example:** 30
Exclude Day Supply	EXCLUDESUPPLY	**Description:** indicates if lookback period to define new use of a medical product looks for evidence of a dispensing or evidence of days supply, when NDCs are used to define criteria. Allowable values : • **N:** lookback period should search for evidence of days supply • **Y:** lookback period should search for evidence of a dispensing date **Named by:** Requester **Input type:** Optional **Format:** Character (1)
Lab Date Selection Algorithm	RAWLABDATETYPE	**Details:** relevant for requests that query laboratory result values. Field specifies in what sequence date(s) in the SCDM Laboratory Result table should be considered to select one relevant date for a laboratory result of interest. The parameter will allow the requester to either specify 1) a single date variable (Lab_dt, Order_dt, or Result_dt) to use; or 2) a

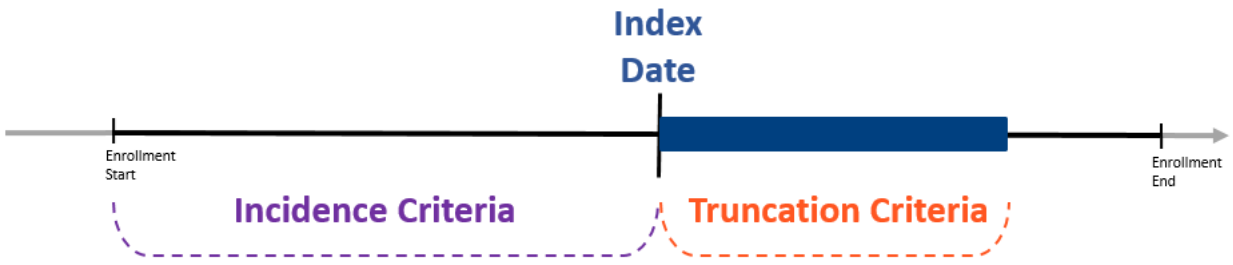
Parameter	Field Name	Description
		hierarchy to choose a date variable (*e.g.,* select Lab_dt else if missing select Result_dt else if missing select Order_dt). Valid values are any combination of the following: • **L:** Lab Date • **O:** Order Date • **R:** Result Date **Note 1:** leave blank if CODECAT ≠ LB. **Note 2:** many Data Partners do not populate all three date fields. Use of an algorithm for date selection is recommended. **Defined by:** Requester **Input type:** Required **Format:** SAS character $3. **Example 1:** RawLabDateType =LRO. In this case, the program will use Lab_dt else if missing use Result_dt else if missing use Order_dt. **Example 2:** RawLabDateType =L. In this case, the program will use Lab_dt only.
Lab Result Values	RAWLABRESULT	**Details:** specifies the lab result value or lab result range for querying. RAWLABRESULT allows for values or ranges of quantitative laboratory results (*e.g.,* 100; 100-200) and values of qualitative laboratory results (*e.g.,* "POSITIVE"). Valid values are: • <=X (less than or equal to X) • <X (less than X) • >=X (greater than or equal to X) • >X (greater than X) • ~=X (not equal to X) • X:Y (between X and Y) Any string of relevant characters is allowed for qualitative results querying. **Note 1:** leave blank if CODECAT ≠ LB. **Note 2:** There are two fields in the Laboratory Result table that include results: MS_Result_C (contains results for qualitative tests) and MS_Result_N (contains results for quantitative tests). The field where the result will be queried will depend on the RAWCODETYPE value.

Parameter	Field Name	Description
		Note 3: Ranges cannot be specified with hyphens. Must use ":". **Defined by:** Requester Input type: Required for laboratory results **Format:** Alphanumeric; **Example 1:** RawLabResult=20:50 **Example 2:** RawLabResult=POSITIVE
Product Approval Date	PRODUCTAPPROVALDATE	**Details:** defines user-specified approval date for product. **Note 1**: Valide for Type 6 only **Defined by:** Requester **Input type:** Required **Format:** Date9. **Example:** 01Mar2012
Product Start Marketing Date	PRODUCTMARKETINGDATE	**Details:** defines user-specified start marketing date for product. If unknown, leave null. **Note 1:** Valid for Type 6 only **Defined by:** Requester **Input type:** Optional **Format:** Date9. **Example:** 01Mar2012
Other Product Date	OTHERPRODUCTDATE	**Details:** defines user-specified other product-related date. If unknown or not needed, leave null. **Note 1**: Valid for Type 6 only **Defined by:** Requester **Input type:** Optional **Format:** Date9. **Example:** 01Mar2012

Figure 38 below illustrates the differences between various values of parameters T2_INDEX and T2_FUP and how each should be used depending on if the codes are for the identification of index date, incidence criteria or truncation criteria.

Figure 38. Differences Between Values of Parameters that Indicate Which Role Each Codes Plays in Defining Type 2 (exposures and follow-up time) index date/exposure episode or HOI

T2_INDEX and T2_FUP Parameters	Index Date Codes	Incidence Criteria Codes	Truncation Criteria Codes
DEF	✓	✓	✗
IOT	✗	✓	✓
IOD or IOC	✗	✓	✗
FUT	✗	✗	✓



j) User-defined Strata Levels Lookup Table

The User-defined Strata Levels Lookup table is required in the *inputfiles* folder. Programmatically this table functions as a look-up table. However as values are user-defined per each package it is considered an input file.

The User-defined Strata Levels Lookup Table (Table 27) defines both the ouput tables that will be returned as well as the stratifications of each output table. All output tables desired must be identified in the Output Table Identifier field with the exception of ([RUNID]_baseline_[PERIODID].sas7bdat, [RUNID]_signature.sas7bdat, and [RUNID]_attrition.sas7bdat which will be automatically produced. Besides these exceptions, **Output tables not specified within the User-defined Strata Levels Lookup Table will not be returned.** ([RUNID]_baseline_[PERIODID].SAS7bdat is not available for Type 6 cohort identification).

For each output table requested, strata levels must be identified with a Strata Level ID and defined in the Strata Level Variables field. When performing a Type 1 or Type 2 analysis that calculates denominators (eligible members and eligible member-days), the CIDA tool will only populate these fields for Standard Strata Levels. This table lists Standard Strata Level IDs (Levels 000 – 199) with pre-defined stratifications and the accompanying output tables that must use these strata to calculate denominators. Other output tables from different cohort identification strategies may also use these Standard Strata Levels, but there is no dependency or limitation to the Standard Strata Levels for calculations.

Alternatively, custom strata levels may be created using any combination of valid stratification variables identified for each particular output table. When creating custom strata levels the user must consult this table to determine what stratification variables are available per each output table requested.

Table 27. STRATA Specification

Parameter	Field Name	Description
Output Table Identifier	TableID	**Details**: CIDA output table identifier for which strata applies. **Valid values**: t1cida: [*RUNID*]_t1_cida.sas7bdat t1censor: [*RUNID*]_censor_cida.sas7bdat t2cida: [*RUNID*]_t2_cida.sas7bdat t2censor: [*RUNID*]_censor_cida.sas7bdat t3cida: [*RUNID*]_t3_cida.sas7bdat t4elig: [*RUNID*]_t4_cida_elig.sas7bdat t4treat: [*RUNID*]_t4_cida_treat.sas7bdat t4ctrl: [*RUNID*]_t4_cida_ctrl.sas7bdat t4treatgestwK: *[RUNID]*_t4_cida_treat_gestwk.sas7bdat t4ctrlgestwK: *[RUNID]*_t4_cida_ctrl_gestwk.sas7bdat t5disp: [*RUNID*]_t5_cida_disp_by_daysupp.sas7bdat t5first: [*RUNID*]_t5_cida_firsteps.sas7bdat t5alleps: [*RUNID*]_t5_cida_alleps.sas7bdat t5episdur: [*RUNID*]_t5_cida_epidur.sas7bdat t5censor: [*RUNID*] _t5_cida_episdur_censor.sas7bdat t5gaps: [*RUNID*]_t5_cida_gaps.sas7bdat t6counts: [RUNID]_t6_utilcounts t6disp: [RUNID]_t6_utildispstats t6episdur: [RUNID]_t6_utilepisdurstats t6censor: [RUNID]_t6_utilepis_censor t6uptake: [RUNID]_t6_utiluptakestats t6trend: [RUNID]_t6_trendcounts t6switchepisdur: [RUNID]_t6_switchepisdurstats t6plota: [RUNID]_t6_switchplota t6plotb: [RUNID]_t6_switchplotb t2multevent: [RUNID]_t2_multevent.sas7bdat t2epigap: [RUNID]_t2_epigap t2overlap: [RUNID]_t2_overlap.sas7bdat t2conc: [RUNID]_t2_concomitance.sas7bdat t2its: [RUNID]_t2_its.sas7bdat **Note 1:** There will need to be a separate row per each tableID to apply a particular strata to. For example, if a user wishes to output overall counts to t1cida and t1censor, two rows will need to exist in the lookup for these two tableIDs with the same levelID and levelVars. **Note 2:** By default all output will be stratified by group. **Note 3:** By default, t3cida output will be stratified by a censor indicator variable (CENSOR), to indicate whether the analytic cohort contains episodes that do not have sufficient post-exposure enrollment

Parameter	Field Name	Description
		Note 4: If a table is not specified, it will not be produced. In a type 4 analysis, if t4ctrl is not specified, a non-pregnant control cohort will not be created. **Defined by:** Request programmer **Input type:** Required **Format:** Alphanumeric **Example1**: TableID = t1cida t1censor
Strata Level Identifier	LEVELID	**Details**: level identifier applied to each distinct combination of strata/row of data in the look-up table. **Note 1:** should follow the format ### and value must be distinct to row of data. **Note 2**: For standard values, utilize predefined values as shown in Table 28. For new values, first-digit should start with at least a 2. **Defined by:** Request programmer **Input type:** Required **Format:** Alphanumeric **Example1**: LEVELID = 003
Strata Level Variable(s)	LEVELVARS	**Details**: variable to include in distinct level identifier. **Note 1:** to included more than one variable in strata definition, i.e. Hispanic * sex, create space-delimited list of all variables to be cross tabulated. **Defined by:** Request programmer **Input type:** Required **Format:** Alphanumeric **Example 1**: LEVELVARS = ageGroup **Example 2**: LEVELVARS = ageGroup sex

Table 28. Standard Strata Level IDs

OutTableID	LEVELID	LEVELVARS
t1_cida t2_cida t2_multevent t2_overlap	000	<BLANK>
t1_cida t2_cida t2_multevent t2_overlap	001	year
t1_cida t2_cida t2_multevent t2_overlap	002	sex
t1_cida t2_cida t2_multevent t2_overlap	003	agegroup
t1_cida t2_cida t2_multevent t2_overlap	004	sex agegroup
t1_cida t2_cida t2_multevent t2_overlap	005	sex agegroup year
t1_cida t2_cida t2_multevent t2_overlap	006	sex agegroup year month
t1_cida t2_cida t2_multevent t2_overlap	007	agegroup year
t1_cida t2_cida t2_multevent t2_overlap	008	agegroup year month
t1_cida t2_cida t2_multevent t2_overlap	009	sex year
t1_cida t2_cida t2_multevent t2_overlap	010	sex year month
t1_cida t2_cida t2_multevent t2_overlap	011	year month
t1_cida t2_cida t2_multevent t2_overlap	020	zip3

OutTableID	LEVELID	LEVELVARS
t1_cida t2_cida t2_multevent t2_overlap	021	zip3 zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	022	zip3 sex
t1_cida t2_cida t2_multevent t2_overlap	023	zip3 sex zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	024	zip3 agegroup
t1_cida t2_cida t2_multevent t2_overlap	025	zip3 agegroup zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	026	zip3 year
t1_cida t2_cida t2_multevent t2_overlap	027	zip3 year zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	028	zip3 race
t1_cida t2_cida t2_multevent t2_overlap	029	zip3 race zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	030	zip3 hispanic
t1_cida t2_cida t2_multevent t2_overlap	031	zip3 hispanic zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	040	state
t1_cida t2_cida t2_multevent t2_overlap	041	state zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	042	state sex
t1_cida t2_cida t2_multevent t2_overlap	043	state sex zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	044	state agegroup
t1_cida t2_cida t2_multevent t2_overlap	045	state agegroup zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	046	state year
t1_cida t2_cida t2_multevent t2_overlap	047	state year zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	048	state race
t1_cida t2_cida t2_multevent t2_overlap	049	state race zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	050	state hispanic
t1_cida t2_cida t2_multevent t2_overlap	051	state hispanic zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	060	zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	061	zip_uncertain sex
t1_cida t2_cida t2_multevent t2_overlap	062	zip_uncertain agegroup
t1_cida t2_cida t2_multevent t2_overlap	063	zip_uncertain year
t1_cida t2_cida t2_multevent t2_overlap	070	hhs_reg
t1_cida t2_cida t2_multevent t2_overlap	071	hhs_reg zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	072	hhs_reg sex
t1_cida t2_cida t2_multevent t2_overlap	073	hhs_reg sex zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	074	hhs_reg agegroup
t1_cida t2_cida t2_multevent t2_overlap	075	hhs_reg agegroup zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	076	hhs_reg year
t1_cida t2_cida t2_multevent t2_overlap	077	hhs_reg year zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	078	hhs_reg race
t1_cida t2_cida t2_multevent t2_overlap	079	hhs_reg race zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	080	hhs_reg hispanic
t1_cida t2_cida t2_multevent t2_overlap	081	hhs_reg hispanic zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	090	cb_reg
t1_cida t2_cida t2_multevent t2_overlap	091	cb_reg zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	092	cb_reg sex
t1_cida t2_cida t2_multevent t2_overlap	093	cb_reg sex zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	094	cb_reg agegroup

OutTableID	LEVELID	LEVELVARS
t1_cida t2_cida t2_multevent t2_overlap	095	cb_reg agegroup zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	096	cb_reg year
t1_cida t2_cida t2_multevent t2_overlap	097	cb_reg year zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	098	cb_reg race
t1_cida t2_cida t2_multevent t2_overlap	099	cb_reg race zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	100	cb_reg hispanic
t1_cida t2_cida t2_multevent t2_overlap	101	cb_reg hispanic zip_uncertain
t1_cida t2_cida t2_multevent t2_overlap	110	race
t1_cida t2_cida t2_multevent t2_overlap	111	race sex
t1_cida t2_cida t2_multevent t2_overlap	112	race agegroup
t1_cida t2_cida t2_multevent t2_overlap	113	race year
t1_cida t2_cida t2_multevent t2_overlap	114	race year month
t1_cida t2_cida t2_multevent t2_overlap	115	hispanic
t1_cida t2_cida t2_multevent t2_overlap	116	hispanic sex
t1_cida t2_cida t2_multevent t2_overlap	117	hispanic agegroup
t1_cida t2_cida t2_multevent t2_overlap	118	hispanic year
t1_cida t2_cida t2_multevent t2_overlap	119	hispanic year month

Table 29. Valid Stratification Variables for a Type 1 Analysis (Background Rates)

Variable Name	MSOC Output Tables	
	t1cida	t1censor
agegroup	X	X
cb_reg	X	
censdays_value		X
censdays_value_cat		X
covar*n*	X	
hhs_reg	X	
hispanic	X	
month	X	
race	X	
sex	X	X
state	X	
year	X	X
zip_uncertain	X	
zip3	X	

Table 30. Valid Stratification Variables for a Type 2 Analysis (Exposure and Follow-up Time)

Variable Name	Standard Type 2 MSOC Output Tables		Concomitant Use, Multiple Events, Overlap MSOC Output Tables			
	t2_cida	t2_censor	t2_conc	t2_multevent	t2_epigap	t2_overlap
agegroup	X	X	X	X	X	X
cb_reg	X					
censdays_value		X				
censdays_value_cat		X				
covarn	X					
Event_Flag		X				
hhs_reg	X		X	X		X
hispanic	X		X	X		X
month	X		X	X	X	X
race	X		X	X		X
sex	X	X	X	X	X	X
state	X			X		
year	X	X	X	X	X	X
zip_uncertain	X		X	X		X
zip3	X		X	X		X
Epi_gap					X	
tte_cat				X		X
Epi_count				X		
Time_to_epi				X		
Adherence				X		X
Adherence_#				X		
Total_days_overlap						X

Table 31. Valid Stratification Variables for a Type 3 Analysis (SCRI)

Variable Name	MSOC Output Table
	t3cida
agegroup	X
cb_reg	X
covarn	X
hhs_reg	X
hispanic	X
month	X
race	X
sex	X
state	X
TTC	X
TTE	X
year	X
zip_uncertain	X
zip3	X

Table 32. Valid Stratification Variables for a Type 4 Analysis (Pregnancy)

Variable Name	MSOC Output Tables				
	t4elig	t4preg	t4nopreg	t4preggestwk	T4nopreggestwk
agegroup		X	X		
cb_reg		X	X		
covarn		X	X		
eligdays	X				
eligible	X				
hhs_reg		X	X		
hispanic		X	X		
MoiName		X	X	X	X
pregdurcode			X		
prepostind		X	X		
prepostind	X	X	X		
race		X			
sex		X	X		
state		X	X		
year		X	X		
zip_uncertain		X	X		
zip3		X	X		

Table 33. Valid Stratification Variables for a Type 5 Analysis (Drug Utilization)

Variable Name	MSOC Output Tables					
	t5disp	t5first	t5alleps	t5episdur	t5censor	t5gaps
agegroup	X	X	X	X		X
cb_reg	X	X	X	X		X
covar*n*		X	X	X	X	X
DaySupp	X					
episodeLength				X	X	
episodeNum				X	X	
gaplength						X
gapnum						X
hhs_reg	X	X	X	X		X
hispanic	X	X	X	X		X
mntsfromstart		X	X			
race	X	X	X	X		X
sex	X	X	X	X		X
state	X	X	X	X		X
zip_uncertain	X	X	X	X		X
zip3	X	X	X	X		X

Table 34. Valid Stratification Variables for a Type 6 Analysis (Utilization and Switching)

Variable Name	Type 6 MSOC Output Tables						Switching MSOC Output Tables		
	t6counts	t6disp	t6episdur	t6censor	t6uptake	t6trend	t6switch episdur	t6plota	t6plotb
agegroup	X	X	X	X	X	X	X	X	X
DaySupp		X							
CumEpisode Length			X						
Uptakedays					X				
EpisodeLength				X			X		
TTSwitch								X	X
hhs_reg	X	X	X	X	X	X	X	X	X
hispanic	X	X	X	X	X	X	X	X	X
race	X	X	X	X	X	X	X	X	X
sex	X	X	X	X	X	X	X	X	X
state	X	X	X	X	X	X	X	X	X
year	X	X	X	X	X	X	X	X	X
zip_uncertain	X	X	X	X	X	X	X	X	X
zip3	X	X	X	X	X	X	X	X	X
month	X	X	X	X	X	X	X	X	X

k) Inclusion/Exclusion Codes File

The Inclusion/Exclusion Codes File is optional. It contains the comprehensive set of codes used to define additional cohort inclusion/exclusion criteria (*e.g.,* restrict cohort to individuals with evidence of a pre-existing condition 183 days before the index date). NDCs, ICD procedure and diagnosis codes, HCPCS codes, and/or laboratory result values can be used in any combination and can be restricted to specific care settings and diagnosis code positions (*e.g.,* principal discharge diagnoses only). The evaluation window to assess criteria can vary by code. Table 35 contains detailed specifications for this file.

Table 35. INCLUSIONCODES Specification

Parameter	Field Name	Description
Name of Cohort	GROUP	**Details:** standardized name used to differentiate cohorts. **Note 1**: multiple cohorts can be defined within the same Inclusion/Exclusion Codes File. In this case all cohorts are queried independently and results are reported separately and labeled using each GROUP name specified. **Note 2:** GROUP is the primary key linking cohorts across input files; GROUP values must match (including case) between the INCLUSIONCODES file and other input files. **Note 3**: GROUP values must remain consistent during the course of a surveillance activity. **Note 4**: GROUP values must match ANALYSISGRP in Treatment Pathways File when applying switch pattern inclusion/exclusion criteria. **Named by:** Request programmer **Input type:** Required **Format:** Alphanumeric; SAS character $40; no special characters (*e.g.,* commas, periods, hyphens, etc.) allowed, and underscores must be used to mark spaces. **Example:** Insulin
Name of Stockgroup within the Cohort	STOCKGROUP	**Details:** standardized name used to refer to a specific inclusion/exclusion code/criterion within a given GROUP. **Note 1:** the STOCKGROUP field is used by the stockpiling algorithm as group categories to adjust service dates. **Note 2:** useful when a GROUP contains multiple exposures of interest. For example, if GROUP= "Insulin" STOCKGROUP could take values of "Insulin_Oral" and "Insulin_Injectable".

Parameter	Field Name	Description
		Note 3: no output will be presented by STOCKGROUP. All output is presented at the GROUP level. **Named by:** Requester **Input type:** Required **Format:** Alphanumeric; SAS character $30; special characters (*e.g.,* commas, periods, hyphens, etc.) allowed and underscores must be used to mark spaces. **Example:** Insulin_Oral
Code Category	CODECAT	**Details:** type of each code category value included in the CODETYPE field (below) of this file. Valid values include: • **RX:** NDC • **DX:** Diagnosis code • **PX:** Procedure code • **LB:** Lab code **Defined by:** Requester **Input type:** Required **Format:** SAS character $2. **Example:** DX
Code Type	CODETYPE	**Details:** type of each code value included in the CODE field (below) of this file. Valid values include: If CODECAT = RX: • **09**: 9 digits NDC • **11**: 11 digits NDC If CODECAT = DX: • **09**: ICD-9-CM • **10**: ICD-10-CM • **11:** ICD-11-CM • **OT:** Other If CODECAT = PX: • **09**: ICD-9-CM • **10**: ICD-10-CM • **11**: ICD-11-CM • **C4**: CPT-4 (*i.e.,* HCPCS Level I) • **HC**: HCPCS (*i.e.,* HCPCS Level II) • **H3**: HCPCS Level III • **C2**: CPT Category II • **C3**: CPT Category III • **ND**: 11-digit NDC • **RE**: Revenue • **LO**: Local homegrown

Parameter	Field Name	Description

		• **OT**: Other
		
		If CODECAT = LB:
		• **01N**: extract quantitative lab test result using SOC-defined lab code (see Lab Code Lookup Table)
		• **02N**: extract quantitative lab test result using LOINC
		• **'px'N**: extract quantitative lab test result using CPT
		  • **09N**: ICD-9-CM
		  • **10N**: ICD-10-CM
		  • **11N**: ICD-11-CM
		  • **C4N**: CPT-4 (*i.e.,* HCPCS Level I)
		  • **HCN**: HCPCS (*i.e.,* HCPCS Level II)
		  • **H3N:** HCPCS Level III
		  • **C2N:** CPT Category II
		  • **C3N:** CPT Category III
		  • **NDN:** 11-digit NDC
		  • **REN:** Revenue
		  • **LON:** Local homegrown
		• **01C:** extract qualitative lab test result using SOC-defined lab code (see Lab Code Lookup Table)
		• **02C:** extract qualitative lab test result using LOINC
		• **'px'C:** extract qualitative lab test result using the following codes
		  • **09C**: ICD-9-CM
		  • **10C**: ICD-10-CM
		  • **11C**: ICD-11-CM
		  • **C4C**: CPT-4 (*i.e.,* HCPCS Level I)
		  • **HCC**: HCPCS (*i.e.,* HCPCS Level II)
		  • **H3C:** HCPCS Level III
		  • **C2C:** CPT Category II
		  • **C3C:** CPT Category III
		  • **NDC:** 11-digit NDC
		  • **REC:** Revenue
		  • **LOC:** Local homegrown
		
		Note 1: as the LOINC field is not populated by all Data Partners in the SCDM Laboratory Result table and the CPT code may not be specific to a particular lab test, it is strongly recommended that the Laboratory Result table be queried using SOC-defined lab codes.

Parameter	Field Name	Description
		Defined by: Requester **Input type:** Required **Format:** Alphanumeric; SAS character $3. **Example:** 09
Code	CODE	**Details:** NDC, procedure, diagnosis, and/or lab code of interest. **Note 1**: Codes are matched using exact values (*i.e.,* 3-digit code lookup requires an exact 3-digit code match). Wildcard match (*) functionality is also available for ICD-9 diagnosis codes (*e.g.,* querying "250*0" would be used to find any ICD-9-CM diagnosis codes for diabetes type II, or "250**" to find ICD-9-CM diagnosis codes for all diabetes codes in the range "250.00 - 250.99"). To get "starts with" codes, the user will have to specify 250, 250*, 250**. **Note 2**: For NDCs, either 9 or 11 digit codes can be entered. **Note 3:** remove decimal points in the code value. **Note 4**: CODETYPE/CODECAT must be consistent with the expected format of the CODE value (*e.g.,* the program will not find any valid matches in the data for CODECAT=RX, CODETYPE=11 and a 9-digit NDC value). **Note 5**: Duplicate CODECAT-CODETYPE-CODE-CARESETTING-PRINCIPAL combinations are removed by the MP algorithm. **Note 6**: 'V' and 'E' ICD-9-CM diagnosis codes must be specified using uppercase 'V' and 'E'. **Defined by:** Requester, with support from the SOC as needed **Input type:** Required **Format:** Alphanumeric; SAS character $11. **Example: (CODECAT=NDC; CODETYPE=11):** 12345678911
Care Setting and Diagnosis Position Requirements	CARESETTINGPRINCIPAL	**Details:** defines the care setting and principal diagnosis position requirements for each code. This field uses combination(s) of the SCDM variables care setting (ENCTYPE) and principal discharge diagnosis flag (PDX) to restrict the observance of codes to those in the requested care settings and with the requested diagnosis position. If no restrictions are required (*e.g.,* requester wants all care settings and any value of PDX), leave the field blank. The following

Parameter	Field Name	Description
		are valid entries; all entries must be in single quotes and separated by a space: • **IPP:** inpatient hospital stays, principal diagnoses • **IPS:** inpatient hospital stays, secondary diagnoses • **IPX:** inpatient hospital stays, unclassified diagnoses • **ISP:** non-acute institutional stays, principal diagnoses • **ISS:** non-acute institutional stays, secondary diagnoses • **ISX:** non-acute institutional stays, unclassified diagnoses • **ED*:** emergency department encounters • **AV*:** ambulatory visits • **OA*:** other ambulatory visits **Request Programmer Note 1:** the wildcard symbol (*) can be used to represent "any" values of either care setting or principal discharge diagnosis flag. For example, CARESETTINGPRINCIPAL = 'IP*' will restrict codes to those observed in the inpatient setting irrespective of the principal diagnosis flag value. CARESETTINGPRINCIPAL = '**P' will restrict diagnosis codes to those in the principal position, irrespective of the care setting. **Request Programmer Note 2:** the principal discharge diagnosis flag is only relevant for diagnosis codes. All other codes should use the * wildcard for the third digit of the CARESETTINGPRINCIPAL value. **Note 3:** CARESETTINGPRINCIPAL is allowed to vary between CODEs within the same GROUP. For example, CARESETTINGPRINCIPAL is allowed to equal 'IPP' for one diagnosis code and 'IPP' 'EDP' for another diagnosis code *in the same GROUP*. **Defined by:** Requester **Input type:** Optional; Default: blank (*i.e.,* no restrictions) **Format:** Alphanumeric **Example:** 'IPX' 'ED*' '**P'
Condition Exclusion Indicator	CONDINCLUSION	**Details:** indicates whether each criterion specified (i.e., CONDLEVEL value) is for an inclusion (=1) or exclusion (=0) criterion.

Parameter	Field Name	Description
		Note 1: exclusion criteria (CONDINCLUSION=0) require continuous enrollment during the CONDFROM – CONDTO period (below). **Note 2**: within GROUP values, CONDINCLUSION = 0 and CONDINCLUSION = 1 criteria are separated by an "and" operator. For example, in a scenario with 1) CONDLEVEL = "Diabetes" and CONDINCLUSION=1; and 2) CONDLEVEL = "Heart_Failure" and CONDINCLUSION=0, the program will require presence of Diabetes **and** absence of Heart Failure for a patient to be eligible for cohort entry. **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** 1
Name of inclusion/exclusion condition	CONDLEVEL	**Details:** requester-defined name to represent a unique inclusion or exclusion criterion. **Note 1**: within GROUP and CONDINCLUSION values, CONDLEVEL values indicate criteria separated by an "or" operator. For example, in a scenario with 1) CONDLEVEL = "Diabetes" and CONDINCLUSION=1; and 2) CONDLEVEL = "Heart_Failure" and CONDINCLUSION=1, the program will require presence of Diabetes **or** presence of Heart Failure for a patient to be eligible for cohort entry. **Named by:** Requester **Input type:** Required **Format:** Character **Example:** Diabetes
Name of inclusion/exclusion sub-condition	SUBCONDLEVEL	**Details:** requester-defined name to represent unique inclusion or exclusion criteria **within CONDLEVEL values.** Allows requesters to define an individual inclusion/exclusion criterion (e.g., Diabetes) using a complex algorithm (e.g., diagnosis codes **and** laboratory result values). **Note 1**: within GROUP, CONDINCLUSION, and CONDLEVEL values, criteria specified with the same SUBCONDLEVEL value are separated by an "and" operator. For example, in a scenario with 1) CONDLEVEL = "Diabetes", CONDINCLUSION=1, SUBCONDLEVEL= "diagnoses" and SUBCONDINCLUSION=1; and 2) CONDLEVEL = "Diabetes", CONDINCLUSION=1, SUBCONDLEVEL=

Parameter	Field Name	Description
		"HGBA1C" and SUBCONDINCLUSION=1, the program will define Diabetes as presence of a diagnosis code indicative of diabetes **and** a HGBA1C test result indicative of Diabetes. **Named by:** Requester **Input type:** Required **Format:** Character **Example:** Diabetes_DX
Sub-condition Exclusion Indicator	SUBCONDINCLUSION	**Details:** indicates whether each SUBCONDLEVEL criterion is for an inclusion (=1) or exclusion (=0) criterion. **Note 1:** exclusion criteria (SUBCONDINCLUSION=0) require continuous enrollment during the CONDFROM – CONDTO period (below). **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** 1
Evaluation Period Start	CONDFROM	**Details:** used in combination with CONDTO (below). CONDFROM defines the start of the evaluation period for each CODE value specified, expressed in terms of "days from Index Date". For example, if Index Date=01/08/2009 and CONDFROM for a given condition code is set to -7, the MP algorithm will start looking for that condition code on 01/01/2009. **Note 1**: individual CODE values within a same GROUP are allowed to have different evaluation periods and therefore have different CONDFROM and CONDTO values. **Note 2**: the index date is "day zero". Therefore, if zero is included in the CONDFROM-CONDTO interval for a given CODE value, the index date is included in the evaluation period. **Note 3**: if CONDFROM > 0 then the evaluation period will start after the index date. **Note 4: special case**: when CONDFROM = missing the program considers an inclusion/exclusion criterion met if the member has no evidence of the exposure in their entire available history before the value of CONDTO. In this case, for an exclusion criterion, continuous enrollment is *not required* for the duration of the evaluation period (only explicitly

Parameter	Field Name	Description
		defined enrollment criteria, e.g., specified using the ENRDAYS value, are required). **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** -180
Evaluation Period End	CONDTO	**Details:** used in combination with CONDFROM (above). CONDTO defines the end of the evaluation period for each CODE value specified, expressed in terms of "days from Index Date". For example, if Index Date=01/08/2009 and CONDTO for a given condition code is set to -1, the MP algorithm will look for that condition code between the CONDFROM date through 01/07/2009. **Note 1**: individual CODE values within the same GROUP are allowed to have different evaluation periods and therefore have different CONDFROM and CONDTO values. **Note 2**: the index date is "day zero". Therefore if zero is included in the CONDFROM-CONDTO interval for a given CODE value the index date is included in the evaluation period. **Note 3: special case**: when CONDTO = missing the program considers an inclusion/exclusion criterion met if the member has no evidence of the exposure in their entire available history after the index date. In this case, for an exclusion criterion, continuous enrollment is *not required* for the duration of the evaluation period (only explicitly defined enrollment criteria, e.g., specified using the ENRDAYS value, are required). **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** -1
Indicates the number of instances for the condition	CODEDAYS	**Details:** sets the minimum number of times SUBCONDLEVEL should be found in the baseline period. **Note**: multiple codes identified on the same day will only count once (i.e., count code days not code instances). **Defined by:** Requester **Input type:** Required **Format:** Numeric

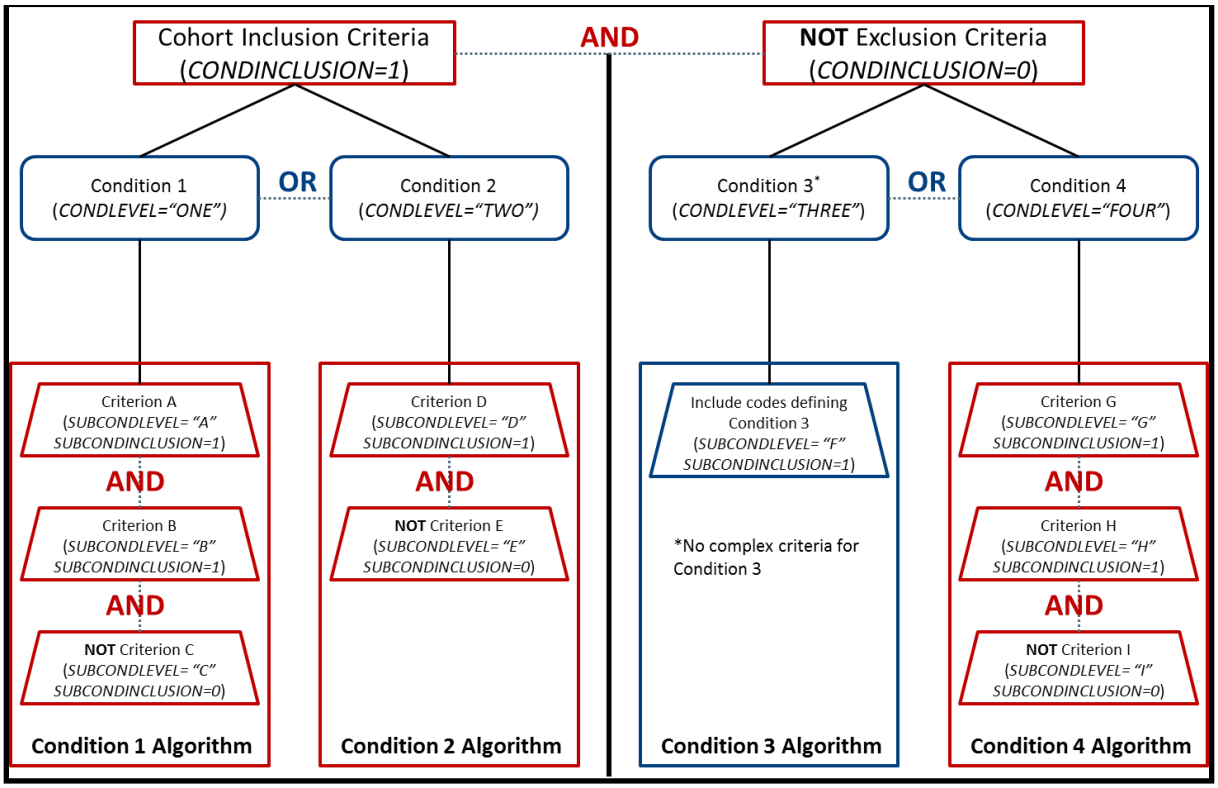
Parameter	Field Name	Description
		Example: 1 (default)
Code Supply	CODESUPPLY	**Description:** indicates, for each code listed in the file, a forced supply that should be attached to the code. The specified code suppled will replace RxSup for RX codes. **Note 1**: Non-RX codes are not stockpiled **Named by:** Requester **Input type:** Optional **Format:** Numeric **Example:** 10
Exclude Day Supply	EXCLUDESUPPLY	**Description:** indicates if lookback period to define inclusion/exclusion criteria looks for evidence of a dispensing or evidence of days supply. Allowable values : • **N:** lookback period should search for evidence of days supply • **Y:** lookback period should search for evidence of a dispensing date **Note 1:** Each CONDLEVEL can only be associated with one EXCLUDESUPPLY value **Named by:** Requester **Input type:** Optional **Format:** Character (1)
Source population, mother or infant, for which code applies	CODEPOP	**Details:** in queries restricting Mother-Infant pregnant cohorts for inferential analyses, , population source to query for code. Valid values include: • **M**: indicates should look in mother claims only • **I**: indicates should look in infant claims only • **MI**: MI indicates should look in mother or infant claims • **<blank>:** note relevant to query **Named by:** Request programmer **Input type:** Required **Format:** Character (2) **Example:** M
Index date to anchor inclusion/exclusion criteria	INDEXDATE	**Details:** Identifies which index date to use to define inclusion or exclusion criteria in queries restricting Mother-Infant pregnant cohorts for inferential analyses. Valid values include: • **INDEXDT-** Index date of pregnancy start

Parameter	Field Name	Description
		• **INDEXDT_EXP** - First date of dispensing or procedure code of drug or comparator of interest/exposure. • **INDEXDT_DELIV** – Index date of delivery admission date • **<blank>** - exclusion/inclusion criteria will be applied during initial cohort extraction **Note 1:** Only applicable for Type 4 inferential analyses (i.e., when an MILCOHORTFILE is specified). When left blank, exclusion/inclusion criteria will be applied during initial cohort extraction. When specified, criteria will only be applied to determine MIL cohorts and not pregnancy cohort in Type 4 metrics. **Note 2:** Enrollment criteria will be reassessed when CODEPOP=M or MI in the mother's claim for the period CONDFROM-INDEXDATE **Note 3:** When CODEPOP = I or MI and searching in the infant's claims, INDEXDT_DELIV will start searching from the infant's birth date **Note 4:** When INDEXDATE=INDEXDT_EXP, and the exposure episode begins prior to the period of interest, indexdt_exp will be set to the start of the period of interest. **Defined by:** Requester **Input type:** Required (cannot be left blank) **Format:** character(30) **Example:** INDEXDT_EXP
Lab Date Selection Algorithm	RAWLABDATETYPE	**Details:** relevant for requests that query laboratory result values. Field specifies in what sequence date(s) in the SCDM Laboratory Result table should be considered to select one relevant date for a laboratory result of interest. The parameter will allow the requester to either specify 1) a single date variable (Lab_dt, Order_dt, or Result_dt) to use; or 2) a hierarchy to choose a date variable (*e.g.,* select Lab_dt else if missing select Result_dt else if missing select Order_dt). Valid values are any combination of the following: • **L:** Lab Date • **O:** Order Date • **R:** Result Date

Parameter	Field Name	Description
		Note 1: leave blank if CODECAT ≠ LB. **Defined by:** Requester **Input type:** Required **Format:** SAS character $3. **Example 1:** RawLabDateType =LRO. In this case, the program will use Lab_dt else if missing use Result_dt else if missing use Order_dt. **Example 2:** RawLabDateType =L. In this case, the program will use Lab_dt only.
Lab Result Values	RAWLABRESULT	**Details:** specifies the lab result value or lab result range for querying. RAWLABRESULT allows for values or ranges of quantitative laboratory results (*e.g.,* 100; 100-200) and values of qualitative laboratory results (*e.g.,* "POSITIVE"). Valid values are: • <=X (less than or equal to X) • <X (less than X) • >=X (greater than or equal to X) • >X (greater than X) • ~=X (not equal to X) • X:Y (between X and Y) Any string of relevant characters is allowed for qualitative results querying. **Note 1:** leave blank if CODECAT ≠ LB. **Note 2:** there are two fields in the Laboratory Result table that include results: MS_Result_C (contains results for qualitative tests) and MS_Result_N (contains results for quantitative tests). The field where the result will be queried will depend on the RAWCODETYPE value. **Note 3:** ranges cannot be specified with hyphens. Must use ":". **Defined by:** Requester Input type: Required for laboratory results **Format:** Alphanumeric; **Example 1:** RawLabResult=20:50 **Example 2:** RawLabResult=POSITIVE

Specifying the Inclusion/Exclusion Codes File, and understanding the relationships among parameters, can be challenging. Figure 39 describes file parameters and the interactions between "and" and "or" operators. The fictitious example includes two inclusion criteria and two exclusion criteria. Each inclusion criterion is defined using a complex algorithm (i.e., Condition 1 is defined as Criterion A and Criterion B and not Criterion C); one of the two exclusion criteria is defined using a complex algorithm (i.e., Condition 4 is defined as Criterion G and Criterion H and not Criterion I). If a patient meets the definition of Condition 4, they are excluded from the cohort.

Figure 39. Relationship Between Parameters Defining Inclusion/Exclusion Criteria



In terms of creating input files, the following two examples demonstrate how input files should be created to ensure different inclusion/exclusion criteria.

Example 1:

Inclusion criteria: Condition A or Condition B
Exclusion criteria: Condition C and Condition D

Group	Stockgroup	CondInclusion	CondLevel	SubcondLevel	Subcondinclusion
Group A	Condition A	1	Cond1	SubCond1	1
Group A	Condition B	1	Cond2	Subcond1	1
Group A	Condition C	0	Cond3	Subcond2	1
Group A	Condition D	0	Cond3	Subcond3	1

Example 2:

Inclusion criteria: Condition A and not Condition B
Exclusion criteria: (Condition C and Condition D) or (Condition E and Condition F)

Group	Stockgroup	Condinclusion	CondLevel	SubcondLevel	Subcondinclusion
Group A	Condition A	1	Cond1	SubCond1	1
Group A	Condition B	1	Cond1	Subcond2	0
Group A	Condition C	0	Cond3	Subcond3	1
Group A	Condition D	0	Cond3	Subcond4	1
Group A	Condition E	0	Cond4	Subcond5	1
Group A	Condition F	0	Cond4	Subcond6	1

l) Covariate Codes File

The Covariate Codes File is required for requests evaluating the presence of covariates at baseline and requests using covariates for analytic adjustment. Otherwise, it should not be specified. NDCs, ICD procedure and diagnosis codes, and HCPCS codes can be used in any combination and can be restricted to specific care settings and diagnosis code positions (*e.g.,* principal discharge diagnoses only).

Only one Covariate Codes File can be specified per execution of the CIDA tool. Therefore, all cohorts (GROUPs) included in a CIDA tool execution will extract information for the same covariates.

The Covariate Codes File also asks requesters to specify which covariates may be used for future subgroup analyses. This information is necessary to inform the PSA tool what information must be extracted and saved to the MSOC folder. Table 36 contains detailed specifications for this file.

Table 36. COVARIATECODES Specification

Parameter	Field Name	Description
Covariate Name	STUDYNAME	**Details:** unique name for each covariate defined in the file. **Defined by:** Request programmer **Input type:** Required **Format:** SAS character $50. **Example:** Diabetes
Name of Stockpiling Group within the Cohort	STOCKGROUP	**Details:** standardized name used to refer to a specific exposure/HOI within a given GROUP. **Note 1:** the STOCKGROUP field is used by the stockpiling algorithm as group categories to adjust service dates. **Note 2:** useful when a GROUP contains multiple exposures of interest. For example, if GROUP= "Insulin" STOCKGROUP could take values of "Insulin_Oral" and "Insulin_Injectable". **Note 3:** no output will be presented by STOCKGROUP. All output is presented at the GROUP level. **Named by:** Request programmer

Parameter	Field Name	Description
		Input type: Required **Format:** Alphanumeric; SAS character $30; special characters (*e.g.,* commas, periods, hyphens, etc.) allowed and underscores must be used to mark spaces. **Example:** Insulin_Oral
Numeric Covariate Indicator	COVARNUM	**Details:** a numeric indicator for each covariate specified, to identify covariates for further processing by the PSA tool and order of covariates in output. **Note 1:** each unique STUDYNAME should have a unique COVARNUM value. **Note 2:** COVARNUM must start at 1 and be incremented by 1 for each additional covariate. **Defined by:** Request programmer **Input type:** Required **Format:** Numeric **Example:** 11
Code Category	CODECAT	**Details:** type of each code category value included in the CODETYPE field (below) of this file. Valid values include: • **RX:** NDC • **DX:** Diagnosis code • **PX:** Procedure code • **CC:** indicates combination of multiple COVARNUM values to define a covariate. The referenced COVARNUM values to define the combination must be previously defined (i.e., be listed on rows preceding the CODECAT=CC row). **Defined by:** Requester **Input type:** Required **Format:** SAS character $2. **Example:** DX
Code Type	CODETYPE	**Details:** type of each code value included in the CODE field (below) of this file. Valid values include: If CODECAT = RX: • **09**: 9 digits NDC • **11**: 11 digits NDC If CODECAT = DX: • **09**: ICD-9-CM • **10**: ICD-10-CM • **11:** ICD-11-CM

Parameter	Field Name	Description
		• **OT:** Other If CODECAT = PX: • **09**: ICD-9-CM • **10**: ICD-10-CM • **11**: ICD-11-CM • **C4**: CPT-4 (*i.e.,* HCPCS Level I) • **HC**: HCPCS (*i.e.,* HCPCS Level II) • **H3**: HCPCS Level III • **C2**: CPT Category II • **C3**: CPT Category III • **ND**: 11-digit NDC • **RE**: Revenue • **LO**: Local homegrown • **OT**: Other If CODECAT = CC, leave this field blank. **Defined by:** Requester **Input type:** Required **Format:** Alphanumeric; SAS character $3. **Example:** 09
Code	CODE	**Details:** NDC, diagnosis, and/or procedure code of interest. **Note 1**: Codes are matched using exact values (*i.e.,* 3-digit code lookup requires an exact 3-digit code match). Wildcard match (*) functionality is also available for ICD-9 diagnosis codes (*e.g.,* querying "250*0" would be used to find any ICD-9-CM diagnosis codes for diabetes type II, or "250**" to find ICD-9-CM diagnosis codes for all diabetes codes in the range "250.00 - 250.99"). To get "starts with" codes, the user will have to specify 250, 250*, 250**. **Note 2**: For NDCs, either 9 or 11 digit codes can be entered. **Note 3:** remove decimal points in the code value. **Note 4**: CODETYPE/CODECAT must be consistent with the expected format of the CODE value (*e.g.,* the program will not find any valid matches in the data for CODECAT=RX, CODETYPE=11 and a 9-digit NDC value). **Note 5**: Duplicate CODECAT-CODETYPE-CODE-CARESETTING-PRINCIPAL combinations are removed by the MP algorithm.

Parameter	Field Name	Description
		Note 6: 'V' and 'E' ICD-9-CM diagnosis codes must be specified using uppercase 'V' and 'E'. **Note 7**: If CODECAT = CC, this field contains the algorithm for the combination of COVARNUM values, e.g., "1 and (2 or 3)" to describe an algorithm requesting presence of COVARNUM=1 and (COVARNUM=2 or COVARNUM=3). Combiniation covariates must be defined after all traditional covariates have been defined. (i.e., be listed on rows preceding all CODECAT=CC row). Any combination of "and", "or", and "not" may be used in combination with parentheses to define algorithms. Example algorithms (these are all different): • (1 or 9) and not 2 • 1 or 9 and not 2 • Not 2 and 1 or 9 Unless all expressions are of one type (i.e. 1 and 2 and 3), it is highly recommended to use parentheses for algorithm clarity (i.e., bullets 2 and 3 above should be avoided). **Defined by:** Requester, with support from the SOC as needed **Input type:** Required **Format:** Alphanumeric; SAS character $50. **Example1: (CODECAT=NDC; CODETYPE=11):** 12345678911
Care Setting and Principal Diagnosis Indicator	CARESETTINGPRINCIPAL	**Details:** defines the care setting and principal diagnosis position requirements for each code. This field uses combination(s) of the SCDM variables care setting (ENCTYPE) and principal discharge diagnosis flag (PDX) to restrict the observance of codes to those in the requested care settings and with the requested diagnosis position. If no restrictions are required (*e.g.,* requester wants all care settings and any value of PDX), leave the field blank. The following are valid entries; all entries must be in single quotes and separated by a space: • **IPP:** inpatient hospital stays, principal diagnoses • **IPS:** inpatient hospital stays, secondary diagnoses

Parameter	Field Name	Description
		• **IPX:** inpatient hospital stays, unclassified diagnoses • **ISP:** non-acute institutional stays, principal diagnoses • **ISS:** non-acute institutional stays, secondary diagnoses • **ISX:** non-acute institutional stays, unclassified diagnoses • **ED*:** emergency department encounters • **AV*:** ambulatory visits • **OA*:** other ambulatory visits **Request Programmer Note 1:** the wildcard symbol (*) can be used to represent "any" values of either care setting or principal discharge diagnosis flag. For example, CARESETTINGPRINCIPAL = 'IP*' will restrict codes to those observed in the inpatient setting irrespective of the principal diagnosis flag value. CARESETTINGPRINCIPAL = '**P' will restrict diagnosis codes to those in the principal position, irrespective of the care setting. **Request Programmer Note 2:** the principal discharge diagnosis flag is only relevant for diagnosis codes. All other codes should use the * wildcard for the third digit of the CARESETTINGPRINCIPAL value. **Note 3:** CARESETTINGPRINCIPAL is allowed to vary between CODEs within the same GROUP. For example, CARESETTINGPRINCIPAL is allowed to equal 'IPP' for one diagnosis code and 'IPP' 'EDP' for another diagnosis code *in the same GROUP*. **Note 4**: If CODECAT = CC, leave this field blank. **Defined by:** Requester **Input type:** Optional; Default: blank (*i.e.,* no restrictions) **Format:** Alphanumeric **Example:** 'IPX' 'ED*' '**P'
Covariate Evaluation Period Start	COVFROM	**Details:** used in combination with COVTO (below). COVFROM defines the start of the evaluation period for each CODE value specified, expressed in terms of "days from Index Date". For example, if Index Date=01/08/2009 and COVFROM for a given condition code is set to -7, the MP algorithm will start looking for that condition code on 01/01/2009.

Parameter	Field Name	Description
		Note 1: individual CODE values within a same GROUP are allowed to have different evaluation periods and therefore have different COVFROM and COVTO values. **Note 2**: the index date is "day zero". Therefore, if zero is included in the COVFROM-COVTO interval for a given CODE value, the index date is included in the evaluation period. **Note 3**: if COVFROM > 0 then the evaluation period will start after the index date. **Note 4: special case**: when COVFROM = missing the program considers a covariate met if the member has the code in their entire available history before the value of COVTO. In this case, continuous enrollment is *not required* for the duration of the evaluation period (only explicitly defined enrollment criteria, e.g., specified using the ENRDAYS value, are required). **Note 5**: If CODECAT = CC, leave this field blank. **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** -180
Covariate Evaluation Period End	COVTO	**Details:** used in combination with COVFROM (above). COVTO defines the end of the evaluation period for each CODE value specified, expressed in terms of "days from Index Date". For example, if Index Date=01/08/2009 and COVTO for a given condition code is set to -1, the MP algorithm will look for that condition code between the COVFROM date through 01/07/2009. **Note 1**: individual CODE values within the same GROUP are allowed to have different evaluation periods and therefore have different COVFROM and COVTO values. **Note 2**: the index date is "day zero". Therefore if zero is included in the COVFROM-COVTO interval for a given CODE value the index date is included in the evaluation period. **Note 3: special case**: when COVTO = missing the program considers a covariate met if the member has the code in their entire available history after the index date. In this case, continuous enrollment is *not required* for the duration of the evaluation period

Parameter	Field Name	Description
		(only explicitly defined enrollment criteria, e.g., specified using the ENRDAYS value, are required). **Note 4**: If CODECAT = CC, leave this field blank. **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** -1
Indicator that Covariate is Needed for Subgroup Analysis	KEEP	**Details:** indicates if the variable is needed for additional analyses. **Valid values include:** • **0:** Do not keep information needed for additional analyses/covariate will not be used in additional analyses • **1:** Keep information needed for additional analyses **Note 1:** if KEEP=1, covariate information will be output to the [*RUNID*] _matched_ [*COMP_ORDER*] _ [*PERIODID*].sas7bdat output file. **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** 0 (default)
Indicates the number of instances for the condition	CODEDAYS	**Details:** sets the minimum number of times COVARNUM should be found in the baseline period. **Note:** multiple codes identified on the same day will only count once (i.e., count code days not code instances). **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** 1 (default)
Code Supply	CODESUPPLY	**Description:** indicates, for each code listed in the file, a forced supply that should be attached to the code. The specified code supply will replace the RxSup for RX codes. **Note 1:** Non-RX codes are not stockpiled. **Named by:** Requester **Input type:** Optional **Format:** Numeric **Example:** 10

Parameter	Field Name	Description
Exclude Day Supply	EXCLUDESUPPLY	**Description:** indicates if lookback period to define inclusion/exclusion criteria looks for evidence of a dispensing or evidence of days supply. Allowable values : • **N:** lookback period should search for evidence of days supply • **Y:** lookback period should search for evidence of a dispensing date **Named by:** Requester **Input type:** Optional **Format:** Character (1)

m) Comorbidity Score File

The Comorbidity Score File is required if a request is calculating the Charlson/Elixhauser combined comorbidity score (required for propensity score matched analyses). If the file is specified, the Comorbidity Score Codes Lookup Table must be included in the *inputfiles* folder. Specification of this file requests calculation of a combined Charlson/Elixhauser comorbidity score.

If the output of the CIDA tool is used for subsequent propensity score matched analyses using the PSA tool, the comorbidity score is available to estimate the propensity score along with the other covariates in the Covariate Codes File.

The Comorbidity Score File uses two parameters, COMORBFROM and COMORBTO, to define periods for observing the medical condition of interest and calculating the two scores.. When COMORBTO = 0, the index date is included in the period for observing the medical condition of interest. When COMORBTO > 0, post-index days are included.

Table 37 contains detailed specifications for the Comorbidity Score File.

Table 37. COMORBFILE Specification

Parameter	Field Name	Description
Name of Cohort	GROUP	**Details:** standardized name used to differentiate cohorts. **Note 1**: multiple cohorts can be defined within the same Comorbidity Score File. In this case all cohorts are queried independently and results are reported separately and labeled using each GROUP name specified. **Note 2:** GROUP is the primary key linking cohorts across input files; GROUP values must match (including case) between the COMORBFILE file and other input files. **Note 3**: GROUP values must remain consistent during the course of a surveillance activity. **Note 4:** For concomitant use, multiple events, and overlap analyses, GROUP must match ANALYSISGRP

Parameter	Field Name	Description
		Named by: Request programmer **Input type:** Required **Format:** Alphanumeric; SAS character $40; no special characters (*e.g.,* commas, periods, hyphens, etc.) allowed, and underscores must be used to mark spaces. **Example:** Insulin
Comorbidity Score Evaluation Period Start	COMORBFROM	**Details:** defines the start of the comorbidity score evaluation period, expressed in terms of "days from Index Date". For example, if Index Date=01/08/2009 and COMORBFROM is set to -7, the algorithm will start evaluating the score on 01/01/2009. **Note 1:** when COMORBFROM = missing, the program evaluates comorbidities in the entire available patient history before the value of COMORBTO. **Note 2:** An individual is not required to be enrolled during the entire COMORBFROM-COMORBTO period. **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** -180
Comorbidity Score Evaluation Period End	COMORBTO	**Details:** used in combination with COMORBFROM (above). COMORBTO defines the end of the comorbidity score evaluation period, expressed in terms of "days from Index Date". For example, if Index Date=01/08/2009 and COMORBTO is set to -1, the algorithm will evaluate the score between the COMORBFROM date through 01/07/2009. **Note 1:** the index date is "day zero". Therefore, if zero is included in the COMORBFROM–COMORBTO interval, the index date is included in the evaluation period. If COMORBTO ≥1, the evaluation period will include days after the index date. **Note 2**: when COMORBTO = missing, the program evaluates comorbidities in the entire available patient history before the index date. **Note 3:** An individual is not required to be enrolled during the entire COMORBFROM-COMORBTO period. **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** 0

Parameter	Field Name	Description
Comorbidity Score Stratification Groups	CCIGROUP	**Details:** the grouping to apply for stratification of the combined comorbidity scores. Groups must be separated by a space and "+" used to make the last group open ended. To leave the first group open-ended, use "low-". In the output "low-", will be replaced with "<=". If a negative is desired as the upper bound of a group, do not include a space in the group (*e.g.,* use low--1 for low to -1). Note that groups should not have overlapping values. In the event that overlapping values are entered, the value will be mapped to the first group in the list. **Note 1:** the theoretical range of scores is -2 -26. **Defined by:** Requester **Input type:** Required **Format:** Alphanumeric; SAS character $20. **Example:** low-0 1 2-3 4-7 8+

n) Utilization File

The Utilization File is required if a request is calculating medical or drug utilization metrics. Specification of this file will:

- Calculate three drug utilization metrics: 1) number of dispensings; 2) number of unique generics dispensed; and 3) number of unique drug classes dispensed during a requester-defined number of days around the index date and output these values.
- Calculate a single medical utilization metric (*i.e.,* number of medical encounters per individual during a requester-defined period of time) and stratify CIDA tool output metrics by requester-defined groupings of this metric. If the output of the CIDA tool is used for subsequent analyses (e.g., ANALYSIS=PS or ADS), CIDA will also calculate the number of medical encounters per individual *per encounter type* (*i.e.,* SCDM variable ENCTYPE in the Encounter table). As there are five ENCTYPE values in the SCDM Encounter table (AV = Ambulatory Visit, ED = Emergency Department, IP = Inpatient Hospital Stay, IS = Non-Acute Institutional Stay, and OA = Other Ambulatory Visit [*e.g.,* home health visits, telemedicine, telephone and email consultations]), 5 variables will be output for further analysis.

Table 38 contains detailed specifications for this file.

Table 38. UTILFILE Specification

Parameter	Field Name	Description
Name of Cohort	GROUP	**Details:** standardized name used to differentiate cohorts. **Note 1**: multiple cohorts can be defined within the same Utilization File. In this case all cohorts are queried independently and results are reported separately and labeled using each GROUP name specified. **Note 2:** GROUP is the primary key linking cohorts across input files; GROUP values must match (including case) between the UTILFILE file and other input files. **Note 3**: GROUP values must remain consistent during the course of a surveillance activity. **Note 4:** For concomitant use, multiple events, and overlap analyses, GROUP must match ANALYSISGRP **Named by:** Request programmer **Input type:** Required **Format:** Alphanumeric; SAS character $40; no special characters (*e.g.,* commas, periods, hyphens, etc.) allowed, and underscores must be used to mark spaces. **Example:** Insulin
Medical Utilization Evaluation Period Start	MEDUTILFROM	**Details:** used in combination with MEDUTILTO (below). MEDUTILFROM defines the start of the medical utilization evaluation period, expressed in terms of "days from Index Date". For example, if Index Date=01/08/2009 and MEDUTILFROM is set to -7, the MP algorithm will start evaluating drug utilization on 01/01/2009. **Note 1**: the index date is "day zero". Therefore, if zero is included in the MEDUTILFROM – MEDUTILTO interval, the index date is included in the evaluation period. **Note 2**: if MEDUTILFROM > 0 then the evaluation period will start after the index date. **Note 3: special case**: when MEDUTILFROM = missing, the program considers all codes in their entire available history until the value of MEDUTILTO. **Note 4:** An individual is not required to be enrolled during the entire MEDUTILFROM period. **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** -180

Parameter	Field Name	Description
Medical Utilization Evaluation Period End	MEDUTILTO	**Details:** used in combination with MEDUTILFROM (above). MEDUTILTO defines the end of the medical utilization evaluation period, expressed in terms of "days from Index Date". For example, if Index Date=01/08/2009 and MEDUTILTO for a given condition code is set to -1, the MP algorithm will evaluate medical utilization between the MEDUTILFROM date through 01/07/2009. **Note 1**: the index date is "day zero". Therefore, if zero is included in the MEDUTILFROM - MEDUTILTO interval, the index date is included in the evaluation period. **Note 2: special case**: when MEDUTILTO = missing the program evaluates medical utilization in their entire available history after the index date. **Note 3:** An individual is not required to be enrolled during the entire MEDUTILTO period. **Note 4:** MEDUTILTO may be negative. **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** -1
Care Settings to Define Medical Visits	CARESETTINGS	**Details:** the SCDM encounter types to identify and count medical visits. Valid values must be quoted and separated by a space: • **IP:** inpatient hospital stays • **IS:** non-acute institutional stays • **ED:** emergency department visits • **AV:** ambulatory visits • **OA:** other ambulatory visits **Note 1:** if the output of the CIDA tool will be used for subsequent analyses (ANALYSIS=PS or ADS), all care settings must be specified. In this case, CIDA tool output to the *msoc* folder will contain one overall utilization metric that counts all medical visits across all SCDM encounter types. The output dataset to the *dplocal* folder for the PSA tool, however, will contain the number of visits by each of the 5 encounter types. These metrics will be used to estimate the propensity score along with other covariates specified in the Covariate Codes File. **Note 2:** metrics are always computed allowing one visit per patient per day. If all care settings are specified by the CIDA tool, the overall utilization metric will only allow one visit per patient per day across all care settings specified in the

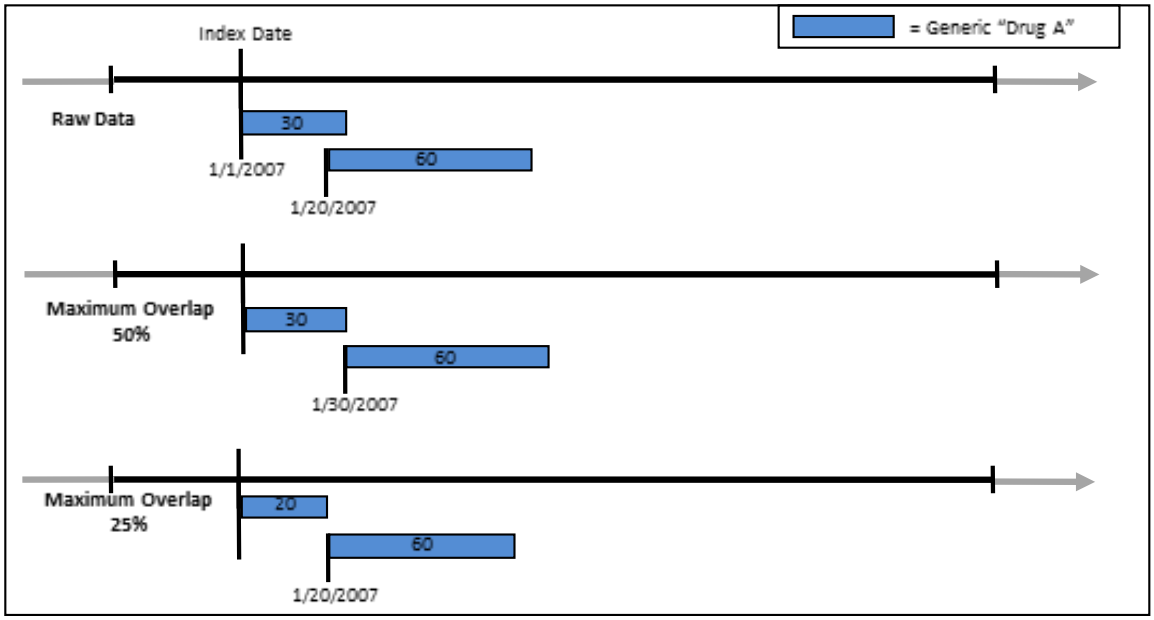
Parameter	Field Name	Description
		CARESETTINGS field (e.g., an outpatient and inpatient encounter on the same day is counted as one encounter). If all care settings are specified by the CIDA tool, the 5 metrics calculated for the PSA tool will count one visit per patient per care setting per day. **Note 3**: for inpatient visits (IP), length of stay is not taken into account. In other words, the module will count the number of inpatient admissions as reported in the data. Moreover, inpatient episodes that may overlap are not combined by the module. **Defined by:** Requester **Input type:** Required **Format:** Alphanumeric; SAS character $25. **Example:** 'ED' 'IP' 'IS' 'OA' 'AV'
Counts of Medical Visits Group Stratification	CSSTRAT	**Details:** groupings of counts of medical visits for CIDA tool output stratification. **Note 1:** groupings must be separated by a space; use "+" if the last group is open ended. **Note 2:** stratification is done by a single, overall utilization metric (for all visits observed for the encounter types provided in the CARESETTINGS field). **Defined by:** Requester **Input type:** Required **Format:** Alphanumeric; SAS character $20. **Example:** 0 1 2-7 8+
Drug Utilization Evaluation Period Start	DRUGUTILFROM	**Details:** used in combination with DRUGUTILTO (below). DRUGUTILFROM defines the start of the drug utilization evaluation period, expressed in terms of "days from Index Date". For example, if Index Date=01/08/2009 and DRUGUTILFROM is set to -7, the MP algorithm will start evaluating drug utilization on 01/01/2009. **Note 1**: the index date is "day zero". Therefore, if zero is included in the DRUGUTILFROM – DRUGUTILTO interval, the index date is included in the evaluation period. **Note 2**: if DRUGUTILFROM > 0 then the evaluation period will start after the index date. **Note 3: special case**: when DRUGUTILFROM = missing, the program considers all codes in their entire available history until the value of DRUGUTILTO. **Note 4:** An individual is not required to be enrolled during the entire DRUGUTILFROM period.

Parameter	Field Name	Description
		Named by: Requester **Input type:** Required **Format:** Numeric **Example:** -180
Drug Utilization Evaluation Period End	DRUGUTILTO	**Details:** used in combination with DRUGUTILFROM (above). DRUGUTILTO defines the end of the medical utilization evaluation period, expressed in terms of "days from Index Date". For example, if Index Date=01/08/2009 and DRUGUTILTO for a given condition code is set to -1, the MP algorithm will evaluate medical utilization between the DRUGUTILFROM date through 01/07/2009. **Note 1**: the index date is "day zero". Therefore if zero is included in the DRUGUTILFROM - DRUGUTILTO interval, the index date is included in the evaluation period. **Note 2: special case**: when DRUGUTILTO = missing the program evaluates medical utilization in their entire available history after the index date. **Note 3:** An individual is not required to be enrolled during the entire DRUGUTILTO period. **Note 4:** DRUGUTILTO may be negative. **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** -1

o) Stockpiling File

The Stockpiling File is optional. It is used to instruct the MP algorithm on how valid dispensings are selected and used by the stockpiling algorithm to create exposure episodes. Requesters can require restrictions on days supplied and amount supplied values for dispensings that are considered by the modular program. Requesters can also specify how the program adjusts dispensing dates based on the maximum overlap between adjacent dispensings. For example, consider the dispensing pattern in Figure 40 where the first dispensing and second dispensing overlap by 10 days.

Figure 40. Use of Maximum Percentage Overlap in Stockpiling File



If a requester specifies a maximum overlap of 50%, the stockpiling algorithm will only augment dispensing dates if the number of days of overlap between the two dispensing is less than (30 days * .5) = 15 days. Since the dispensings overlap by 10 days (<15 days) the start date of the second dispensing is adjusted to 1/30/2007 (30 days after the first dispensing date). However, if a requester specifies a maximum overlap of 25%, the stockpiling algorithm will only augment dispensing dates if the number of days of overlap between the two dispensing is less than (30 days * .25) = 7 days (value is rounded down). Since the dispensings overlap by 10 days (> 7 days) the start date of the second dispensing is not adjusted and the first dispensing days supply value is truncated at 20 days.

Default values of each parameter are described below in the detailed specifications for this file (Table 39). If default values are requested, this input file does not need to be specified in the program package.

Table 39. STOCKPILINGFILE Specification

Parameter	Field Name	Description
Name of Cohort	GROUP	**Details:** standardized name used to differentiate cohorts. **Note 1**: multiple cohorts can be defined within the same Stockpiling File. In this case all cohorts are queried independently and results are reported separately and labeled using each GROUP name specified. **Note 2:** GROUP is the primary key linking cohorts across input files; GROUP values must match (including case) between the STOCKPILINGFILE file and other input files. **Note 3**: GROUP values must remain consistent during the course of a surveillance activity. **Named by:** Request programmer **Input type:** Required **Format:** Alphanumeric; SAS character $40; no special characters (*e.g.,* commas, periods, hyphens, etc.) allowed, and underscores must be used to mark spaces. **Example:** Insulin
Same Day Dispensing Processing Indicator	SAMEDAY	**Details**: defines how same day dispensings are processed. The first position indicates how days supplied (RxSup in the SCDM) is handled; the second position indicates how amount supplied (RxAmt in the SCDM) is handled. Valid values (for each position are): • **a:** adds all (amount supplied or days supplied) values for dispensings in the same GROUP/STOCKGROUP on the same day • **n:** uses minimum (amount supplied or days supplied) value for dispensings in the same GROUP/ STOCKGROUP on the same day • **x:** uses maximum (amount supplied or days supplied) value for dispensings in the same GROUP/ STOCKGROUP on the same day • **m:** uses mean (amount supplied or days supplied) value for dispensings in the same GROUP/ STOCKGROUP on the same day **Note 1**: a total of 16 combinations are possible (*e.g.,* aa, an, etc.). **Note 2**: default value is “aa”. **Defined by:** Requester **Input type:** Required

Parameter	Field Name	Description
		Format: SAS character $2 **Example:** SAMEDAY = aa
Range of Allowable Days Supplied Values	SUPRANGE	**Details:** specifies the allowable range of days supplied values (variable RxSup in the SCDM) that are allowed for a dispensing to be used to create valid treatment episodes. Valid values are: • **x<-HIGH:** value must be > x • **y-HIGH**: value must be >= y • **LOW-<x**: value must be < x • **x-y**: value must be between x and y inclusively • **x<-y**: value must be greater than x and less or equal than y • **x-<y**: value must be greater or equal than x and less than y • **x<-<y**: value must be between x and y but not equal **Note 1**: allowable values can also be discrete, *e.g.,* "10", "20". **Note 2**: failing to be in the specified range excludes a dispensing from consideration. **Note 3:** default is "0<-HIGH", indicating that the program will not consider days supplied values of 0 or less. **Defined by**: Requester **Input type**: Required **Format**: SAS character $40 **Examples**: SUPRANGE=5-<80; SUPRANGE = 0<-HIGH
Range of Allowable Amount Supplied Values	AMTRANGE	**Details:** specifies the allowable range of amount supplied values (variable RxAmt in the SCDM) that are allowed for a dispensing to be used to create valid treatment episodes. Valid values are: • **x<-HIGH:** value must be > x • **y-HIGH**: value must be >= y • **LOW-<x**: value must be < x • **x-y**: value must be between x and y inclusively • **x<-y**: value must be greater than x and less or equal than y • **x-<y**: value must be greater or equal than x and less than y

Parameter	Field Name	Description
		• **x<-<y**: value must be between x and y but not equal **Note 1**: allowable values can also be discrete, *e.g.,* "10", "20". **Note 2**: failing to be in the specified range excludes a dispensing from consideration. **Note 3:** default is "0<-HIGH", indicating that the program will not consider amount supplied values of 0 or less. **Defined by**: Requester **Input type**: Required **Format**: SAS character $40 **Examples**: SUPRANGE=5-<80; SUPRANGE = 0<-HIGH
Overlap Percentage Processing	PERCENTDAYS	**Details:** the maximum percentage overlap of previous dispensing's days supply allowed for pushing dispensing dates forward. When this percentage is exceeded, the previous dispensing's days supply is truncated at the day prior to the next dispensing date. If this parameter is left blank, no truncation will occur and any overlap of supply between dispensing will be corrected by pushing overlapping days supplied forward. **Note 1**: default is blank. **Note 2**: Although rare, when PERCENTDAYS >0 it is possible for the overlap to exceed 100%. If this occurs, the dispensing will be replaced by the eclipsed claim. **Defined by:** Requester **Input type:** Optional **Format:** Numeric **Example:** PERCENTDAYS = 0.25

p) Concomitant Use File

The Concomitant Use file allows requesters to specify GROUP values from a Type 2 analysis and perform additional analyses. In this file, requesters can specify a primary treatment episode (defined by a GROUP in TYPE2FILE), , evaluate the occurrence of secondary episodes (defined by a GROUP in TYPE2FILE), and evaluate if an outcome of interest occurs during concomitant use via CONC_FUP = "DEF" in cohortcodes.

If a Type 2 concomitant use and Type 2 multiple event analyses are both requested, they must be specified in two separate program runs.

Table 40. CONCFILE Parameters

Parameter	Field Name	Description
Group name for analysis	ANALYSISGRP	**Details:** COHORTGRP name to differentiate primary/secondary pairs. **Named by:** Requester **Input type:** Required **Format:** Character (40) **Example:** drug_a_and_drug_b
Primary Episode	PRIMARY	**Details:** COHORTGRP name of the primary episode of interest **Named by:** Requester **Input type:** Required **Format:** Character (40) **Example:** drug_a
Secondary Episode	SECONDARY	**Details:** GROUP name of the secondary episode of interest **Note 1:** Requester can specify multiple GROUPs as one secondary episode. GROUP episodes will be collapsed into single secondary episodes **Named by:** Requester **Input type:** Required **Format:** Character (100) **Example:** drug_b
Exposure Order Indicator	EXPORDER	**Details:** indicates whether the order of primary and secondary exposure is relevant when creating valid concomitant treatment episodes (e.g., primary exposure must always be initiated before secondary exposure). **Allowable values:** **P:** always require primary exposure to be initiated before secondary exposure for concomitant episodes. **S:** restrict concomitant episodes to those where primary and secondary episodes are initiated on the same day. **N**: will not enforce order restriction. **Named by:** Requester

Parameter	Field Name	Description
		Input type: Required (default value is N) **Format:** Character $1 **Example:** N
Event Washout Period for concomitant episodes.	CONCFUPWASHPER	**Details:** length of event washout period in days. The washout period is a period before an incident concomitant treatment episode during which a member cannot have any evidence of event(s) of interest or any other event(s) specified in the CONC_FUP parameter. If a member has fewer than CONCFUPWASHPER days of enrollment before the concomitant episode index date, the treatment episode is excluded from the incident evaluation. **Note 1**: the event washout period looks back from the concomitant episode index date. **Named by:** Requester **Input type:** Required (0 must be entered if no EVENTWASHPER is required) **Format:** Numeric **Example:** 365
Event Blackout Period for concomitant period	CONCBLACKOUTPER	**Details:** the event blackout period in days. The requester can specify a period at the start of a concomitant treatment episode during which valid events found by the concomitant algorithm are ignored. That is, the at-risk period starts at the end of the blackout period. Moreover, if an event occurs during the blackout period, the episode will not be considered incident with respect to the event (and thus excluded from output metrics). **Named by:** Requester **Input type:** Required (0 must be entered if no BLACKOUTPER is required) **Format:** Numeric **Example:** 7
Number of concomitant episodes to count	CONCCOHORTDEF	**Details:** Defines whether to count the first eligible concomitant episode, or to count all concomitant episodes **Note 1:** This is different from selecting the T2COHORTDEF. T2COHORTDEF defines whether to retain the first or all eligible episodes in the entire query period. CONCCOHORTDEF determines whether to keep the first or all eligible **concomitant** episodes. **Note 2:** Valid evalues are: • 01: Count the first concomitant episode • 02: Count all concomitant episodes

Parameter	Field Name	Description
		Named by: Requester **Input type:** Required **Format:** Alphanumeric; SAS Chatacter $2. **Example:** 01
HOI Characterization De-duplication Process for concomitant episodes	CONCEVENTCOUNT	**Details:** by design, individuals stop contributing days at risk during an exposure episode when an HOI occurs. HOIs/Days at-risk metrics reported allow individuals to contribute, at most, one HOI per episode. However, the MP algorithm is able to characterize the number of total HOIs observed during valid treatment episodes. Requesters can use this field to determine how this characterization should count the number of HOIs. Again, this is for characterization only, and will not affect HOI/Days at-risk metrics. Valid values are: • **0:** counts all occurrences of an HOI during an exposure episode. • **1:** de-duplicates occurrences of the *same HOI code and code type* on the same day (*i.e.,* de-duplicates at the exact match code level). Note: a patient may have the same HOI code and code type on the same day if they were recorded by different providers and/or occurred in different care settings. • **2:** de-duplicates occurrences of the *same HOI GROUP* on the same day (*e.g.,* de-duplicates at the GROUP level). Consider the example where the HOI is defined with ICD-9-CM diagnosis codes 250.01 and 250.11 in any care setting. A member has an occurrence of code=250.01 on two separate AV records and of code=250.11 on another AV record on the same date during his/her incident treatment episode. CONCEVENTCOUNT=0 will identify three HOIs. CONCEVENTCOUNT=1 will identify two HOIs. CONCEVENTCOUNT=2 will identify one HOI. **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** 2

q) Multiple Events File

The MULTEVENTFILE allows requesters to specify COHORTGRP values from a Type 2 analysis and perform additional analyses. In this file, requesters can specify a primary treatment episode (defined by a GROUP in TYPE2FILE), define an observation window relative to the primary treatment episode, and evaluate the occurrence of secondary episodes/events (defined by a GROUP in TYPE2FILE). Events can be defined as an interval (i.e., an episode) or as a single point in time.

If a Type 2 concomitant/overlap use and Type 2 multiple event analysis are requested, they must be specified in two separate program runs.

Table 41. MULTEVENTFILE Parameters

Parameter	Field Name	Description
Group name for analysis	ANALYSISGRP	**Details:** GROUP name to differentiate primary/secondary pairs. **Named by:** Requester **Input type:** Required **Format:** Character (40) **Example:** drug_a_and_drug_b
Primary Episode	PRIMARY	**Details:** COHORTGRP name of the primary episode of interest **Named by:** Requester **Input type:** Required **Format:** Character (40) **Example:** drug_a
Secondary Episode	SECONDARY	**Details:** COHORTGRP name of the secondary episode of interest **Note 1:** Requester can specify multiple GROUPs as one secondary episode. GROUP episodes will be collapsed into single secondary episodes **Named by:** Requester **Input type:** Required **Format:** Character (100) **Example:** drug_b
Start of Observation Window	OBSFROM	**Details:** Number of days from OBSFROMANCHOR to start observation window. **Note 1:** If OBSFROM =., all enrollment history prior to OBSFROMANCHOR will be considered. **Note 2:** If OBSFROM > 0, then the evaluation period will start after the OBSFROMANCHOR date **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** -60

Parameter	Field Name	Description
End of Observation Window	OBSTO	**Details:** Number of days from OBSTOANCHOR to end observation window **Note 1** If OBSTO= ., all enrollment history after RISKANCHORTO will be considered. **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** 60
Day 0 for start of observation window	OBSFROMANCHOR	**Details:** Defines point in time to calculate start of observation window **Note 1:** Valid values are: Index, EpisodeEnd **Note 2:** To only consider primary episode duration as the observation window, OBSFROMANCHOR = Index and OBSFROM= 0. **Named by:** Requester **Input type:** Required **Format:** Alphanumeric **Example:** Index
Day 0 for end of observation window	OBSTOANCHOR	**Details:** Defines point in time to calculate end of observation window **Note 1:** Valid values are: Index, EpisodeEnd **Note 2:** To only consider primary episode duration as the observation window, OBSTOANCHOR = EpisodeEnd and OBSTO = 0. **Named by:** Requester **Input type:** Required **Format:** Alphanumeric **Example:** Index
Secondary Episode to use for time to secondary episode output	EPISODENUM	**Details:** Defines which secondary episode to output time to secondary episode statistics **Note 1:** Valid values = 1 (time to 1st episode), 2 (time to 2nd episode), 3 (time to 3rd episode) **Note 2:** When EPISODENUM = 2 or 3, then time to secondary episode statistics are only computed for episodes with ≥ 2 or ≥3 secondary episodes **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** 1

Parameter	Field Name	Description
Minimum cutoff to be considered adherent	MINADHERENCE	**Details:** Defines the minimum number or percent of adherent episodes a patient must have in order for the patient to be considered adherent. **Note 1:** This is specified in conjunction with the parameter MINADHERENCE_SCALE. **Note 2:** Leave blank if adherence is not being evaluated. **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** 50
Defines scale for MINADHERENCE parameter	MINADHERENCE_SCALE	Details: Defines the scale used (either count or percent) for determining whether a patient is considered adherent. Valid values are: • Percent • Count **Note 1:** This is specified in conjunction with the parameter MINADHERENCE **Note 2:** For example, if MINADHERENCE = 50 and MINADHERENCE_SCALE = percent, then 50% or more of a patient's primary episodes must meet adherence in order for the patient to be considered adherent **Note 3:** Leave blank if adherence is not being evaluated. **Named by:** Requester **Input type:** Required **Format:** Character (10) **Example:** percent
Primary episode categories	TTE_OUTPUT_CAT	Stratification variable. Categorizes length of the primary episode variable in user-defined ranges. **Note 1.** This may be left blank if not requested Format: Character (10)

r) Multiple Events Adherence Definition File

The optional MULTEVENTFILE_ADHERE file allows requesters to specify multiple criteria to determine overall adherence for a Type 2 multiple events analysis. In this analysis, adherence can be specified via user-defined parameters: primary episode duration, minimum number of secondary episodes/events, time to first secondary episode/event, and secondary episode/event gap (includes censoring as an event).

Each criterion is specified as a unique ADHERENCEID. Multiple adherence patterns for each pair will be considered "OR" criteria (i.e. episode meets adherence if ADHERENCEID 1 is met OR ADHERENCEID 2 or ADHERENCEID 3, etc.).

For multiple events analysis, any number of combinations of EPISODELENGTH_START/END, EPIGAP_START/TO, EPICOUNT_START/END, and TTEPI_START/END can be used to define adherence.

If a Type 2 concomitant use and Type 2 multiple event analysis are both requested, they must be specified in two separate program runs.

Table 42. MULTEVENTFILE_ADHERE Specification

Parameter	Field Name	Description
Group name for analysis	ANALYSISGRP	**Details:** GROUP name to differentiate primary/secondary pairs. **Named by:** Requester **Input type:** Required **Format:** Character (40) **Example:** drug_a_and_drug_b
Adherence Identifier	ADHERENCEID	**Details:** Numeric identifier to identify adherence pattern **Note 1:** ADHERENCEID should start with 1 for each ANALYSISGRP **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** 1
Length of Primary Episode Start	EPISODELENGTH_START	**Details:** Minimum episode length **Note 1:** Missing EPISODELENGTH_START will consider all episode lengths ≤ EPISODELENGTH_END **Named by:** Requester **Input type:** Optional **Format:** Numeric **Example:** 1
Length of Primary Episode End	EPISODELENGTH_END	**Details:** Maximum episode length **Note 1:** Missing EPISODELENGTH_END will consider all episode lengths ≥ EPISODELENGTH_START **Named by:** Requester **Input type:** Required **Format:** Numeric

Parameter	Field Name	Description
		Example: 365
Number of Secondary Episodes Start	EPICOUNT_START	**Details:** Minimum number of secondary episodes **Note 1:** Missing EPICOUNT_START will consider the number of secondary episodes ≤ EPICOUNT_END **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** 1
Number of Secondary Episodes End	EPICOUNT_END	**Details:** Maximum number of secondary episodes **Note 1:** Missing EPICOUNT_END will consider the number of secondary episodes ≥ EPICOUNT_START **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** 5
Secondary Episode Gap Start	EPIGAP_START	**Details:** Minimum gap between secondary episodes **Note 1:** Missing EPIGAP_START will consider all gaps ≤ EPIGAP_END **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** 1
Secondary Episode Gap End	EPIGAP_END	**Details:** Maximum gap between secondary episodes **Note 1:** Missing EPIGAP_END will consider all gaps ≥ EPIGAP_START **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** 30
Time to Secondary Episode Start	TTEPI_START	**Details:** Minimum time to EPISODENUM secondary episode **Note 1:** Missing TTEPI_START will consider any time to secondary episode ≤ TTEPI_END **Note 2:** EPISODENUM parameter is specified in the MULTEVENTFILE input file. When EPISODENUM = 2, then time to secondary episode will be calculated from RISKANCHORFROM to time to 2nd secondary episode. In this situation, time to secondary episode will not be calculated for primary episodes with 0 or 1 secondary episodes – i.e. the first secondary episode will be

Parameter	Field Name	Description
		ignored for TTEPI_START/END and EPIGAP_START/END parameters. **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** 1
Time to Secondary Episode Start	TTEPI_END	**Details:** Maximum time to EPISODENUM secondary episode **Note 1:** Missing TTEPI_END will consider time to EPISODENUM secondary episodes ≥ TTEPI_START **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** 10

s) Overlap File

The OVERLAPFILE characterizes an overlap of primary and secondary treatment episodes during the observation window. It allows requesters to specify COHORTGRP values from a Type 2 analysis to perform additional analyses. In this file, requesters can specify a primary treatment episode (defined by a GROUP in TYPE2FILE), define an observation window relative to the primary treatment episode,

If a Type 2 concomitant use and Type 2 multiple event analyses are both requested, they must be specified in two separate program runs.

Table 43. OVERLAPFILE Specifications

Parameter	Field Name	Description
Group name for analysis	ANALYSISGRP	**Details:** COHORTGRP name to differentiate primary/secondary pairs. **Named by:** Requester **Input type:** Required **Format:** Character (40) **Example:** drug_a_and_drug_b
Primary Episode	PRIMARY	**Details:** COHORTGRP name of the primary episode of interest **Named by:** Requester **Input type:** Required **Format:** Character (30) **Example:** drug_a
Secondary Episode	SECONDARY	**Details:** GROUP name of the secondary episode of interest **Note 1:** Requester can specify multiple GROUPs as one secondary episode. GROUP episodes will be collapsed into single secondary episodes

Parameter	Field Name	Description
		Named by: Requester **Input type:** Required **Format:** Character (100) **Example:** drug_b
Start of Observation Window	OBSFROM	**Details:** Number of days from OBSFROMANCHOR to start observation window. **Note 1:** If OBSFROM=., all enrollment history prior to OBSFROMANCHOR will be considered. **Note 2:** If OBSFROM> 0, then the evaluation period will start after the OBSFROMANCHOR date **Note 3:** overlap will calculate the number of days the secondary episode overlaps the observation window. If overlap of primary episode is desired, observation window should be primary treatment episode. **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** -60
End of Observation Window	OBSTO	**Details:** Number of days from OBSTOANCHOR to end observation window **Note 1** If OBSTO = ., all enrollment history after RISKANCHORTO will be considered. **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** 60
Day 0 for start of observation window	OBSFROMANCHOR	**Details:** Defines point in time to calculate start of observation window **Note 1:** Valid values are: Index, EpisodeEnd **Note 2:** To only consider primary episode duration as the observation window, OBSFROMANCHOR = Index and OBSFROM= 0. **Named by:** Requester **Input type:** Required **Format:** Alphanumeric **Example:** Index
Day 0 for end of observation window	OBSTOANCHOR	**Details:** Defines point in time to calculate end of observation window **Note 1:** Valid values are: Index, EpisodeEnd

Parameter	Field Name	Description
		Note 2: To only consider primary episode duration as the observation window, OBSTOANCHOR = EpisodeEnd and OBSTO = 0. **Named by:** Requester **Input type:** Required **Format:** Alphanumeric **Example:** Index
Defines categories for overlap output	CUTOFFCAT	**Details:** Defines the categories to bin secondary and observation window overlap. Multiple categories should be separated by a space **Note 1:** Each category will be evaluated separately. Categories do not have to be mutually exclusive **Note 2:** Valid special characters are <, ≤, >, ≥, - **Note 3:** This is specified in conjunction with CUTOFFCAT_SCALE. **Named by:** Requester **Input type:** Required **Format:** Character (50) **Example:** 0-20 21-50 ≥50 <30 <60
Defines scale for CUTOFFCAT parameter	CUTOFFCAT_SCALE	**Details:** Defines the scale for CUTOFFCAT categories. Valid values are: • Percent • Day **Note 1:** This is specified in conjunction with the parameter CUTOFFCAT **Named by:** Requester **Input type:** Required **Format:** Character (10) **Example:** percent
Minimum cutoff to be considered adherent	MINADHERENCE	**Details:** Defines the minimum number or percent of adherent episodes a patient must have in order for the patient to be considered adherent. **Note 1:** This is specified in conjunction with the parameter MINADHERENCE_SCALE. **Note 2:** Leave blank if adherence is not being evaluated. **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** 50

Parameter	Field Name	Description
Defines scale for MINADHERENCE parameter	MINADHERENCE_SCALE	Details: Defines the scale used (either count or percent) for determining whether a patient is considered adherent. Valid values are: • Percent • Count **Note 1:** This is specified in conjunction with the parameter MINADHERENCE **Note 2:** For example, if MINADHERENCE = 50 and MINADHERENCE_SCALE = percent, then 50% or more of a patient's primary episodes must meet adherence in order for the patient to be considered adherent **Note 3:** Leave blank if adherence is not being evaluated. **Named by:** Requester **Input type:** Required **Format:** Character (10) **Example:** percent
Primary episode categories	TTE_OUTPUT_CAT	Stratification variable. Categorizes length of the primary episode variable in user-defined ranges. **Note 1:** May be left blank if not requested. Format: Character (10)

t) Overlap Adherence Definition File

The optional OVERLAPFILE_ADHERE file allows requesters to specify multiple criteria to determine overall adherence for a concomitant use analysis. Adherence may be based on minimum or maximum % or number of days overlap between a primary episode observation window and a seconday episode. If the observation window falls outside of the primary episode, then % overlap is of the observation window and secondary episode.

Each criterion is specified as a unique ADHERENCEID. Multiple adherence patterns for each pair will be considered "OR" criteria for evaluation of adherence (i.e. episode is counted in the ADHERENCE field in the output if ADHERENCEID 1 is met OR ADHERENCEID 2 or ADHERENCEID 3, etc.). Any number of combinations of OVERLAP_START/END can be used to define adherence.

If a Type 2 overlap use and Type 2 multiple event analysis are both requested, they must be specified in two separate program runs.

Table 44. OVERLAPFILE_ADHERE Specifications

Parameter	Field Name	Description
Group name for analysis	ANALYSISGRP	**Details:** GROUP name to differentiate primary/secondary pairs. **Named by:** Requester **Input type:** Required **Format:** Character (40) **Example:** drug_a_and_drug_b
Adherence Identifier	ADHERENCEID	**Details:** Numeric identifier to identify adherence pattern **Note 1:** ADHERENCEID should start with 1 for each ANALYSISGRP **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** 1
Minimum overlap value	OVERLAP_START	**Details:** Minimum overlap between secondary episodes and observation window to be considered adherent **Note 1:** Missing OVERLAP_START will consider any overlap ≤ OVERLAP_END **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** 50
Maximum overlap value	OVERLAP_END	**Details:** Maximum overlap between secondary episodes and observation window to be considered adherent **Note 1:** Missing OVERLAP_END will consider any overlap ≥ OVERLAP_START **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** 100
Overlap scale	OVERLAP_SCALE	**Details:** Scale to define OVERLAP_START AND OVERLAP_END **Note 1:** Valid values are: • Percent • Days **Named by:** Requester **Input type:** Required **Format:** character (10) **Example:** Percent

u) Pregnancy Duration File

The Pregnancy Duration File is optional and its specification is only required for a pregnancy episodes identification strategy (Type 4 analysis). It is used to instruct the MP algorithm on how to calculate pregnancy episode duration and index date. Requesters can specify which codes are used to calculate pregnancy duration and the priority given to each code when multiple codes are identified. Table 45 contains detailed specifications for this file.

Table 45. PREGDUR Specification

Parameter	Field Name	Description
Name of Cohort	GROUP	**Details:** standardized name used to differentiate cohorts. **Note 1**: multiple cohorts can be defined within the same pregdur file. In this case all cohorts are queried independently and results are reported separately and labeled using each GROUP name specified. **Note 2:** GROUP is the primary key linking cohorts across input files; GROUP values must match (including case) between the PREGDUR file and other input files. **Note 3**: GROUP values must remain consistent during the course of a surveillance activity. **Named by:** Request programmer **Input type:** Required **Format:** Alphanumeric; SAS character $30; no special characters (*e.g.,* commas, periods, hyphens, etc.) allowed, and underscores must be used to mark spaces. **Example:** Insulin
Name of Stockgroup within the Cohort	STOCKGROUP	**Details:** standardized name used to refer to a specific preterm/postterm codes within a given GROUP. **Named by:** Requester **Input type:** Required **Format:** Alphanumeric; SAS character $30; special characters (*e.g.,* commas, periods, hyphens, etc.) allowed and underscores must be used to mark spaces. **Example:** PreTerm_2weeks
Code Category	CODECAT	**Details:** type of each code category value included in the CODETYPE field (below) of this file. Valid values include: • **RX:** NDC • **DX:** Diagnosis code • **PX:** Procedure code • **LB:** Lab code **Defined by:** Requester **Input type:** Required

Parameter	Field Name	Description
		Format: SAS character $2. **Example:** DX

Code Type	CODETYPE	**Details:** type of each code value included in the CODE field (below) of this file. Valid values include:
		
		If CODECAT = RX:
		• **09**: 9 digits NDC
		• **11**: 11 digits NDC
		
		If CODECAT = DX:
		• **09**: ICD-9-CM
		• **10**: ICD-10-CM
		• **11:** ICD-11-CM
		• **OT:** Other
		
		If CODECAT = PX:
		• **09**: ICD-9-CM
		• **10**: ICD-10-CM
		• **11**: ICD-11-CM
		• **C4**: CPT-4 (*i.e.,* HCPCS Level I)
		• **HC**: HCPCS (*i.e.,* HCPCS Level II)
		• **H3**: HCPCS Level III
		• **C2**: CPT Category II
		• **C3**: CPT Category III
		• **ND**: 11-digit NDC
		• **RE**: Revenue
		• **LO**: Local homegrown
		• **OT**: Other
		
		If CODECAT = LB:
		• **01N**: extract quantitative lab test result using SOC-defined lab code (see Lab Code Lookup Table)
		• **02N**: extract quantitative lab test result using LOINC
		• **'px'N**: extract quantitative lab test result using the following codes
		  • **09N**: ICD-9-CM
		  • **10N**: ICD-10-CM
		  • **11N**: ICD-11-CM
		  • **C4N**: CPT-4 (*i.e.,* HCPCS Level I)
		  • **HCN**: HCPCS (*i.e.,* HCPCS Level II)
		  • **H3N:** HCPCS Level III
		  • **C2N:** CPT Category II
		  • **C3N:** CPT Category III
		  • **NDN:** 11-digit NDC
		  • **REN:** Revenue

Parameter	Field Name	Description

		• **LON:** Local homegrown
		• **01C:** extract qualitative lab test result using SOC-defined lab code (see Lab Code Lookup Table)
		• **02C:** extract qualitative lab test result using LOINC
		• **'px'C:** extract qualitative lab test result using the following codes
		  • **09C**: ICD-9-CM
		  • **10C**: ICD-10-CM
		  • **11C**: ICD-11-CM
		  • **C4C**: CPT-4 (*i.e.,* HCPCS Level I)
		  • **HCC**: HCPCS (*i.e.,* HCPCS Level II)
		  • **H3C:** HCPCS Level III
		  • **C2C:** CPT Category II
		  • **C3C:** CPT Category III
		  • **NDC:** 11-digit NDC
		  • **REC:** Revenue
		  • **LOC:** Local homegrown
		
		Note 1: as the LOINC field is not populated by all Data Partners in the SCDM Laboratory Result table and the CPT code may not be specific to a particular lab test, it is strongly recommended that the Laboratory Result table be queried using SOC-defined lab codes.
		
		Defined by: Requester
		Input type: Required
		Format: Alphanumeric; SAS character $3.
		Example: 09

Code	CODE	**Details:** NDC, procedure, diagnosis, and/or lab code of interest. **Note 1**: Codes are matched using exact values (*i.e.,* 3-digit code lookup requires an exact 3-digit code match). Wildcard match (*) functionality is also available for ICD-9 diagnosis codes (*e.g.,* querying "250*0" would be used to find any ICD-9-CM diagnosis codes for diabetes type II, or "250**" to find ICD-9-CM diagnosis codes for all diabetes codes in the range "250.00 - 250.99"). To get "starts with" codes, the user will have to specify 250, 250*, 250**. **Note 2**: For NDCs, either 9 or 11 digit codes can be entered. **Note 3:** remove decimal points in the code value.

Parameter	Field Name	Description
		Note 4: CODETYPE/CODECAT must be consistent with the expected format of the CODE value (*e.g.,* the program will not find any valid matches in the data for CODECAT=RX, CODETYPE=11 and a 9-digit NDC value). **Note 5**: Duplicate CODECAT-CODETYPE-CODE-CARESETTING-PRINCIPAL combinations are removed by the MP algorithm. **Note 6**: 'V' and 'E' ICD-9-CM diagnosis codes must be specified using uppercase 'V' and 'E'. **Defined by:** Requester, with support from the SOC as needed **Input type:** Required **Format:** Alphanumeric; SAS character $11. **Example: (CODECAT=DX; CODETYPE=09):** 64421
Care Setting and Diagnosis Position Requirements	CARESETTINGPRINCIPAL	**Details:** defines the care setting and principal diagnosis position requirements for each code. This field uses combination(s) of the SCDM variables care setting (ENCTYPE) and principal discharge diagnosis flag (PDX) to restrict the observance of codes to those in the requested care settings and with the requested diagnosis position. If no restrictions are required (*e.g.,* requester wants all care settings and any value of PDX), leave the field blank. The following are valid entries; all entries must be in single quotes and separated by a space: • **IPP:** inpatient hospital stays, principal diagnoses • **IPS:** inpatient hospital stays, secondary diagnoses • **IPX:** inpatient hospital stays, unclassified diagnoses • **ISP:** non-acute institutional stays, principal diagnoses • **ISS:** non-acute institutional stays, secondary diagnoses • **ISX:** non-acute institutional stays, unclassified diagnoses • **ED*:** emergency department encounters • **AV*:** ambulatory visits • **OA*:** other ambulatory visits **Request Programmer Note 1:** the wildcard symbol (*) can be used to represent "any" values of either care setting or principal discharge diagnosis flag. For example, CARESETTINGPRINCIPAL = 'IP*' will restrict codes to those observed in the inpatient setting

Parameter	Field Name	Description
		irrespective of the principal diagnosis flag value. CARESETTINGPRINCIPAL = '**P' will restrict diagnosis codes to those in the principal position, irrespective of the care setting. **Request Programmer Note 2:** the principal discharge diagnosis flag is only relevant for diagnosis codes. All other codes should use the * wildcard for the third digit of the CARESETTINGPRINCIPAL value. **Note 3:** CARESETTINGPRINCIPAL is allowed to vary between CODEs within the same GROUP. For example, CARESETTINGPRINCIPAL is allowed to equal 'IPP' for one diagnosis code and 'IPP' 'EDP' for another diagnosis code *in the same GROUP*. **Defined by:** Requester **Input type:** Optional; Default: blank (*i.e.,* no restrictions) **Format:** Alphanumeric **Example:** 'IPX' 'ED*' '**P'
Indicator for Priority Group 1 Delivery Codes	PRIORITYGROUP1	**Details:** indicates if a code is a first priority code **Note 1:** valid values are 0 or 1 **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** 1
Indicator for Priority Group 2 Delivery Codes	PRIORITYGROUP2	**Details:** indicates if a code is a second priority code **Note 1:** valid values are 0 or 1 **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** 0
Priority Assigned to Each Code	PRIORITY	**Details:** for members with many different codes, the priority is used to indicate which code will be kept to define the pregnancy episode start date and duration. **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** 3
Duration Assigned to Each Code	DURATION	**Details:** Duration in days attached to the pregnancy episode when a delivery has a gestational age code. **Note 1:** Codes with the same priority should have the same duration. **Defined by:** Requester

Parameter	Field Name	Description
		Input type: Required **Format:** Numeric **Example:** 268
Lab Date Selection Algorithm	RAWLABDATETYPE	**Details:** relevant for requests that query laboratory result values. Field specifies in what sequence date(s) in the SCDM Laboratory Result table should be considered to select one relevant date for a laboratory result of interest. The parameter will allow the requester to either specify 1) a single date variable (Lab_dt, Order_dt, or Result_dt) to use; or 2) a hierarchy to choose a date variable (*e.g.,* select Lab_dt else if missing select Result_dt else if missing select Order_dt). Valid values are any combination of the following: • **L:** Lab Date • **O:** Order Date • **R:** Result Date **Note 1:** leave blank if CODECAT ≠ LB. **Note 2:** many Data Partners do not populate all three date fields. Use of an algorithm for date selection is recommended. **Defined by:** Requester **Input type:** Required **Format:** SAS character $3. **Example 1:** RawLabDateType =LRO. In this case, the program will use Lab_dt else if missing use Result_dt else if missing use Order_dt. **Example 2:** RawLabDateType =L. In this case, the program will use Lab_dt only.
Lab Result Values	RAWLABRESULT	**Details:** specifies the lab result value or lab result range for querying. RAWLABRESULT allows for values or ranges of quantitative laboratory results (*e.g.,* 100; 100-200) and values of qualitative laboratory results (*e.g.,* "POSITIVE"). Valid values are: • <=X (less than or equal to X) • <X (less than X) • >=X (greater than or equal to X) • >X (greater than X) • ~=X (not equal to X) • X:Y (between X and Y)

Parameter	Field Name	Description
		Any string of relevant characters is allowed for qualitative results querying. **Note 1:** leave blank if CODECAT ≠ LB. **Note 2:** There are two fields in the Laboratory Result table that include results: MS_Result_C (contains results for qualitative tests) and MS_Result_N (contains results for quantitative tests). The field where the result will be queried will depend on the RAWCODETYPE value. **Note 3:** Ranges cannot be specified with hyphens. Must use ":". **Defined by:** Requester Input type: Required for laboratory results **Format:** Alphanumeric; **Example 1:** RawLabResult=20:50 **Example 2:** RawLabResult=POSITIVE

v) Mother-Infant Cohort File

The Mother-Infant Cohort File is optional and its specification is only required for a pregnancy episodes identification strategy (Type 4 analysis) that will create cohorts for further processing with the PSA tool. It is used to evaluate pregnancy and birth outcomes using Type 4 and perform PS-based inferential analyses. Table 456 contains detailed specifications for this file.

Table 46. MILCOHORTFILE Specification

Parameter	Field Name	Description
Analysis Group	ANALYSISGRP	**Details:** Requester defined name of group to include in output tables **Note 1:** Exposed group will be created with the Group name &ANALYSISGRP_exp **Note 2:** Comparator group will be created with the group name &ANALYSISGRP_ctrl **Defined by:** Requester **Input type:** Required (cannot be left blank) **Format:** Character(40) **Example:** preg_exposed
Exposure Group Name	GROUPNAME	**Details:** standardized name used to determine which CIDA cohort group to use to evaluate exposures and comparators **Named by:** Request programmer **Input type:** Required

Parameter	Field Name	Description
		Format: Alphanumeric; SAS character $40; no special characters (*e.g.,* commas, periods, hyphens, etc.) allowed, and underscores must be used to mark spaces. **Example:** exposed_drug1
Medical code for Exposure group	EXPMP	**Details:** Name of the MPn that links to the exposure group MPn in the T4_DEF variable in cohortcodes file. **Defined by:** Requester **Input type:** Required (cannot be left blank) **Format:** Character(5) **Example:** MP1
Medical code for comparator group	CONTROLMP	**Details:** Name of the MPn that links to the comparator group MPn in the T4_DEF variable in cohortcodes file. **Note:** If your comparator group is unexposed to the MP designated in COMPMP, please leave this blank. **Defined by:** Requester **Input type:**) **Format:** Character(5) **Example:** MP2
Sex criteria to apply to linked infant cohort	CSEX	**Details:** optional parameter to restrict infant cohort to only specified Sex values. Blank will ensure that all Sex values are included in analyses. **Note 1:** valid values will be in single quotes and separated by a space. Valid values are: • **F**: Female • **M**: Male • **O**: Other **Note 2:** Restriction will be applied to the Pregnant Exposed, Pregnant Unexposed, and Pregnant Comparator cohorts, and will not be applied to the Non-Pregnant Cohort. **Named by:** Requester **Input type:** Required **Format:** Char (3) **Example:** 'F' 'M'
Unit for defining exposure time	EXPOSUREUNIT	**Details:** Type of time interval that defines the exposure time period. Valid values include: • **T**: indicates trimester • **W**: indicates gestational weeks **Defined by:** Requester **Input type:** Required (cannot be left blank) **Format:** Character(1) **Example:** T

Parameter	Field Name	Description
Exposure Period Start	EXPOSUREFROM	**Details:** defines the start of the evaluation period for exposure of interest and will be used in combination with INTERVAL. **Note 1:** If INTERVAL is "T", valid values are -1 to 3 where negative value represents the pre-pregnancy period. **Note 2:** If INTERVAL is "W", valid values are integers, including negative values if period is prior to pregnancy start date **Defined by:** Requester **Input type:** Required (cannot be left blank) **Format:** Numeric **Example:** 1
Exposure Period End	EXPOSURETO	**Details:** defines the end of the evaluation period for exposure of interest and will be used in combination with INTERVAL. **Note 1:** If INTERVAL is "T", valid values are -1 to 3 where negative value represents the pre-pregnancy period. **Note 2:** If INTERVAL is "W", valid values are integers, including negative values if period is prior to pregnancy start date **Defined by:** Requester **Input type:** Required (cannot be left blank) **Format:** Numeric **Example:** 3
Identifying claims to look for outcomes	OUTCOMEPOP	**Details:** Identifying outcomes in claims, in conjunction with T4_FUP = "DEF" in cohortcodes. Valid values include: • **M**: indicates should look in mother claims only • **I**: indicates should look in infant claims only • **MI**: indicates should look in mother and infant claims **Defined by:** Requester **Input type:** Required (cannot be left blank) **Format:** character(2) **Example:** MI
Index date for covariates and PS risk set.	INDEXDATE	**Details:** Identifies which index date to use for covariate evaluation period and propensity score risk-set creation. Also identifies which index date to use to anchor OUTCOMEFROM. Valid values include:

Parameter	Field Name	Description
		• **INDEXDT-** Index date of pregnancy start • **INDEXDT_EXP** - First date of dispensing or procedure code of drug or comparator of interest/exposure. • **INDEXDT_DELIV** – Index date of delivery admission date (or infant birth date for a matched infant cohort) **Note 1:** If control cohort is unexposed, the index date can either be INDEXDT or INDEXDT_DELIV **Note 2:** OUTCOMEFROM will be anchored on INDEXDATE **Note 3:** If exposure occurs before and overlaps the start of the evaluation period (i.e., trimester or gestational week) INDEXDT_EXP will be set to the first day of the evaluation period). **Defined by:** Requester **Input type:** Required (cannot be left blank) **Format:** character(15) **Example:** INDEXDT_EXP
Anchor date to end outcome follow-up	OutcomeToAnchor	**Details:** Identifies which index date to use to anchor OUTCOMETO • **INDEXDT-** Index date of pregnancy start • **INDEXDT_EXP** - First date of dispensing or procedure code of drug or comparator of interest/exposure. • **INDEXDT_DELIV** – Index date of delivery admission date **Note 1:** If control cohort is unexposed, the index date can either be INDEXDT or INDEXDT_DELIV **Defined by:** Requester **Input type:** Required (cannot be left blank) **Format:** character(15) **Example:** INDEXDT_EXP
Outcome Period Start	OUTCOMEFROM	**Details:** used in combination with OUTCOMETO (below). OUTCOMEFROM defines the start of the evaluation period for outcome evaluation, expressed in terms of "days from INDEXDATE" (index date used for covariate assessment). For example, if INDEXDATE = INDEXDT_DELIV and delivery Date=01/08/2009 and OUTCOMEFROM is set to 0, the algorithm will start looking for that outcome code starting on 01/08/2009. **Defined by:** Requester

Parameter	Field Name	Description
		Input type: Required (cannot be left blank) **Format:** Numeric **Example:** -90
Outcome Period End	OUTCOMETO	**Details:** used in combination with OUTCOMEFROM (above). OUTCOMETO defines the end of the evaluation period for outcome evaluation, expressed in terms of "days from OUTCOMETOANCHOR". For example, if OUTCOMETOANCHOR = INDEXDT_DELIV and delivery Date=01/08/2009 and OUTCOMETO for a given condition code is set to 10, the MP algorithm will look for that outcome code between the OUTCOMEFROM date through 01/18/2009. **Defined by:** Requester **Input type:** Required (cannot be left blank) **Format:** Numeric **Example:** 10

w) Most Frequent Utilization File

Optional file to request most frequent utilization assessment. Specifications in this file apply to all GROUPs in a single execution of the program. A single row in this table represents a unique combination of ANALYSISNUM, CODECAT, and CODETYPE. This file is available to specify for all Types of analysis and reference a previously defined index date. Table 47 contains detailed specifications for this file.
If lab codes are requested, lab date will be assigned as the first non missing value in the order of lab date, result date, order date.

Table 47. MFUFILE Specifications

Parameter	Field Name	Description
Analysis Number	ANALYSISNUM	**Details:** numeric indicator to define a unique analysis. A unique analysis can have multiple rows in this file if multiple combinations of CODECAT/CODETYPE are requested. **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** 1
Number of Codes to Return	TOPXX	**Details:** number of codes to return in output, across all specified code types. For example, if TOPXX = 50, the program will return the most frequent 50 codes observed during the MFUFROM - MFUTO period across all of code types specified in CODECAT/CODETYPE **Named by:** Requester **Input type:** Required **Format:** Numeric

Parameter	Field Name	Description
		Example: 50
Evaluation Period Start	MFUFROM	**Details:** used in combination with MFUTO (below). MFUFROM defines the start of the evaluation period to identify most frequent codes, expressed in terms of "days from Index Date". For example, if Index Date=01/08/2009 and MFUFROM is set to -7, the algorithm will start looking for codes on 01/01/2009. **Note 1:** members must be enrolled for the duration of the MFUFROM – Index Date period. **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** -7
Evaluation Period End	MFUTO	**Details:** used in combination with MFUFROM (above). MFUTO defines the end of the evaluation period to identify most frequent codes, expressed in terms of "days from Index Date". For example, if Index Date=01/08/2009 and MFUTO is set to 20, the algorithm will look for codes until 01/28/2009. **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** 20
Code Category	CODECAT	**Details:** type of each code category value included in the CODETYPE field (below) of this file. Valid values include: • **RX:** NDC • **DX:** Diagnosis code • **PX:** Procedure code • **LB:** Lab code **Defined by:** Requester **Input type:** Required **Format:** SAS character $2. **Example:** DX
Code Type	CODETYPE	**Details:** type of each code value included in the CODE field (below) of this file. Valid values include: If CODECAT = RX: • **09**: 9 digits NDC • **11**: 11 digits NDC If CODECAT = DX: • **09**: ICD-9-CM • **10**: ICD-10-CM

Parameter	Field Name	Description

		• **11:** ICD-11-CM
		• **OT:** Other
		
		If CODECAT = PX:
		• **09**: ICD-9-CM
		• **10**: ICD-10-CM
		• **11**: ICD-11-CM
		• **C4**: CPT-4 (*i.e.,* HCPCS Level I)
		• **HC**: HCPCS (*i.e.,* HCPCS Level II)
		• **H3**: HCPCS Level III
		• **C2**: CPT Category II
		• **C3**: CPT Category III
		• **ND**: 11-digit NDC
		• **RE**: Revenue
		• **LO**: Local homegrown
		• **OT**: Other
		
		If CODECAT = LB:
		• **01N**: extract quantitative lab test result using SOC-defined lab code (see Lab Code Lookup Table)
		• **02N**: extract quantitative lab test result using LOINC
		• **'px'N**: extract quantitative lab test result using the following codes
		  • **09N**: ICD-9-CM
		  • **10N**: ICD-10-CM
		  • **11N**: ICD-11-CM
		  • **C4N**: CPT-4 (*i.e.,* HCPCS Level I)
		  • **HCN**: HCPCS (*i.e.,* HCPCS Level II)
		  • **H3N:** HCPCS Level III
		  • **C2N:** CPT Category II
		  • **C3N:** CPT Category III
		  • **NDN:** 11-digit NDC
		  • **REN:** Revenue
		  • **LON:** Local homegrown
		• **01C:** extract qualitative lab test result using SOC-defined lab code (see Lab Code Lookup Table)
		• **02C:** extract qualitative lab test result using LOINC
		• **'px'C:** extract qualitative lab test result using the following codes
		  • **09C**: ICD-9-CM
		  • **10C**: ICD-10-CM

Parameter	Field Name	Description
		• **11C**: ICD-11-CM • **C4C**: CPT-4 (*i.e.*, HCPCS Level I) • **HCC**: HCPCS (*i.e.*, HCPCS Level II) • **H3C:** HCPCS Level III • **C2C:** CPT Category II • **C3C:** CPT Category III • **NDC:** 11-digit NDC • **REC:** Revenue • **LOC:** Local homegrown **Note 1:** as the LOINC field is not populated by all Data Partners in the SCDM Laboratory Result table and the CPT code may not be specific to a particular lab test, it is strongly recommended that the Laboratory Result table be queried using SOC-defined lab codes. **Defined by:** Requester **Input type:** Required **Format:** Alphanumeric; SAS character $3. **Example:** 09
Care Setting and Diagnosis Position Requirements	CARESETTINGPRINCIPAL	**Details:** defines the care setting and principal diagnosis position requirements for each CODETYPE requested. This field uses combination(s) of the SCDM variables care setting (ENCTYPE) and principal discharge diagnosis flag (PDX) to restrict the observance of codes to those in the requested care settings and with the requested diagnosis position. If no restrictions are required (*e.g.,* requester wants all care settings and any value of PDX), leave the field blank. The following are valid entries; all entries must be in single quotes and separated by a space: • **IPP:** inpatient hospital stays, principal diagnoses • **IPS:** inpatient hospital stays, secondary diagnoses • **IPX:** inpatient hospital stays, unclassified diagnoses • **ISP:** non-acute institutional stays, principal diagnoses • **ISS:** non-acute institutional stays, secondary diagnoses • **ISX:** non-acute institutional stays, unclassified diagnoses • **ED*:** emergency department encounters • **AV*:** ambulatory visits • **OA*:** other ambulatory visits

Parameter	Field Name	Description
		Request Programmer Note 1: the wildcard symbol (*) can be used to represent "any" values of either care setting or principal discharge diagnosis flag. For example, CARESETTINGPRINCIPAL = 'IP*' will restrict codes to those observed in the inpatient setting irrespective of the principal diagnosis flag value. CARESETTINGPRINCIPAL = '**P' will restrict diagnosis codes to those in the principal position, irrespective of the care setting. **Request Programmer Note 2:** care setting and principal discharge diagnosis flag is not relevant for RX code types. For these code types leave this field blank. **Request Programmer Note 3:** the principal discharge diagnosis flag is only relevant for diagnosis codes. Procedure code types should use the * wildcard for the third digit of the CARESETTINGPRINCIPAL value. **Defined by:** Requester **Input type:** Optional; Default: blank (*i.e.,* no restrictions) **Format:** Alphanumeric
Counting Method	COUNTMETHOD	**Details:** Indicates if the top most frequent codes are output or of the largest number of patients are output. Allowable values: **C**: the most frequent utilization output file produced is sorted by code counts **P**: the most frequent utilization output file produced is sorted by patient counts **Defined by:** Requester **Input type:** Required; Default: C **Format:** Alphanumeric

x) Treatment Pathways File

This input file is required when conducting a Type 6 analysis. It provides requester-defined attributes to the identification and computation of switch pattern episodes. Requester will specify in this file which group products will be evaluated for switching patterns and the number of switch patterns (up to two allowed). Identification of episodes that qualify as a switch (or not) will be in accordance with user-specified overlap and gap thresholds in this file.

Table 48. TreatmentPathways Specifications

Parameter	Field Name	Description
Name of Switch Pattern	ANALYSISGRP	**Details:** standardized name used to differentiate switch patterns. **Named by:** Request programmer **Input type:** Required **Format:** Alphanumeric; SAS character $40; no special characters (*e.g.,* commas, periods, hyphens, etc.) allowed, and underscores must be used to mark spaces. **Example:** SwitchPatternA
Name of Cohort	GROUP	**Details:** standardized name used to differentiate cohorts. **Note 1**: in the Type 6 File, GROUP represents the product group. When analyzing switch patterns, ANALYSISGRP may contain mulitiple GROUPs, with each group representing a product that is being analyzed for the ANALYSISGRP cohort. **Note 2:** GROUP is the primary key linking cohorts across input files; GROUP values must match (including case) between the TYPE6FILE and other input files. **Named by:** Request programmer **Input type:** Required **Format:** Alphanumeric; SAS character $40; no special characters (*e.g.,* commas, periods, hyphens, etc.) allowed, and underscores must be used to mark spaces. **Example:** Insulin
Allowed Number of Switch Pattern Episodes per Individual	SWITCHCOHORTDEF	**Details:** indicates how many switch pattern periods an individual can contribute. Options include: • **01:** Only the first valid switch pattern episode during the query period • **02:** All switch pattern episodes during the query period.

Parameter	Field Name	Description
		Note 1: This value must be the same for the same SWITCHPATTERN values. **Note 2**: If 02 is selected, there needs to be a corresponding GAPTOL value in SWITCHEVALSTEP value 0 row in order to indicate a period of an allowable gap to consider between multiple switching episodes per person. **Defined by:** Requester **Input type:** Required **Format:** Alphanumeric; SAS character $2 **Example:** 01
Switch evaluation step value	SWITCHEVALSTEP	**Details:** value used to differentiate evaluation step. Note that each switch pattern (SWITCHPATTERN) can support up to 2 evaluation steps. 0=Switch pattern evaluation start 1= first evaluation 2=second evaluation **Named by:** Request programmer **Input type:** Required **Format:** Numeric. Valid values include 0, 1 or 2. **Example:** 1
Gap Tolerance	GAPTOL	**Details:** value used to indicate allowable number of gap days in between treatment episodes identified in current and prior switch pattern evaluation steps, in order for treatment pattern to be identified as a switch. **Note 1:** values provided for SWITCHEVALSTEP value 0 will function as the tolerance threshold, in days, between the end of one switch pattern and when the tool will start looking for another one for the same patient. This is only relevant for SWITCHCOHORTDEF="02", where the tool will allow for more than one switch episode per patient. **Note 2:** values provided for GAPTOL values 1 and 2 will serve as a gap tolerance threshold between the current evaluation step and the prior evaluation step. This will be in effect for any SWITCHCOHORTDEF value. **Named by:** Request programmer **Input type:** Required **Format:** Numeric. **Example:** 30

Parameter	Field Name	Description
Overlap Tolerance	OVERLAPTOL	**Details:** value used to indicate allowable number of overlap days OR percent (of prior evaluation step treatment episode) in between treatment episodes identified in current and prior switch pattern evaluation steps, in order for treatment pattern to be identified as a switch. **Note 1:** should be null for SWITCHEVALSTEP value 0. **Named by:** Request programmer **Input type:** Required **Format:** Numeric. **Example:** 30
Overlap Type: Days or Percent	OVERLAPTYPE	**Details:** value to denote whether value in OVERLAPTOL represents a days or a percent. **Note 1:** should be null for SWITCHEVALSTEP value 0. **Note 2:** Allowable values are: • Days • Percent **Named by:** Request programmer **Input type:** Required **Format:** Alphanumeric; SAS character $7 **Example:** percent
Switch Pattern Cohort Inclusion Date	SWITCHCOHORTINCLDATE	**Details:** indicates which date to use for inclusion into the switch pattern cohort of interest as well as optionally as the index date of the treatment episode initiating the switch pattern. If SWITCHCOHORTINCLDATE value is provided, observed patterns of switching will only be counted as such if the SWITCHCOHORTINCLDATE occurs on or before the last day of the first treatment episode of the pattern, inclusive of the GAPTOL value provided for the SWITCHEVALSTEP SWITCHCOHORTINCLDATE. Optional parameter. If null, all switch pattern episodes will be included in analysis and the initial switch pattern step index date will be each patient's initial product RxDate. **Note 1:** only one value per unique SWITCHPATTERN value is allowed.Tool will issue custom warning if more than one row per unique SWITCHPATTERN value contain a non-null SWITCHCOHORTINCLDATE value. **Note 2**: values should be provided for the SWITCHEVALSTEP value that represents the GROUP

Parameter	Field Name	Description
		date that will be used. For example, if SWITCHEVALSTEP value 1 represents a generic product, and the approval date to use for this SWITCHCOHORTINCLDATE is that generic product's PRODUCTAPPROVALDATE, then PRODUCTAPPROVALDATE should be specified in the SWITCHEVALSTEP value 1 row for that SWITCHPATTERN value. **Valid values:** PRODUCTAPPROVALDATE PRODUCTMARKETINGDATE OTHERPRODUCTDATE COMPUTEDSTARTMARKETINGDATE **Named by:** Requester **Input type:** Required **Format:** Alphanumeric; SAS character $30 **Example:** PRODUCTAPPROVALDATE
Switch Pattern Cohort Inclusion Date Strategy Indicator	SWITCHDATEUSE	**Details:** indicates how the SWITCHCOHORTINCLDATE will be used. It is an optional parameter. It must be null if SWITCHCOHORTINCLDATE is null. It requires a value if SWITCHCOHORTINCLDATE is not null. The non-null SWITCHDATE value must be in the same row as the non-null SWITCHCOHORTINCLDATE. **Note 1:** values should be provided for the SWITCHEVALSTEP value that represents the GROUP date that will be used. For example, if SWITCHEVALSTEP value 1 represents a generic product, and the approval date to use for this SWITCHCOHORTINCLDATE is that generic product's PRODUCTAPPROVALDATE, then PRODUCTAPPROVALDATE should be specified in the SWITCHEVALSTEP value 1 row for that SWITCHPATTERN value. **Valid values:** **1** = SWITCHCOHORTINCLDATE used only as switch cohort entry date. First treatment episode RxDate is used as index for computing time to first switch. **2**= SWITCHCOHORTINCLDATE used as switch cohort entry date AND as initial switch step index date for computing time to first switch. **Named by:** Requester **Input type:** Required **Format:** Numeric; SAS character $2

Parameter	Field Name	Description
		Example: 1
Computation of switch pattern duration includes gaps?	SWITCHGAPINCL	**Details:** indicator for whether gaps in treatment episodes that are included in a switch episode will be counted as part of the switch episode duration. **Note 1:** This value must be the same for the same SWITCHPATTERN values. **Valid values:** **Y**=Yes, gaps between episodes will be counted as part of the overall switch pattern duration **N=**No, gaps between episodes will not be counted as part of the overall switch pattern duration **Defined by:** Requester **Input type:** Required **Format:** SAS alphanumeric $1 **Example:** 10
Minimum pre-index enrollment days for inclusion/exclusion criteria in switch episodes.	SWITCHENRDAYS	**Details:** parameter to specify the number of days of continuous enrollment required before the index date for inclusion/exclusion criteria when evaluating switch episodes. **Note 1**: if not specified, a default value of 0 days is used. **Named by:** Requester **Input type:** Optional **Format:** Numeric **Example:** 183
Minimum post-index enrollment days for inclusion/exclusion criteria in switch episodes.	SWITCHENRDAYSAFTIND	**Details:** parameter to specify the number of days of continuous enrollment required after the index date for inclusion/exclusion criteria when evaluating switch episodes. **Note 1:** may be left blank if no post-index enrollment is required. **Named by:** Requester **Input type:** Optional **Format:** Numeric **Example:** 183

y) Create Report File

Optional file to request a PDF report produced following the execution of CIDA. This file defines the parameters to customize the report. Table 49contains detailed specifications for this file for a Type 1 and Type 2 Report.

Table 49. CREATEREPORT_FILE Specifications

Parameter	Field Name	Description
Request ID	REQUESTID	**Details:** Workplan/Request ID for report. Used to name output files. **Note 1:** REQUESTID should be < 23 characters **Named by:** Requester **Input type:** Required **Format:** Char (23) **Example:** cber_mpl1p_wp001
Name of SAS input file to define groups	GROUPS_TABLE	**Details:** Input file defining groups to include in report, group headers, group labels, and group order **Named by:** Requester **Input type:** Required **Format:** Char (30) **Example:** report_groups_info
Name of SAS input file to define report columns	COLUMNS_TABLE	**Details:** Input file defining columns to include in report and customized column headers **Note 1:** Input file is optional. Only need to include if either 1) requesting the inclusion of non-default columns or 2) modifying the column header for either default or non-default columns **Named by:** Requester **Input type:** Optional **Format:** Char (30) **Example:** report_columns_info
Indicator for request type	TYPE	**Details:** Designates whether request is a type 1 or type 2 request. **Note 1:** Enter 1 for TYPE 1 and 2 for TYPE 2 **Named by:** Requester **Input type:** Required **Format:** Num (1) **Example:** 1
Name(s) of request package typefile input file	ALLTYPEFILES	**Details:** Name of typefile used in request package. Multiple typefile should be separated by space. **Named by:** Requester **Input type:** Required **Format:** Char (50)

Parameter	Field Name	Description
		Example: wp001_type2file wp002_type2file wp003_type2file
Name of request package monitoring input file	MONITORINGFILE	**Details:** Name of monitoringfile used in request package **Named by:** Requester **Input type:** Required **Format:** Char (30) **Example:** wp001_monitoring
Name of request package cohort input file	COHORTFILE	**Details:** Name of cohortfile used in request package **Named by:** Requester **Input type:** Required **Format:** Char (30) **Example:** wp001_cohort
Name of request package user strata input file	USERSTRATA	**Details:** Name of userstrata used in request package **Named by:** Requester **Input type:** Required **Format:** Char (30) **Example:** wp001_userstrata
Name to insert in title template	CUSTOMTITLE	**Details:** Customize report table titles by including request specific information in the table title. Standard title format is: "Summary of <customtitle> in the Sentinel Distributed Database between <start follow-up> and <end follow-up>. **Named by:** Requester **Input type:** Optional **Format:** Char (50) **Example:** incident beta-blocker users
Columns to exclude	EXCLUDE	**Details:** List of default columns to exclude from report. **Note 1:** Columns are identified using the following numbering scheme: 0. Do not exclude any columns 1. Exclude NPTS 2. Exclude EPISODES 3. Exclude ADJUSTEDCODECOUNT 4. Exclude RAWCODECOUNT 5. Exclude DAYSUPP 6. Exclude AMTSUPP 7. Exclude EPS_WEVENTS 8. Exclude ALL_EVENTS 9. Exclude TTE 10. Exclude TTE/365.25 11. Exclude DENNUMPTS 12. Exclude DENNUMMEMDAYS 13. Exclude DENNUMMEMDAYS/365.25

Parameter	Field Name	Description
		Note 2: Excluded column number should be separated by a space **Named by:** Requester **Input type:** Optional **Format:** Char (50) **Example:** 2 6 7
Stratification levels	STRATIFY_BY_LEVEL	**Details:** Stratification level values to include in report. **Note 1:** Separate stratification levels by a space **Note 2:** Refer to output_level_key.xlsx to determine stratification-level mapping scheme **Named by:** Requester **Input type:** Required **Format:** Char (50) **Example:** 000 002 003
Zip code lookup file	ZIPFILE	**Details:** If stratifying results by 3-digit zip code, include the name of the zip lookup file in order to map 3-digit zip code to state **Named by:** Requester **Input type:** Optional **Format:** Char (15) **Example:** zip_lkp
Labels for age groupings to display	AGEGROUPFMT	**Details:** Specify how age groups should be displayed, if age-group stratified output is produced. **Note 1:** For example, if "00-01" is the age group in CIDA and "< 1 year" is the desired display, the parameter should be entered as: %let agegroupfmt = "00-01" = "<1 year" **Note 2:** Repeat above for multiple entries. For example: %let agegroupfmt = "00-01" = "<1 year" "01-02" = "1-2 years"; **Named by:** Requester **Input type:** Optional **Format:** Char (100) **Example:** "00-01" = "<1 year" "01-02" = "1-2 years"
File name for report logo	LOGO	**Details:** Specify file name (including extension) for logo to display in report. **Note 1:** To ensure correct formatting, logo should be a JPG **Note 2:** If blank, no logo will be displayed **Named by:** Requester

Parameter	Field Name	Description
		Input type: Optional **Format:** Char (30) **Example:** Sentinel_logo.jpg
Produce baseline table	OUTPUT_BASELINETABLE	**Details:** Y/N indicator to produce baseline table. **Named by:** Requester **Input type:** Required **Format:** Char (1) **Example:** Y
Query Period Start Identifier	LOOK_START	**Details:** Identifies at what time period the report should begin outputting results **Note 1:** Only applicable to baseline table **Named by:** Requester **Input type:** Required **Format:** Num (1) **Example:** 1
Query Period End Identifier	LOOK_END	**Details:** Identifies at what time period the report should end outputting results **Note 1:** Only applicable to baseline table **Named by:** Requester **Input type:** Required **Format:** Num (1) **Example:** 1
Output CDF plot	OUTPUT_CDF_KM	**Details:** Y/N indicator to produce CDF plot (reasons for censor) and KM plot (time to event). **Note 1:** If OUTPUT_CDF_KM = Y, then the censor dataset must be returned to the MSOC **Named by:** Requester **Input type:** Required **Format:** Char (1) **Example:** N
Title of CDF plot	CDF_TITLE1	**Details:** Display title for CDF plot. Title will automatically include each group label. **Note 1:** If left blank, title will default to: Time to Censor <group label> **Named by:** Requester **Input type:** Optional **Format:** Char (50) **Example:** Follow-up time to censor
Title 1 of KM Plot	KM_TITLE1	**Details:** Display title 1 for KM plot. All groups will appear on the same plot.

Parameter	Field Name	Description
		Note 1: If left blank, title will default to: Time to Event **Named by:** Requester **Input type:** Optional **Format:** Char (50) **Example:** Days to Seizure
Title 2 of KM Plot	KM_TITLE2	**Details:** Optional title 2 for KM plot. Will display underneath title 1. **Note 1:** If left blank, only KM_TITLE1 will be displayed **Named by:** Requester **Input type:** Optional **Format:** Char (50) **Example:** for individuals with prior MI
Footnote 1 for CDF plot	CDF_FOOTNOTE1	**Details:** Optional footnote to add to CDF plot **Note 1:** If left blank, default footnote is: "A single episode may contribute to multiple categories if a patient was censored due to multiple criteria on the same day" **Named by:** Requester **Input type:** Optional **Format:** Char (50) **Example:** Individuals were censored at the earliest of: 1) end of episode, 2) death, 3) end of enrollment
Footnote 2 for CDF plot	CDF_FOOTNOTE2	**Details:** Optional footnote 2 to add to CDF plot **Named by:** Requester **Input type:** Optional **Format:** Char (50) **Example:** Individuals were censored at the earliest of: 1) end of episode, 2) death, 3) end of enrollment
Footnote 1 for KM plot	KM_FOOTNOTE1	**Details:** Optional footnote 1 to add to KM plot **Note 1:** If not specified, no footnote will be displayed **Named by:** Requester **Input type:** Optional **Format:** Char (50) **Example:** Seizures were considered only in the inpatient setting
Footnote 2 for KM plot	KM_FOOTNOTE2	**Details:** Optional footnote 1 to add to KM plot **Named by:** Requester **Input type:** Optional **Format:** Char (50) **Example:** Seizures were considered only in the inpatient setting

Parameter	Field Name	Description
Minimum X axis value for CDF plot	CDF_XMIN	**Details:** Requester can optionally specify the minimum x-axis value on the CDF plot **Note 1:** If CDF_XMIN is specified, CDF_XMAX and CDF_XTICK must also be specified. IF CDF_XMIN is blank, CDF_XMAX and CDF_XTICK must also be blank **Note 2:** Default minimum x-axis value is 0 **Named by:** Requester **Input type:** Optional **Format:** Num (8) **Example:** 50
Maximum X axis value for CDF plot	CDF_XMAX	**Details:** Requester can optionally specify the maximum x-axis value on the CDF plot **Note 1:** If CDF_XMAX is specified, CDF_XMIN and CDF_XTICK must also be specified. IF CDF_XMAX is blank, CDF_XMIN and CDF_XTICK must also be blank **Note 2:** Default maximum x-axis value is the max of data **Named by:** Requester **Input type:** Optional **Format:** Num (8) **Example:** 100
X-axis tick marks for CDF plot	CDF_XTICK	**Details:** Requester can optionally specify the distance between x-axis tick marks on the CDF plot **Note 1:** If CDF_XTICK is specified, CDF_XMIN and CDF_XMAX must also be specified. IF CDF_XTICK is blank, CDF_XMIN and CDF_XMAX must also be blank **Note 2:** Default tick value is determined by SAS **Note 3:** If CDF_XMIN = 0, CDF_XMAX = 200, and CDF_XTICK = 50, this will create x-axis tick values of (0 50 100 150 200) **Named by:** Requester **Input type:** Optional **Format:** Num (8) **Example:** 20
Minimum X axis value for KM plot	KM_XMIN	**Details:** Requester can optionally specify the minimum x-axis value on the KM plot **Note 1:** If KM_XMIN is specified, KM_XMAX and KM_XTICK must also be specified. IF KM_XMIN is blank, KM_XMAX and KM_XTICK must also be blank **Note 2:** Default minimum x-axis value is 0 **Named by:** Requester

Parameter	Field Name	Description
		Input type: Optional **Format:** Num (8) **Example:** 50
Maximum X axis value for KM plot	KM_XMAX	**Details:** Requester can optionally specify the maximum x-axis value on the KM plot **Note 1:** If KM_XMAX is specified, KM_XMIN and KM_XTICK must also be specified. IF KM_XMAX is blank, KM_XMIN and KM_XTICK must also be blank **Note 2:** Default maximum x-axis value is the max of data **Named by:** Requester **Input type:** Optional **Format:** Num (8) **Example:** 100
X-axis tick marks for KM plot	KM_XTICK	**Details:** Requester can optionally specify the distance between x-axis tick marks on the KM plot **Note 1:** If KM_XTICK is specified, KM_XMIN and KM_XMAX must also be specified. IF KM_XTICK is blank, KM_XMIN and KM_XMAX must also be blank **Note 2:** Default tick value is determined by SAS **Note 3:** If KM_XMIN = 0, KM_XMAX = 200, and KM_XTICK = 50, this will create x-axis tick values of (0 50 100 150 200) **Named by:** Requester **Input type:** Optional **Format:** Num (8) **Example:** 20
Minimum Y axis value for CDF plot	CDF_YMIN	**Details:** Requester can optionally specify the minimum y-axis value on the CDF plot **Note 1:** If CDF_YMIN is specified, CDF_YMAX and CDF_YTICK must also be specified. IF CDF_YMIN is blank, CDF_YMAX and CDF_YTICK must also be blank **Note 2:** Default minimum y-axis value is 0 **Named by:** Requester **Input type:** Optional **Format:** Num (8) **Example:** 0.80
Maximum Y axis value for CDF plot	CDF_YMAX	**Details:** Requester can optionally specify the maximum y-axis value on the CDF plot

Parameter	Field Name	Description
		Note 1: If CDF_YMAX is specified, CDF_YMIN and CDF_YTICK must also be specified. IF CDF_YMAX is blank, CDF_YMIN and CDF_YTICK must also be blank **Note 2:** Default maximum y-axis value is 1 **Named by:** Requester **Input type:** Optional **Format:** Num (8) **Example:** 0.90
Y-axis tick marks for CDF plot	CDF_YTICK	**Details:** Requester can optionally specify the distance between y-axis tick marks on the CDF plot **Note 1:** If CDF_YTICK is specified, CDF_YMIN and CDF_YMAX must also be specified. IF CDF_YTICK is blank, CDF_YMIN and CDF_YMAX must also be blank **Note 2:** Default tick value is determined by SAS **Note 3:** If CDF_YMIN = .5, CDF_YMAX = 1, and CDF_YTICK = .1, this will create x-axis tick values of (.5 .6 .7 .8 .9 1) **Named by:** Requester **Input type:** Optional **Format:** Num (8) **Example:** .1
Minimum Y axis value for KM plot	KM_YMIN	**Details:** Requester can optionally specify the minimum y-axis value on the KM plot **Note 1:** If KM_YMIN is specified, KM_YMAX and KM_YTICK must also be specified. IF KM_YMIN is blank, KM_YMAX and KM_YTICK must also be blank **Note 2:** Default minimum y-axis value is 0 **Named by:** Requester **Input type:** Optional **Format:** Num (8) **Example:** 0.80
Maximum Y axis value for KM plot	KM_YMAX	**Details:** Requester can optionally specify the maximum y-axis value on the KM plot **Note 1:** If KM_YMAX is specified, KM_YMIN and KM_YTICK must also be specified. IF KM_YMAX is blank, KM_YMIN and KM_YTICK must also be blank **Note 2:** Default maximum y-axis value is 1 **Named by:** Requester **Input type:** Optional **Format:** Num (8) **Example:** 0.90

Parameter	Field Name	Description
Y-axis tick marks for KM plot	KM_YTICK	**Details:** Requester can optionally specify the distance between y-axis tick marks on the KM plot **Note 1:** If KM_YTICK is specified, KM_YMIN and KM_YMAX must also be specified. IF KM_YTICK is blank, KM_YMIN and KM_YMAX must also be blank **Note 2:** Default tick value is determined by SAS **Note 3:** If KM_XMIN = .5, KM_XMAX = 1, and KM_XTICK = .1, this will create x-axis tick values of (.5 .6 .7 .8 .9 1) **Named by:** Requester **Input type:** Optional **Format:** Num (8) **Example:** .1
Minimum X axis value for time to censor KM plot	KM_EP_XMIN	**Details:** Requester can optionally specify the minimum x-axis value on the time to censor KM plot **Note 1:** KM_EP_XMIN is specified, KM_EP_XMAX and KM_EP_XTICK must also be specified. IF KM_EP_XMIN is blank, KM_EP_XMAX and KM_EP_XTICK must also be blank **Note 2:** Default minimum x-axis value is 0 **Named by:** Requester **Input type:** Optional **Format:** Num (8) **Example:** 50
Maximum X axis value for time to censor KM plot	KM_EP_XMAX	**Details:** Requester can optionally specify the maximum x-axis value on the time to censor KM plot **Note 1:** If KM_EP_XMAX is specified, KM_EP_XMIN and KM_EP_XTICK must also be specified. IF KM_EP_XMAX is blank, KM_EP_XMIN and KM_EP_XTICK must also be blank **Note 2:** Default maximum x-axis value is the max of data **Named by:** Requester **Input type:** Optional **Format:** Num (8) **Example:** 100
X-axis tick marks for time to censor KM plot	KM_EP_XTICK	**Details:** Requester can optionally specify the distance between x-axis tick marks on the time to censor KM plot **Note 1:** If KM_EP_XTICK is specified, KM_EP_XMIN and KM_EP_XMAX must also be specified. IF KM_EP_XTICK is blank, KM_EP_XMIN and KM_EP_XMAX must also be blank

Parameter	Field Name	Description
		Note 2: Default tick value is determined by SAS **Note 3:** If KM_EP_XMIN = 0, KM_EP_XMAX = 200, and KM_EP_XTICK = 50, this will create x-axis tick values of (0 50 100 150 200) **Named by:** Requester **Input type:** Optional **Format:** Num (8) **Example:** 20
Minimum Y axis value for time to censor KM plot	KM_EP_YMIN	**Details:** Requester can optionally specify the minimum y-axis value on the time to censor KM plot **Note 1:** If KM_EP_YMIN is specified, KM_EP_YMAX and KM_EP_YTICK must also be specified. IF KM_EP_YMIN is blank, KM_EP_YMAX and KM_EP_YTICK must also be blank **Note 2:** Default minimum y-axis value is 0 **Named by:** Requester **Input type:** Optional **Format:** Num (8) **Example:** 0.80
Maximum Y axis value for time to censor KM plot	KM_EP_YMAX	**Details:** Requester can optionally specify the maximum y-axis value on the time to censor KM plot **Note 1:** If KM_EP_YMAX is specified, KM_EP_YMIN and KM_EP_YTICK must also be specified. IF KM_EP_YMAX is blank, KM_EP_YMIN and KM_EP_YTICK must also be blank **Note 2:** Default maximum y-axis value is 1 **Named by:** Requester **Input type:** Optional **Format:** Num (8) **Example:** 0.90
Y-axis tick marks for time to censor KM plot	KM_EP_YTICK	**Details:** Requester can optionally specify the distance between y-axis tick marks on the time to censor KM plot **Note 1:** If KM_EP_YTICK is specified, KM_EP_YMIN and KM_EP_YMAX must also be specified. IF KM_EP_YTICK is blank, KM_EP_YMIN and KM_EP_YMAX must also be blank **Note 2:** Default tick value is determined by SAS **Note 3:** If KM_EP_XMIN = .5, KM_EP_XMAX = 1, and KM_EP_XTICK = .1, this will create x-axis tick values of (.5 .6 .7 .8 .9 1)

Parameter	Field Name	Description
		Named by: Requester **Input type:** Optional **Format:** Num (8) **Example:** .1
Censoring criteria to display on CDF plot	CENSORING_DISPLAY	**Details:** List of censoring criteria to display on CDF plot **Note 1:** If left blank, all censoring criteria will be displayed **Note 2:** Separate censoring criteria by a space **Note 3:** Current options for censoring are: 1) cens_elig, 2) cens_dth, 3) cens_dpend, 4) cens_qryend, 5) cens_episend, 6) cens_spec, and 7) cens_event **Named by:** Requester **Input type:** Optional **Format:** Char (40) **Example:** cens_elig cens_dth cens_dpend cens_event
Label for eligibility censoring	CENS_ELIG	**Details:** Label to display for individuals censored due to end of enrollment **Note 1:** If blank, default label is "Disenrollment" **Named by:** Requester **Input type:** Optional **Format:** Char (20) **Example:** End of Enrollment
Label for death censoring	CENS_DTH	**Details:** Label to display for individuals censored due to death **Note 1:** If blank, default label is "Evidence of Death" **Named by:** Requester **Input type:** Optional **Format:** Char (20) **Example:** Death
Label for end of data partner data censoring	CENS_DPEND	**Details:** Label to display for individuals censored due to end of available data partner data **Note 1:** If blank, default label is "End of Data Partner Data" **Named by:** Requester **Input type:** Optional **Format:** Char (20) **Example:** End of available data
Label for end of query period censoring	CENS_QRYEND	**Details:** Label to display for individuals censored due to end of query period **Note 1:** If blank, default label is "End of Query Period" **Named by:** Requester

Parameter	Field Name	Description
		Input type: Optional **Format:** Char (20) **Example:** End of query period (9/30/2015)
Label for end of treatment episode censoring	CENS_EPISEND	**Details:** Label to display for individuals censored due to end of treatment episode **Note 1:** If blank, default label is "End of Exposure Episode" **Named by:** Requester **Input type:** Optional **Format:** Char (20) **Example:** End of beta-blocker treatment
Label for requester-specified censoring criteria	CENS_SPEC	**Details:** Label to display for individuals censored due requester-specified episode truncation criteria **Note 1:** If blank, default label is "Occurrence of request-defined censoring criteria" **Named by:** Requester **Input type:** Optional **Format:** Char (20) **Example:** Evidence of stroke
Label for event censoring	CENS_EVENT	**Details:** Label to display for individuals censored due to occurrence of event **Note 1:** If blank, default label is "Occurrence of Event" **Named by:** Requester **Input type:** Optional **Format:** Char (20) **Example:** Event of interest
Display number of episodes	DISPLAYN	**Details:** Y/N indicator to display number of episodes within CDF/KM plots **Named by:** Requester **Input type:** Required **Format:** Char (1) **Example:** N
Spacing between table rows	LINE_SPACING	**Details:** Defines the amount of white space between table rows. This parameter can be modified to increase or decrease white space, primarily to prevent orphan rows **Note 1:** Set to 1.75 as default **Named by:** Requester **Input type:** Required **Format:** Num (8) **Example:** 1.75

z) Groups Table

Required file to request a Type 1 and Type 2 PDF report produced following the execution of CIDA. This file defines the scenarios to include in the report, and to customize each scenario. Table 50 contains detailed specifications for this file.

Table 50. GROUPS_TABLE Specifications

Parameter	Field Name	Description
Label to group multiple groups	HEADER	**Details:** Label to group multiple groups under a header **Note 1:** If left blank, all groups will be displayed together with no header **Named by:** Requester **Input type:** Optional **Format:** Char (60) **Example:** With 183 day washout
Name of group to include in report	GROUP1	**Details:** Group name to include in report **Named by:** Requester **Input type:** Required **Format:** Char (30) **Example:** beta_blocker
RunID which contains group name from GROUP1	RUNID1	**Details:** runID assigned to the run that corresponds to the group name listed in GROUP1 **Named by:** Requester **Input type:** Required **Format:** Char (10) **Example:** r01
Name of group to combine with group1	GROUP2	**Details:** If a combined group is being produced, list the second group name **Note 1:** Leave blank if a combined group is not being created **Named by:** Requester **Input type:** Optional **Format:** Char (30) **Example:** beta_blocker_wexl
RunID which contains group name from GROUP2	RUNID2	**Details:** runID assigned to the run that corresponds to the group name listed in GROUP2 **Note 1:** Leave blank if GROUP2 is blank. Required if GROUP2 is specified **Named by:** Requester **Input type:** Optional **Format:** Char (10) **Example:** r02

Parameter	Field Name	Description
Group name to assign to combined group	COMBINEDGROUPNAME	**Details:** Custom group name for new combined groups. **Note 1:** If scenario only consists of 1 group (i.e. group2 and runid2 are blank), COMBINEDGROUPNAME should be blank **Note 2:** This parameter is assigned by the requester to distinguish groups during program execution. It is not the display name for the group. **Named by:** Requester **Input type:** Optional **Format:** Char (50) **Example:** Beta_blocker_lisinopril
Group label to display	GROUPLABEL	**Details:** Display name for group in report **Note 1:** If blank, raw group name (i.e. GROUP1 name) will be displayed in report **Named by:** Requester **Input type:** Optional **Format:** Char (50) **Example:** Beta Blockers
Group display order	ORDER	**Details:** Order to display groups in report. All groups should receive a unique order value **Named by:** Requester **Input type:** Required **Format:** Num (8) **Example:** 1
Covariates to display under History of Use	HISTORYOFUSE	**Details:** To include a history of use section in the baseline table, include a list of covariates to include in this section, separated by a comma. **Note 1:** By default, all covariates will be grouped under 'Recorded History of' in the output and all utilization vars will be excluded. **Note 2:** If specified, only covariates specifically listed in HISTORYOFUSE, RECORDEDHISTORY, and UTILIZATIONINTENSITY will be displayed **Note 3:** Can use 'dash' notation (i.e. COVAR1-COVAR9) **Named by:** Requester **Input type:** Optional **Format:** Char (50) **Example:** covar1, covar1, covar3

Parameter	Field Name	Description
Covariates to display under Recorded History	RECORDEDHISTORY	**Details:** Include a list of covariates to include under 'Recorded History of' **Note 1:** By default, all covariates will be grouped under 'Recorded History of' in the output and all utilization vars will be excluded. **Note 2:** If specified, only covariates specifically listed in HISTORYOFUSE, RECORDEDHISTORY, and UTILIZATIONINTENSITY will be displayed **Note 3:** Can use 'dash' notation (i.e. COVAR1-COVAR9) **Named by:** Requester **Input type:** Optional **Format:** Char (50) **Example:** covar1, covar1, covar3
Covariates to display under Utilization History	UTILIZATIONINTENSITY	**Details:** To include a utilization history section in the baseline table, include a list of covariates to include in this section, separated by a comma. **Note 1:** By default, all covariates will be grouped under 'Recorded History of' in the output and all utilization vars will be excluded. **Note 2:** If specified, only covariates specifically listed in HISTORYOFUSE, RECORDEDHISTORY, and UTILIZATIONINTENSITY will be displayed **Note 3:** Can use 'dash' notation (i.e. COVAR1-COVAR9) **Named by:** Requester **Input type:** Optional **Format:** Char (50) **Example:** numav, numip, numed, numoa, numis, numclass, numgeneric, numrx
Covariates to italicize in baseline table	HIGHLIGHT_VARS	**Details:** List of covariates, utilization, and characteristic variables to italicize in baseline table, separated by a space **Named by:** Requester **Input type:** Optional **Format:** Char (60) **Example:** covar1 covar2 covar25
Covariate sort order	ALPHABETICAL_COVARSORT	**Details:** Y/N indicator to determine whether to sort covariates by alphabetical order (Y) or covarnum (N) **Named by:** Requester

Parameter	Field Name	Description
		Input type: Required **Format:** Char (1) **Example:** Y
Baseline table group label	BASELINELABEL	**Details:** Group-specific label for baseline table. **Note 1:** If blank, GROUPLABEL will be used in baseline table. **Named by:** Requester **Input type:** Optional **Format:** Char (50) **Example:** Beta Blockers

aa) Columns Table

Optional file to include when requesting a Type 1 and Type 2 PDF report produced following the execution of CIDA. This file defines the summary columns and customizes column headers in the report. Table 51 contains detailed specifications for this file.

Table 51. COLUMNS_TABLE Specifications

Parameter	Field Name	Description
Standardized column notation	VARIABLE	**Details:** Standardized name for report column **Named by:** Requester **Input type:** Required **Format:** Char (60) **Example:** DaySupp/Npts
Column Header	DESCRIPTION	**Details:** Display header for VARIABLE **Named by:** Requester **Input type:** Required **Format:** Char (30) **Example:** Day Supply per Patient

B. OUTPUT

1. MSOC Folder

The CIDA tool generates output to the MSOC folder based on the type of analysis specified.

Background Rate Calculations (Type 1 Analyses)

This analysis generates up to six output tables:

- [*RUNID*]_baseline_[*PERIODID*].sas7bdat
- [*RUNID*]_t1_cida.sas7bdat
- [*RUNID*]_signature.sas7bdat
- [*RUNID*]_attrition.sas7bdat
- [*RUNID*]_MFU.sas7bdat (if MFU analysis is requested)
- [*RUNID*]_profile_[*PERIODID*] (if covariate profile is requested)

If reason for censoring eligibility output is specified, the analysis generates an additional table:

- [*RUNID*]_censor_CIDA.sas7bdat

If the output for index code distribution is specified, the analysis generates two additional tables:

- [*RUNID*]_distindex.sas7bdat
- [*RUNID*]_distindexmap.sas7bdat

Where RUNID is the request programmer-defined execution identifier (defined in main macro parameters) and PERIODID is the time period under analysis specified in the Monitoring File.

Exposures and Follow-up (Type 2 Analyses)

This analysis generates up to eight output tables:

- [*RUNID*]_baseline_[*PERIODID*].sas7bdat
- [*RUNID*]_t2_cida.sas7bdat
- [*RUNID*]_signature.sas7bdat
- [*RUNID*]_attrition.sas7bdat
- [*RUNID*]_MFU.sas7bdat (if MFU analysis is requested)
- [*RUNID*]_profile_[*PERIODID*] (if covariate profile is requested)

Where RUNID is the requester-defined execution identifier (defined in main macro parameters) and PERIODID is the time period under analysis specified in the Monitoring File.

If reason for censoring follow-uptime output is specified, the analysis generates an additional table:

- [*RUNID*]_censor_CIDA.sas7bdat

If the output for index code distribution is specified, the analysis generates two additional tables:

- [*RUNID*]_distindex.sas7bdat
- [*RUNID*]_distindexmap.sas7bdat

Self-controlled Risk Interval Design (Type 3 Analyses)

This analysis generates up to ten output tables:

- [*RUNID*]_baseline_[*PERIODID*].sas7bdat
- [*RUNID*]_baseline_an_[*PERIODID*].sas7bdat (if baseline table is requested)
- [*RUNID*]_baseline_an_censor_[*PERIODID*].sas7bdat (if baseline table is requested)
- [*RUNID*]_t3_cida.sas7bdat
- [*RUNID*]_signature.sas7bdat
- [*RUNID*]_attrition.sas7bdat
- [*RUNID*]_metadata_for_time_period_[*PERIODID*].sas7bdat
- [*RUNID*]_MFU.sas7bdat (if MFU analysis is requested)
- [*RUNID*]_MFU_an.sas7bdat (if MFU analysis is requested)
- [*RUNID*]_MFU_an_censor.sas7bdat (if MFU analysis is requested)
- [*RUNID*]_profile_[*PERIODID*] (if covariate profile is requested)
- [*RUNID*]_profile_an_[*PERIODID*] (if covariate profile is requested)
- [*RUNID*]_profile_an_censor_[*PERIODID*] (if covariate profile is requested)

If the output for index code distribution is specified, the analysis generates two additional tables:

- [*RUNID*]_distindex.sas7bdat
- [*RUNID*]_distindexmap.sas7bdat

Where RUNID is the request programmer-defined execution identifier (defined in main macro parameters) and PERIODID and the time period under analysis specified in the Monitoring File.

Pregnancy Episodes and Medical Product Use (Type 4 Analyses)

This analysis generates up to eight output tables:

- [*RUNID*]_baseline_Preg_[*PERIODID*].sas7bdat
- [*RUNID*]_baseline_NoPreg_[*PERIODID*].sas7bdat
- [*RUNID*]_baseline_MI_[*PERIODID*].sas7bdat
- [*RUNID*]_signature.sas7bdat
- [*RUNID*]_t4_cida_elig.sas7bdat
- [*RUNID*]_t4_cida_Preg.sas7bdat
- [*RUNID*]_t4_cida_NoPreg.sas7bdat
- [RUNID]_t4_Preg_gestwk.sas7bdat
- [RUNID]_t4_NoPreg_gestwk.sas7bdat
- [*RUNID*]_MFU.sas7bdat (if MFU analysis is requested)
- [*RUNID*]_profile_Preg_[*PERIODID*] (if covariate profile is requested)
- [*RUNID*]_profile_NoPreg_[*PERIODID*] (if covariate profile is requested)

If the output for index code distribution is specified, the analysis generates two additional tables:

- [*RUNID*]_distindex.sas7bdat
- [*RUNID*]_distindexmap.sas7bdat

Where RUNID is the request programmer-defined execution identifier (defined in main macro parameters).

Medical Product Utilization (Type 5 Analyses)

This analysis generates up to eleven output tables:

- [*RUNID*]_baseline_[*PERIODID*].sas7bdat (if baseline table is requested)
- [*RUNID*]_t5_cida_ disp_by_daysupp.sas7bdat
- [*RUNID*]_t5_cida_ firsteps
- [*RUNID*]_t5_cida_alleps
- [*RUNID*]_t5_cida_episdur.sas7bdat
- [*RUNID*]_t5_cida_episdur_censor.sas7bdat
- [*RUNID*]_t5_cida_gaps.sas7bdat
- [*RUNID*]_signature.sas7bdat
- [*RUNID*]_attrition.sas7bdat
- [*RUNID*]_MFU.sas7bdat (if MFU analysis is requested)
- [*RUNID*]_profile_[*PERIODID*] (if covariate profile is requested)

If the output for index code distribution is specified, the analysis generates two additional tables:

- [*RUNID*]_distindex.sas7bdat
- [*RUNID*]_distindexmap.sas7bdat

Where RUNID is the request programmer-defined execution identifier (defined in main macro parameters) and PERIODID and the time period under analysis specified in the Monitoring File.

All output tables for all types of analyses are described below.

Manufacturer-level Product Utilization and Switching Patterns Cohort Identification Strategy (Type 6 Analyses)

Type 6 analysis generates up to 14 output tables that will be returned to SOC for aggregation and reporting.

- [*RUNID*]_t6_productsdates.sas7bdat
- [*RUNID*]_attrition.sas7bdat
- [*RUNID*]_t6_utilcounts.sas7bdat
- [*RUNID*]_t6_trendcounts.sas7bdat
- [*RUNID*]_t6_utildispstats.sas7bdat
- [*RUNID*]_t6_utilepis_censor.sas7bdat
- [*RUNID*]_t6_utilepisdurstats.sas7bdat
- [*RUNID*]_t6_utiluptakestats.sas7bdat
- [*RUNID*]_t6_switchattrition.sas7bdat
- [*RUNID*]_t6_switchplota.sas7bdat
- [*RUNID*]_t6_switchplotb.sas7bdat
- [*RUNID*]_t6_switchepisdurstats.sas7bdat
- [RUNID]_signature.sas7bdat

a) [*RUNID*]_T1_CIDA.sas7bdat

The [*RUNID*]_T1_CIDA output table includes the number of individuals, index dates, dispensings, dispensing days supplied, dispensing amount supplied, eligible members and eligible member days. All metrics are reported overall and stratified by age group, sex, year, and year-month. Table 52 contains specifications for the [*RUNID*]_T1_CIDA output table.

Table 52. [RUNID]_T1_CIDA Output

Variable	Description
GROUP	Cohort name. **Format:** Character(30)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character(3)
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(1)
RACE	Race. Allowable values are those in the SCDM. **Format:** Character(1)
HISPANIC	A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race. Allowable values are those in the SCDM. **Format:** Character(1)
AGEGROUP	Age Groups. Categories are requester-defined. **Format:** Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
YEAR	Year of index date. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)

Variable	Description
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date N: zip date occurred before the index date **Format:** Character(1)
COVAR*n*	Requester-defined covariate **Note 1:** N will equal COVARNUM specified in the Covariate file Format: Numeric
NPTS	Number of patients. **Format:** Numeric
EPISODES	Number of index date defining records. **Format:** Numeric
ADJUSTEDCODECOUNT	Number of dispensings used to define index date defining records. **Note 1:** When multiple dispensings occur on the same day with same STOCKGROUP, this will be counted as 1 dispensing. **Format:** Numeric
RAWCODECOUNT	Number of dispensings used to define index date defining records. **Note 1:** When multiple dispensings occur on the same day with same STOCKGROUP, each dispensing will be counted. **Format:** Numeric
DAYSUPP	Total days supply associated with dispensings. **Format:** Numeric
AMTSUPP	Total amount supplied associated with dispensings. **Format:** Numeric
EPS_WEVENTS	Will be 0 for a Type 1 analysis. **Format:** Numeric
ALL_EVENTS	Will be 0 for a Type 1 analysis. **Format:** Numeric
TTE	Will be 0 for a Type 1 analysis. **Format:** Numeric
DENNUMPTS	Number of patients eligible to have at least one index date. **Format:** Numeric
DENNUMMEMDAYS	Number of days that patients are eligible to have an index date. **Format:** Numeric

b) *[RUNID]*_censor_CIDA.sas7bdat (Type 1 Analysis)

The [*RUNID]*_censor_CIDA output table includes information on the number of episodes reason for censoring eligibility for every day of follow-up. Table 53 contains specifications for the [*RUNID]* _censor_CIDA output table.

Table 53. [RUNID]_censor_CIDA Output

Variable	Description
GROUP	Cohort name. **Format:** Character (30)
LEVEL	Stratification identifier. Each unique combination of strata (i.e., variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character (3)
CENSDAYS_VALUE	Stratification variable. Represents number of days from index date to censoring for the following four reasons: • End of query period • End of Data Partner (DP) data • Disenrollment • Evidence of death Calculated as "censor date" – indexdt + 1 (episodes censored on indexdt have CENSDAYS_VALUE = 1). **Format**: Numeric (8)
SEX	Sex. Allowable values are "M" (Male), "F" (Female) and "O" (Other). **Format:** Character (1)
AGEGROUP	Age Groups. Categories are requester-defined. **Format:** Character(variable)
YEAR	Year of index date. **Format:** Numeric
CENSOR_OUTPUT_CAT	Stratification variable. Categorizes CENSDAYS_VALUE variable in user-defined ranges. **Format**: Character (10)
EPISODES	Number of episodes by CENSDAYS_VALUE. **Format:** Numeric (8)
CENS_ELIG	Number of episodes censored due to disenrollment. **Format:** Numeric (8)
CENS_DTH	Number of episodes censored due to evidence of death. **Format:** Numeric (8)

Variable	Description
CENS_DPEND	Number of episodes censored due to DP data end date (based on DP_MaxDate in common components). **Format:** Numeric (8)
CENS_QRYEND	Number of episodes censored due to query end date. **Format:** Numeric (8)

c) [*RUNID*]_T2_CIDA.sas7bdat

The [*RUNID*]_T2_CIDA output table includes the number of individuals, exposure episodes, dispensings, dispensing days supplied, dispensing amount supplied, HOIs, days at-risk, eligible members and eligible member days. All metrics are reported overall and stratified by age group, sex, year, and year-month. Table 54 contains specifications for the [*RUNID*]_T2_CIDA output table.

Table 54. [RUNID]_T2_CIDA Output

Variable	Description
GROUP	Cohort name. **Format:** Character(30)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.*, variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character(3)
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(1)
RACE	Race. Allowable values are those in the SCDM. **Format:** Character(1)
HISPANIC	A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race. Allowable values are those in the SCDM. **Format:** Character(1)
AGEGROUP	Age Groups. Categories are requester-defined. **Format:** Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
YEAR	Year of index date. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric

Variable	Description
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date N: zip date occurred before the index date **Format:** Character(1)
COVAR*n*	Requester-defined covariate **Note 1:** N will equal COVARNUM specified in the Covariate file **Format**: Numeric
NPTS	Number of patients. **Format:** Numeric
EPISODES	Number of exposure episodes. **Format:** Numeric
ADJUSTEDCODECOUNT	Number of dispensings used to define index date defining records. **Note 1:** When multiple dispensings occur on the same day with same STOCKGROUP, this will be counted as 1 dispensing. This count will equal the count of the DISPENSINGS metric in prior QRP versions. **Format:** Numeric
RAWCODECOUNT	Number of dispensings used to define index date defining records. **Note 1:** When multiple dispensings occur on the same day with same STOCKGROUP, each dispensing will be counted. **Format:** Numeric
DAYSUPP	Total days supplied for outpatient pharmacy dispensings used to build the exposure episode. For requester-defined follow-up time (i.e., when exposure episodes are *not* created using dispensing days supply), this value will always be populated with the RxSup value associated with the dispensing that defined the index date. **Note 1:** Value will always be =0 for never-exposed cohort. **Format:** Numeric
AMTSUPP	Total amount supplied for outpatient pharmacy dispensings used to build the exposure episode. For requester-defined follow-up time (i.e., when exposure episodes are *not* created using dispensing days supply), this value

Variable	Description
	will always be populated with the RxAmt value associated with the dispensing that defined the index date. **Note 1:** Value will always be =0 for never-exposed cohort. **Format:** Numeric
EPS_WEVENTS	Number of exposure episodes with an HOI. **Format:** Numeric
ALL_EVENTS	Total number of HOIs in all exposure episodes. **For characterization purposes only.** Days-at-risk stop accumulating after the first HOI during an exposure episode. ALL_EVENTS/TTE should never be calculated. This variable value just reports the number of times during treatment episodes that the HOI definition was met. **Format:** Numeric
TTE	Days at-risk. **Format:** Numeric
DENNUMPTS	Number of patients eligible to have at least one exposure episode. **Note 1:** For requests that will use the prospective surveillance with propensity score matched design, this will be blank. **Note 2:** Value will always be blank for never-exposed cohort. **Format:** Numeric
DENNUMMEMDAYS	Number of days that patients are eligible to have an exposure episode. **Note 1:** For requests that will use the prospective surveillance with propensity score matched design, this will be blank. **Note 2:** Value will always be blank for never-exposed cohort. **Format:** Numeric

d) [*RUNID*]_censor_CIDA.sas7bdat (Type 2 Analysis)

The [*RUNID*]_censor_CIDA output table includes information on the number of episodes, event status, and reason for censoring for every day of follow-up. Table 55 contains specifications for the [*RUNID*]_censor_CIDA output table.

Table 55. [RUNID]_censor_CIDA Output

Variable	Description
GROUP	Cohort name. **Format:** Character(30)
LEVEL	Stratification identifier. Each unique combination of strata (i.e., variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character (3)
CENSDAYS_VALUE	Stratification variable. Represents number of days from index date to censoring for the following seven reasons: • Occurrence of event of interest • End of query period • End of Data Partner (DP) data • Disenrollment • End of exposure episode • Occurrence of requester-defined censoring criteria • Evidence of death Calculated as "censor date" – indexdt + 1 (episodes censored on indexdt have CENSDAYS_VALUE = 1). **Format:** Numeric (8)
EVENT_FLAG	Stratification variable. Identifies if CENSDAYS is determined by occurrence of an event. Allowable values: Y: Episode was censored due to occurrence of an event (CENSDAYS_VALUE is time-to-event) N: Episode was censored due to reason other than occurrence of an event **Format:** Character(1)
CENSOR_OUTPUT_CAT	Stratification variable. Categorizes CENSDAYS_VALUE variable in user-defined ranges. **Format:** Character (10)
SEX	Sex. Allowable values are "M" (Male), "F" (Female) and "O" (Other). **Format:** Character (1)
AGEGROUP	Age Groups. Categories are requester-defined. **Format:** Character(variable)

Variable	Description
YEAR	Year of index date. **Format:** Numeric
EPISODES	Number of episodes **Format:** Numeric (8)
CENS_ELIG	Number of episodes censored due to disenrollment. **Format:** Numeric (8)
CENS_DTH	Number of episodes censored due to evidence of death. **Format:** Numeric (8)
CENS_DPEND	Number of episodes censored due to DP data end date (based on DP_MaxDate in common components). **Format:** Numeric (8)
CENS_QRYEND	Number of episodes censored due to query end date. **Format:** Numeric (8)
CENS_EPISEND	Number of episodes censored due to episode end date. **Format:** Numeric (8)
CENS_SPEC	Number of episodes censored due to additional requester-defined criteria (e.g., censor due to occurrence of another set of clinical codes). **Format:** Numeric (8)
CENS_EVENT	Number of episodes censored due to occurrence of request-defined event **Format:** Numeric (8)

e) [*RUNID*]_t2_concomitance.sas7bdat

Table 56 characterizes the episodes of concomitant use and events of interest during these episodes. It includes number of episodes, users, dispensings, dispensing days supplied, all events, and episodes with events.

Episodes will be censored at the occurrence of an event during the overlap period. Variables bolded in the Variable column are stratifiers.

Table 56. [RUNID]_t2_concomitance.sas7bdat Output

Variable	Description
ANALYSISGRP	**Details:** GROUP name to differentiate primary/secondary pairs. **Format:** Character (45)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. **Format:** Character(3)
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(2)
AGEGROUP	Age Groups. Categories are requester-defined. **Format:** Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
YEAR	Year of index date. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric
RACE	Race. Allowable values are those in the SCDM. **Format:** Character(1)
HISPANIC	A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race. Allowable values are those in the SCDM. **Format:** Character(1)
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)

Variable	Description
ZIP_UNCERTAIN	Y / N variable. **Format:** Character(1)
COVAR*n*	Requester-defined covariate **Note 1:** N will equal COVARNUM specified in the Covariate file **Format**: Numeric
RAWCODECOUNT	Number of index defining codes **Format:** Numeric
ADJUSTEDCODECOUNT	Number of index defining codes adjusted for codes incurred on same date. **Format:** Numeric
DAYSUPP	Days supplied **Format:** Numeric
AMTSUPP	Amount supplied **Format**: Numeric
NPTS	Number of patients. **Format:** Numeric
EPISODES	Number of Episodes **Format:** Numeric
EPS_WEVENTS	Number of exposure episodes with an HOI. **Format:** Numeric
ALL_EVENTS	Total number of HOIs in all exposure episodes. **For characterization purposes only.** Days-at-risk stop accumulating after the first HOI during an exposure episode. ALL_EVENTS/TTE should never be calculated. This variable value just reports the number of times during treatment episodes that the HOI definition was met. **Format:** Numeric
TTE	Days at-risk. **Format:** Numeric

f) [RUNID]_concomitance_baseline_[PERIOD].sas7bdat

This table is structured to output one observation per ANALYSISGRP, evaluated at concomitant episode index date. All metrics within a ANALYSISGRP are calculated based on number of concomitant episodes and, depending on the specification of CONCCOHORTDEF values in the CONCFILE an individual may contribute more than one concomitant episode to each ANALYSISGRP. For each ANALYSISGRP, Table 57 includes the number of concomitant episodes *and* number of unique patients, to determine the extent of patient multiple-episode contribution.

Table 57. [RUNID]_concomitance_baseline_[PERIOD].sas7bdat

Variable	Description
ANALYSISGRP	Standardized name used to differentiate cohorts. **Format:** Character(30)
PATIENT	Number of unique patients. **Format:** Numeric(8)
N_EPISODES	Number of episodes. **Format:** Numeric(8)
AGE_XX	For each age group specified in run_programs.sas, the number of patient-episodes classified in that age category. **Format:** Numeric(8)
SEX_X	For each sex, the number of patient-episodes identified. **Format:** Numeric(8)
RACE_X	For each race, the number of patient-episodes identified. **Format:** Numeric(8)
HISPANIC_X	For each Hispanic value, the number of patient-episodes identified. **Format:** Numeric(8)
YEAR_X	For each year, the number of patient-episodes identified. **Format:** Numeric(8)
COVAR_X	For each covariate specified in the Covariate Codes File, the number of patient-episodes identified. **Format:** Numeric(8)
MEAN_AGE	Mean age. **Format:** Numeric(8)
STD_AGE	Standard deviation of age. **Format:** Numeric(8)
MEAN_COMORBIDSCORE	Mean comorbidity score (blank if Comorbidity Score File not specified). **Format**: Numeric(8)
STD_COMORBIDSCORE	Standard deviation of comorbidity score (blank if Comorbidity Score File not specified). **Format:** Numeric(8)

Variable	Description
MEAN_NUMAV	Mean number of AV visits (blank if Utilization File not specified). **Format**: Numeric(8)
MEAN_NUMOA	Mean number of OA visits (blank if Utilization File not specified). **Format**: Numeric(8)
MEAN_NUMIP	Mean number of IP visits (blank if Utilization File not specified). **Format**: Numeric(8)
MEAN_NUMIS	Mean number of IS visits (blank if Utilization File not specified). **Format**: Numeric(8)
MEAN_NUMED	Mean number of ED visits (blank if Utilization File not specified). **Format**: Numeric(8)
MEAN_NUMGENERIC	Mean number of unique generics dispensed (value is set to 0 if Utilization File not specified). **Format**: Numeric(8)
MEAN_NUMCLASS	Mean number of unique drug classes dispensed (value is set to 0 if Utilization File not specified). **Format**: Numeric(8)
MEAN_NUMRX	Mean number of outpatient pharmacy dispensings (value is set to 0 if Utilization File not specified). **Format**: Numeric(8)
STD_NUMAV	Standard deviation of AV visits (value is set to 0 if Utilization File not specified). **Format:** Numeric(8)
STD_NUMOA	Standard deviation of OA visits (value is set to 0 if Utilization File not specified). **Format:** Numeric(8)
STD _NUMIP	Standard deviation of IP visits (blank if Utilization File not specified). **Format:** Numeric(8)
STD _NUMIS	Standard deviation of IS visits (blank if Utilization File not specified). **Format:** Numeric(8)
STD _NUMED	Standard deviation of ED visits (blank if Utilization File not specified). **Format:** Numeric(8)
STD _NUMGENERIC	Standard deviation of unique generics dispensed (value is set to 0 if Utilization File not specified). **Format**: Numeric(8)
STD _NUMCLASS	Standard deviation of unique drug classes dispensed (value is set to 0 if Utilization File not specified). **Format**: Numeric(8)
STD _NUMRX	Standard deviation of outpatient pharmacy dispensings (value is set to 0 if Utilization File not specified). **Format**: Numeric(8)

g) [RUNID]_t2_multevent.sas7bdat

The output table includes the number of episodes, users, dispensings, dispensing days supplied, number of episodes with no secondary episodes, number of episodes with at least 1 secondary episode, number of episodes and users that meet adherence (overall adherence and by each criterion), eligible members and eligible member days.

Table 58. [RUNID]_t2_multevent.sas7bdat Output

Variable	Description
ANALYSISGRP	**Details:** GROUP name to differentiate primary/secondary pairs. **Format:** Character (30)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. **Format:** Character(3)
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(2)
AGEGROUP	Age Groups. Categories are requester-defined. **Format:** Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
YEAR	Year of index date. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric
RACE	Race. Allowable values are those in the SCDM. **Format:** Character(1)
HISPANIC	A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race. Allowable values are those in the SCDM. **Format:** Character(1)
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)

Variable	Description
ZIP_UNCERTAIN	Y / N variable. **Format:** Character(1)
TTE_CAT	Primary episode length categories **Format:** Numeric
EPI_COUNT	Number of secondary episodes during the observation window **Format:** Numeric
TIME_TO_EPI	Time to X secondary episode (As determined by EPISODENUM parameter) **Format:** Numeric
ADHERENCE	Meets adherence **Format:** Numeric
ADHERENCE_#	Meets adherence for each adherence criteria **Format:** Numeric
COVAR*n*	Requester-defined covariate **Note 1:** N will equal COVARNUM specified in the Covariate file **Format**: Numeric
NPTS	Number of patients. **Format:** Numeric
EPISODES	Number of Episodes **Format:** Numeric
RAWCODECOUNT	Number of index defining codes. **Format:** Numeric
ADJUSTEDCODECOUNT	Number of index defining codes adjusted for codes incurred on same date. **Format**: Numeric
DAYSUPP	Days supply associated with dispensings. **Format:** Numeric
TTE	Total duration for the primary episode **Format:** Numeric
EPS_WSECEPI	Number of primary episodes where epi_count ≥1 **Format:** Numeric
EPS_WOSECEPI	Number of primary episodes where epi_count = 0 **Format:** Numeric
ADHERE_EPISODES	Number of primary episodes that meet adherence **Format:** Numeric
ADHERE_NPTS	Number of patients that meet adherence **Format:** Numeric
DENNUMPTS	Number of patients eligible to have at least one index date. Only calculated for overall, age group, sex, and year*month stratified analysis.

Variable	Description
	Format: Numeric
DENNUMMEMDAYS	Number of days that patients are eligible to have an index date. Only calculated for overall, age group, sex, and year*month stratified analysis. **Format:** Numeric

h) [*RUNID*]_t2_epigap.sas7bdat

The episode gap table characterizes secondary episode gaps when using the multievent tool with a Type 2 analysis. There is one row per each secondary episode gap day and a count of secondary episodes with that gap. Gap is reported for overall, and stratified by age group, sex, year/month,

Table 59. [RUNID]_t2_epigap.sas7bdat

Variable	Description
ANALYSISGRP	**Details:** GROUP name to differentiate primary/secondary pairs. **Format:** Character (30)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.*, variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. **Format:** Character(4)
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(2)
AGEGROUP	Age Groups. Categories are requester-defined. **Format:** Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
YEAR	Year of index date. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric
EPI_GAP	Gap day **Format:** Numeric
COUNT1	Number of secondary episodes with a gap corresponding to that day (all secondary episodes) **Format:** Numeric
COUNT2	Number of secondary episodes with a gap between 1st and 2nd secondary episodes. **Format**: Numeric
COUNT3	Number of secondary episodes with a gap between 2nd and 3rd secondary episodes. **Format:** Numeric

i) [*RUNID*]_t2_overlap.sas7bdat

Table 60 characterizes overlap between primary and secondary episodes. It includes number of episodes, users, dispensings, dispensing days supplied, number of episode with no secondary episodes, number of episodes with at least 1 secondary episode, number of users with no secondary episodes, number of users with at least 1 secondary episodes, number of episodes that meet overlap thresholds, minimum and maximum days overlap, number of episodes and users that meet adherence, eligible members and eligible member days.

Table 60. [RUNID]_t2_overlap.sas7bdat Output

Variable	Description
ANALYSISGRP	**Details:** GROUP name to differentiate primary/secondary pairs. **Format:** Character (45)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. **Format:** Character(3)
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(2)
AGEGROUP	Age Groups. Categories are requester-defined. **Format:** Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
YEAR	Year of index date. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric
RACE	Race. Allowable values are those in the SCDM. **Format:** Character(1)
HISPANIC	A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race. Allowable values are those in the SCDM. **Format:** Character(1)
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)

Variable	Description
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. **Format:** Character(1)
COVAR*n*	Requester-defined covariate **Note 1:** N will equal COVARNUM specified in the Covariate file **Format**: Numeric
TTE_CAT	Primary episode length categories **Format:** Numeric
TOTAL_DAYS_OVERLAP	Number of days of overlap **Format:** Numeric
ADHERENCE	Meets adherence **Format:** Numeric
NPTS	Number of patients. **Format:** Numeric
EPISODES	Number of Episodes **Format:** Numeric
RawCodeCount1	Number of index defining codes for primary episode. **Format:** Numeric
RawCodeCount2	Number of index defining codes for secondary episode. **Format:** Numeric
AdjustedCodeCount1	Number of index defining codes for primary episode adjusted for codes incurred on the same date. **Format:** Numeric
AdjustedCodeCount2	Number of index defining codes for secondary episode adjusted for codes incurred on the same date. **Format:** Numeric
DAYSUPP1	Days supply associated with dispensing for primary episode. **Format:** Numeric
DAYSUPP2	Days supply associated with dispensing for secondary episode. **Format:** Numeric
EPS_WSecEpi	Number of primary episodes with at least one secondary episode **Format:** Numeric
EPS_WOSecEpi	Number of primary episodes with no secondary episode **Format:** Numeric
NPTS_WSecEpi	Number of users with at least one secondary episode **Format:** Numeric

Variable	Description
NPTS_WOSecEpi	Number of users with no secondary episode **Format:** Numeric
TTE	Total duration of primary episode **Format:** Numeric
TOTAL_OVERLAP	Total number of days overlap between primary and secondary episodes **Format:** Numeric
EPI_XX_XX (CUTOFFCAT)	Output for each CUTOFFCAT. Number of episodes that where CUTOFFCAT value = 1 **Format:** Numeric
NPTS_XX_XX (CUTOFFCAT)	Output for each CUTOFFCAT. Number of users where at least one episode had a CUTOFFCAT value = 1 **Format:** Numeric
ADHERE_EPISODES	Number of primary episodes that meet adherence **Format:** Numeric
ADHERE_NPTS	Number of patients that meet adherence **Format:** Numeric
DENNUMPTS	Number of patients eligible to have at least one index date. Only calculated for overall, age group, sex, and year/month stratified analysis. **Format:** Numeric
DENNUMMEMDAYS	Number of days that patients are eligible to have an index date. Only calculated for overall, age group, sex, and year/month stratified analysis. **Format:** Numeric

j) [*RUNID*]_T3_CIDA.sas7bdat

The [*RUNID]*_T3_CIDA output table includes the number of individuals and exposure episodes in the exposure and analytic cohorts, number of individuals and exposure episodes censored, and number of HOIs observed. All metrics are reported overall and stratified by age group, sex, year, year-month, and time-to-event (in days). Table 61 contains specifications for the [*RUNID]*_T3_CIDA output table.

Table 61. [RUNID]_T3_CIDA Output

Variable	Description
GROUP	Cohort name. **Format:** Character(30)
LEVEL	Stratification identifier. Each unique combination of strata (i.e., variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character(3)

Variable	Description
CENSOR	Y/N indicator for whether the stratification represents a cohort with incomplete follow-up. **Format:** Character(1)
SEX	Stratification variable. Sex. Allowable values are those in the SCDM. **Format:** Character(1)
RACE	Race. Allowable values are those in the SCDM. **Format:** Character(1)
HISPANIC	A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race. Allowable values are those in the SCDM. **Format:** Character(1)
AGEGROUP	Stratification variable. Age Groups. Categories are requester-defined. **Format:** Character(variable)
AGEGROUPNUM	Stratification variable. Numeric identifier of each AGEGROUP value. **Format:** Numeric
YEAR	Stratification variable. Year of index date. **Format:** Numeric
MONTH	Stratification variable. Month of index date. **Format:** Numeric
TTE_VALUE	Stratification variable. All available time to event values (e.g., -2 -1, 0, 1, 2, 3, etc.). Blank TTE_VALUE may be used to characterize patients in the exposure cohort only. If an HOI is observed on the day of exposure, TTE=0 (i.e., exposure date is day 0). **Format:** Numeric
TTC_VALUE	Stratification variable. All available time to censor values (e.g., 0, 1, 2, 3, etc.). Blank TTC_VALUE may be used to characterize patients in the exposure cohort only. If the last day of follow-up is the index date, TTC_VALUE=0. **Format:** Numeric
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date

Variable	Description
	N: zip date occurred before the index date **Format:** Character(1)
COVAR*n*	Requester-defined covariate **Note 1:** N will equal COVARNUM specified in the Covariate file **Format**: Numeric
NPTS_EXPOSURE	Number of patients identified in the exposure cohort. **Format:** Numeric
EPISODES_EXPOSURE	Number of index dates (exposure episodes) identified for all members in the exposure cohort. Relevant for requests that allow more than one exposure episode per patient. **Format:** Numeric
NPTS_CENSOR_ELIG	Number of patients excluded from the analysis cohort due to insufficient post-exposure continuous enrollment during the risk and control windows. If death and disenrollment occur on the same day, censoring will be attributed to death. **Format:** Numeric
NPTS_CENSOR_DTH	Number of patients excluded from the analysis cohort due to evidence of death during the risk or control windows. If death and disenrollment occur on the same day, censoring will be attributed to death. **Format:** Numeric
NPTS_CENSOR_NOEVENTS	Number of patients excluded from the analysis cohort due to no identified events during either the risk or control windows. **Format:** Numeric
EPISODES_CENSOR_ELIG	Number of exposure episodes excluded from the analysis cohort due to insufficient post-exposure continuous enrollment during the risk and control windows. If death and disenrollment occur on the same day, censoring will be attributed to death. **Format:** Numeric
EPISODES_CENSOR_DTH	Number of exposure episodes excluded from the analysis cohort due to evidence of death during the risk or control windows. If death and disenrollment occur on the same day, censoring will be attributed to death. **Format:** Numeric
EPISODES_CENSOR_NOEVENTS	Number of exposure episodes excluded from the analysis cohort due to no identified events during either the risk or control windows. **Format:** Numeric
NPTS_ANALYSIS	Number of patients identified in the analytic cohort. **Format:** Numeric

Variable	Description
EPISODES_ANALYSIS	Number of index dates (exposure episodes) identified for all members in the analytic cohort. Relevant for requests that allow more than one exposure episode per patient. **Note 1**: EPISODES_ANALYSIS=EVENTS_ANALYSIS_RISK + EVENTS_ANALYSIS_CTRL **Note 2:** EPISODES_ANALYSIS= EPISODES_EXPOSURE – EPISODES_CENSOR_ELIG– EPISODES_CENSOR_DTH – EPISODES_CENSOR_NOEVENTS **Format:** Numeric
EVENTS_ANALYSIS_RISK	Number of events identified in the risk window for patients in the analytic cohort. **Format:** Numeric
EVENTS_ANALYSIS_CTRL	Number of events identified in the control window for patients in the analytic cohort. **Format:** Numeric
MINDAYS_EVENT_ANALYSIS	The minimum number of days from exposure to event (event date – exposure date +1 for post-exposure events; exposure date- event date +1 for pre-exposure events). Relevant for members of the analytic cohort only. **Format:** Numeric
MAXDAYS_EVENT_ANALYSIS	The maximum number of days from exposure to event (event date – exposure date +1 for post-exposure events; exposure date- event date +1 for pre-exposure events). Relevant for members of the analytic cohort only. **Format:** Numeric
MINDAYS_POSTENR_EXPOSURE	The minimum number of days of post-exposure enrollment for the exposure cohort (enrollment end – exposure date). **Format:** Numeric
MAXDAYS_POSTENR_EXPOSURE	The maximum number of days of post-exposure enrollment for the exposure cohort (enrollment end – exposure date). **Format:** Numeric

k) [*RUNID*]_baseline_[T3OUT]_[T4OUT]_[*PERIODID*].sas7bdat

The [*RUNID]*_baseline_[*PERIODID*] output table includes metrics for cohorts of interest during a "baseline" period - a user-defined time period before the index date. The table includes information on comorbidities present, age group, sex, and medical and drug utilization metrics. This output file will be generated for Types 1-5 analyses.

This output table is structured to output one observation per GROUP. All metrics within a GROUP are calculated based on number of episodes and, depending on the specification of T#COHORTDEF values in the Type 1 File or Type 2 File or Type 3 File or Type 4 File or Type 5 File, an individual may contribute more than one episode to each GROUP. For each GROUP in Types 1-4 this output table includes the number of episodes *and* number of unique patients, to determine the extent of patient multiple-episode contribution. For Type 5, the baseline period is at the patient level (variable PATIENT from Table 62 below would equal variable N_EPISODES).

In addition to patients potentially contributing more than one episode within a GROUP, it is also possible that a single patient can contribute to multiple GROUPs. Table 62 contains specifications for the [*RUNID]* _baseline_[*PERIODID*] output table.

For Type 3 analyses, T3OUT takes the value of _an and _an_censor for analysis and analysis_censor datasets; T3OUT is blank for all non-Type 3 analyses.

For Type 4 analyses, T4OUT takes the value of Preg, MI, or NoPreg_for the pregnant cohort, the exposed and comparator/unexposed pregnant cohort, and the non-pregnant cohort, respectively. T4OUT is blank for all non-Type 4 analyses.

Table 62. [RUNID] _baseline_[T3OUT]_[T4OUT]_[PERIODID] Output

Variable	Description
GROUP	Standardized name used to differentiate cohorts. **Format:** Character(30)
PATIENT	Number of unique patients. **Format:** Numeric(8)
N_EPISODES	Number of episodes. **Format:** Numeric(8)
AGE_XX	For each age group specified in run_programs.sas, the number of patient-episodes classified in that age category. **Format:** Numeric(8)
SEX_X	For each sex, the number of patient-episodes identified. **Format:** Numeric(8)
RACE_X	For each race, the number of patient-episodes identified. **Format:** Numeric(8)
HISPANIC_X	For each Hispanic value, the number of patient-episodes identified. **Format:** Numeric(8)
YEAR_X	For each year, the number of patient-episodes identified. **Format:** Numeric(8)

Variable	Description
COVAR_X	For each covariate specified in the Covariate Codes File, the number of patient-episodes identified. **Format:** Numeric(8)
MEAN_AGE	Mean age. **Format:** Numeric(8)
STD_AGE	Standard deviation of age. **Format:** Numeric(8)
MEAN_COMORBIDSCORE	Mean comorbidity score (blank if Comorbidity Score File not specified). **Format**: Numeric(8)
STD_COMORBIDSCORE	Standard deviation of comorbidity score (blank if Comorbidity Score File not specified). **Format:** Numeric(8)
MEAN_NUMAV	Mean number of AV visits (blank if Utilization File not specified). **Format**: Numeric(8)
MEAN_NUMOA	Mean number of OA visits (blank if Utilization File not specified). **Format**: Numeric(8)
MEAN_NUMIP	Mean number of IP visits (blank if Utilization File not specified). **Format**: Numeric(8)
MEAN_NUMIS	Mean number of IS visits (blank if Utilization File not specified). **Format**: Numeric(8)
MEAN_NUMED	Mean number of ED visits (blank if Utilization File not specified). **Format**: Numeric(8)
MEAN_NUMGENERIC	Mean number of unique generics dispensed (value is set to 0 if Utilization File not specified). **Format**: Numeric(8)
MEAN_NUMCLASS	Mean number of unique drug classes dispensed (value is set to 0 if Utilization File not specified). **Format**: Numeric(8)
MEAN_NUMRX	Mean number of outpatient pharmacy dispensings (value is set to 0 if Utilization File not specified). **Format**: Numeric(8)
STD_NUMAV	Standard deviation of AV visits (value is set to 0 if Utilization File not specified). **Format:** Numeric(8)
STD_NUMOA	Standard deviation of OA visits (value is set to 0 if Utilization File not specified). **Format:** Numeric(8)
STD _NUMIP	Standard deviation of IP visits (blank if Utilization File not specified).

Variable	Description
	Format: Numeric(8)
STD _NUMIS	Standard deviation of IS visits (blank if Utilization File not specified). **Format:** Numeric(8)
STD _NUMED	Standard deviation of ED visits (blank if Utilization File not specified). **Format:** Numeric(8)
STD _NUMGENERIC	Standard deviation of unique generics dispensed (value is set to 0 if Utilization File not specified). **Format**: Numeric(8)
STD _NUMCLASS	Standard deviation of unique drug classes dispensed (value is set to 0 if Utilization File not specified). **Format**: Numeric(8)
STD _NUMRX	Standard deviation of outpatient pharmacy dispensings (value is set to 0 if Utilization File not specified). **Format**: Numeric(8)
PREPOSTIND_PRETERM	Number of pregnancy episode with gestational age defined as PRE (0-258 days duration) **Note:** Output for Type 4 analyses only. **Format**: Numeric(8)
PREPOSTIND_TERM	Number of pregnancy episode with gestational age defined as TERM (259-280 days duration) **Note:** Output for Type 4 analyses only. **Format**: Numeric(8)
PREPOSTIND_POSTTERM	Number of pregnancy episode with gestational age defined as POST (281-301 days duration) **Note:** Output for Type 4 analyses only. **Format**: Numeric(8)
PREPOSTIND_NONE	Number of pregnancy episode with gestational age defined as NONE (No PREGDUR codes) **Note:** Output for Type 4 analyses only. **Format**: Numeric(8)
MEAN_GA_BIRTH	Mean gestational age at delivery, in weeks **Note:** Output for Type 4 analyses only. **Format:** Numeric(8)
STD_GA_BIRTH	Standard deviation of gestational age at birth, in weeks **Note:** Output for Type 4 analyses only. **Format**: Numeric(8)
MEAN_GA_FIRST	Mean gestational age at first exposure, in weeks.

Variable	Description
	Note: Output for Type 4 analyses only. Only computed for exposure and comparator cohorts. **Format**: Numeric(8)
STD_GA_FIRST	Standard deviation of first exposure, in weeks. **Note:** Output for Type 4 analyses only. Only computed for exposure and comparator cohorts. **Format**: Numeric(8)
MEAN_DISP_PRE	Mean number of dispensings in pre-pregnancy period **Note:** Output for Type 4 analyses only. Only computed for exposure and comparator cohorts. **Format**: Numeric(8)
STD_DISP_PRE	Standard deviation of number of dispensings in pre-pregnancy period **Note:** Output for Type 4 analyses only. Only computed for exposure and comparator cohorts. **Format**: Numeric(8)
MEAN_AdjustedDisp_T1	Mean number of dispensings in first trimester **Note:** Output for Type 4 analyses only. Only computed for exposure and comparator cohorts. **Format**: Numeric(8))
STD_AdjustedDisp_T1	Standard deviation of number of dispensings in first trimester **Note:** Output for Type 4 analyses only. Only computed for exposure and comparator cohorts. **Format**: Numeric(8)
MEAN_AdjustedDISP_T2	Mean number of dispensings in second trimester **Note:** Output for Type 4 analyses only. Only computed for exposure and comparator cohorts. **Format**: Numeric(8)
STD_AdjustedDISP_T2	Standard deviation of number of dispensings in second trimester **Note:** Output for Type 4 analyses only. Only computed for exposure and comparator cohorts. **Format**: Numeric(8)
MEAN_AdjustedDISP_T3	Mean number of dispensings in third trimester **Note:** Output for Type 4 analyses only. Only computed for exposure and comparator cohorts. **Format**: Numeric(8)
STD_AdjustedDISP_T3	Standard deviation of number of dispensings in third trimester

Variable	Description
	Note: Output for Type 4 analyses only. Only computed for exposure and comparator cohorts. **Format**: Numeric(8)
EXP_T1	The number of pregnancy episodes with exposure episodes overlapping the first trimester. The first trimester is 0-90 days following pregnancy start. **Note:** Output for Type 4 analyses only. Only computed for exposure and comparator cohorts. **Format**: Numeric(8)
EXP_T2	The number of pregnancy episodes with exposure episodes overlapping the second trimester. The second trimester is 91-180 days following pregnancy start. **Note:** Output for Type 4 analyses only. Only computed for exposure and comparator cohorts. **Format**: Numeric(8)
EXP_T3	The number of pregnancy episodes with exposure episodes overlapping the third trimester. The third trimester is 180+ days following pregnancy start and until delivery or child birth date. **Note:** Output for Type 4 analyses only. Only computed for exposure and comparator cohorts. **Format**: Numeric(8)
EXP_PRE	The number of pregnancy episodes with exposure episodes within the pre-pregnancy period. **Note:** Output for Type 4 analyses only. Only computed for exposure and comparator cohorts. **Format**: Numeric(8)
MEAN_BIRTH_ENROLL	Mean time of enrollment after birth, in weeks. **Note:** Output for Type 4 analyses only. Only computed when pregnant cohort is matched to an infant. **Format**: Numeric(8)
STD_BIRTH_ENROLL	Standard deviation of time of enrollment after birth, in weeks. **Note:** Output for Type 4 analyses only. Only computed when pregnant cohort is matched to an infant. **Format**: Numeric(8)
MEAN_ENROLL_DIFF	Mean difference between the date of birth and the date of enrollment of infant, in weeks. **Note:** Output for Type 4 analyses only. Only computed when pregnant cohort is matched to an infant. **Format**: Numeric(8)

Variable	Description
STD_ENROLL_DIFF	Standard deviation of the difference between the date of birth and the date of enrollment of infant, in weeks. **Note:** Output for Type 4 analyses only. Only computed when pregnant cohort is matched to an infant. **Format**: Numeric(8)

l) [*RUNID*]_signature.sas7bdat

The [*RUNID*] _signature output table contains metadata associated with the request, including request identifiers, program identifiers, database version, and run time metrics. Table 61Table 63 contains specifications for the [*RUNID*] _signature output table.

Table 63. [RUNID] _signature Output

Variable	Description
VAR	Metric name. **Format:** Character(15)
VALUE	Metric value. **Format:** Character(200)

m) [*RUNID*]_attrition.sas7bdat

The [*RUNID*] _attrition output table includes the number individuals excluded and remaining at each cohort creation criterion application during the CIDA tool execution. Table 64 contains specifications for the [*RUNID*] _attrition output table. For requests that will use the prospective surveillance with propensity score matched design, attrition table values for levels 1-19 will be based on the current look's ETL. This table will be automatically output for analysis types 1, 2,3, 5, and 6.

Table 64. [RUNID] _attrition Output

Variable	Description
GROUP	Cohort name. **Format:** Character(30)
LEVEL	Criterion identifier. **Format:** Numeric
DESCR	Criterion description. **Format:** Character(500)
REMAINING	Number of individuals remaining after previous exclusion criterion. **Format:** Numeric
EXCLUDED	Number of individuals excluded by the exclusion criterion. **Format:** Numeric

n) [RUNID]_MFU_[OUTCOHORT]_[T3OUT].sas7bdat

This output table includes most frequent utilization assessment, which enables requesters retrive the top XX NDCs, diagnoses, and procedures before and/or after the index date for any Type of analysis in QRP (1-5).

To evaluate the TOP XX codes, the program will look during a requester-defined period relative to index date and count the number of instances of all codes for requested code categories (CODECAT). The top XX codes across all CODECAT requested will be included in the output metrics (e.g., if the Top 50 ICD-9-CM diagnosis, ICD-9-CM procedure, and 9-digit NDCs are requested, the top 50 codes will include a mix of these code types and be selected based on frequency of occurrence).

Codes evaluated will not be pre-processed; meaning, codes evaluated for MFU analyses will not be stockpiled or processed by the envelope macro. Output metrics will report what is observed in the Sentinel Distributed Database (SDD). See Table 65 below contains specifications for this output.

Table 65. [RUNID]_MFU_[outcohort]_[t3out]

Parameter	Field Name	Description
Cohort Name	GROUP	Unique cohort name.
Code Category	CODECAT	CODECAT value associated with identified CODE.
Code Type	CODETYPE	CODETYPE value associated with identified CODE.
Code	CODE	CODE (with no decimal places)
Code Count	CODECOUNT	Count of codes observed during window.
Patient Count	PATCOUNT	Count of patients with identified code during window.

o) [*RUNID*]_metadata_for_time_period_#.sas7bdat

The [*RUNID*] _metadata_for_time_period_# output table is generated for Type 3 requests only. It includes request-specific metadata necessary for the conduct of surveillance activities. For each GROUP value, the output table includes the surveillance start date, exposure assessment period start and end dates, and the requester-defined data completeness date. This information can be used by subsequent executions of the program to ensure mutually exclusive but contiguous exposure assessment periods (e.g., if Look 1 has an exposure assessment period end date of 11/30/2014, Look 2 will need to have an exposure assessment period start date of 12/1/2014).

For example, executing the program package for Look 1 will generate an output file [*RUNID*] _metadata_for_time_period_1 to the *msoc* folder. When the Look 2 package is distribute to partners for execution, the [*RUNID*] _metadata_for_time_period_1 will be included in the *inputfiles* folder for reference by the program. After execution, a new output table ([*RUNID*] _metadata_for_time_period_2) will be output to the *msoc* folder for use in the execution of the Look 3 package. Table 66 contains specifications for the [*RUNID*] _metadata_for_time_period_# output table.

Table 66. [RUNID] _metadata_for_time_period_# Output

Variable	Description
MSREQID	Unique request identifier. A concatenation of MSPROJID, MSWPTYPE, MSWPID, MSDPID, MSVERID. **Format:** Character(variable)
RUNID	Unique run (execution) identifier. **Format:** Character(3)
GROUP	Cohort name. **Format:** Character(30)
TIMEPERIOD	PERIODIDSTART and PERIODIDEND value for the Type 3 analysis. **Format**: Numeric
T3SURVSTARTDATE	Surveillance start date. **Format**: Numeric (date9.)
EXPPERIODSTARTDT	Exposure assessment period start date. For PERIODIDSTART = 1, date will be T3SURVSTARTDATE. For PERIODIDSTART ≥1, the value will equal the EXPPERIODENDDT from the previous execution of the program package + 1. **Format**: Numeric (date9.)
EXPPERIODENDDT	Exposure assessment period end date. This date is determined based on the follow-up required by a particular request. **Format**: Numeric (date9.)
T3ENDOFUPDATE	Requester-defined Data Partner data completeness date. **Format**: Numeric (date9.)

p) [*RUNID*]_t4_CIDA_elig.sas7bdat

The [*RUNID*] _t4_cida_elig output table is generated for Type 4 requests only. Table 67Table 67 contains specifications for the [*RUNID*] _t4_cida_elig output table.

Table 67. [RUNID]_ t4_cida_elig Output

Variable	Description
GROUP	Cohort name. **Format**: Character(30)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character(3)
ELIGIBLE	Y/N . Y: patient meets eligibility requirements (i.e., patient is included in the master file)

Variable	Description
	N: patient does not meet eligibility requirements (i.e., patient is included in alldeliveries, but not the master file) Note: when stratification variable is not included in a LEVEL, values of ELIGIBLE will be blank. **Format:** Character(1)
ELIGDAYS	Number of days of continuous enrollment before delivery date. **Format:** Numeric
PREPOSTIND	Categorical variable, with valid values PRE, POST, TERM, NONE Defined as: • PRE (0-258 days duration) • TERM (259-280 days duration) • POST (281-301 days duration) • NONE (No PregDur codes) **Format:** Character (4)
NPTS	Number of patients. **Format:** Numeric
EPISODES	Number of exposure episodes. **Format:** Numeric

q) [*RUNID*]_t4_CIDA_Preg.sas7bdat

The [*RUNID*] _t4_cida_Preg output table is generated for Type 4 requests only. This output captures pregnancy episodes ending in a live birth. Table 68 contains specifications for the [*RUNID*] _t4_cida_Preg output table.

Table 68. [RUNID]_t4_cida_Preg Output

Variable	Description
GROUP	Cohort name. **Format:** Character(30)
MOINAME	Value of T4_index MP*n*, to denote the medical exposure of interest. Missing MOINAME represents all episodes. **Format:** Character(10)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Note 1:** for level 606, counts of Preterm and Postterm codes reflect the code associated with the delivery. Using Priority variable in pregdur, only the code associated with the delivery is counted in level 606. **Format:** Character(3)

Variable	Description
RACE	Race. Allowable values are those in the SCDM. **Format:** Character(1)
HISPANIC	A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race. Allowable values are those in the SCDM. **Format:** Character(1)
AGEGROUP	Age Groups. Categories are requester-defined. Age is maternal age at delivery. **Format:** Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
YEAR	Year of delivery date. **Format:** Numeric
PREPOSTIND	Categorical variable, with valid values PRE, POST, TERM, NONE Defined as: • PRE (0-258 days duration) • TERM (259-280 days duration) • POST (281-301 days duration) • NONE (No PregDur codes) **Format:** Character(4)
PREGDURCODE	Code used to define pregnancy duration. **Format:** Character(11)
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date N: zip date occurred before the index date **Format:** Character(1)
COVAR*n*	Requester-defined covariate **Note 1:** N will equal COVARNUM specified in the Covariate file **Format**: Numeric
NPTS	Number of patients.

Variable	Description
	Format: Numeric
EPISODES	Number of exposure episodes. **Format:** Numeric
EPISODES_3TRIM	Number of exposure episodes with 3 trimesters. **Format:** Numeric
USEPRE	Pregnancy episodes with product use requester-specified # of days before **Format:** Numeric
ANYT1	Pregnancy episodes with product use during the 1st trimester. The 1st trimester is 0-90 days following pregnancy start. **Format:** Numeric
ANYT2	Pregnancy episodes with product use during the 2nd trimester. The 2nd trimester is 91-180 days following pregnancy start. **Format:** Numeric
ANYT3	Pregnancy episodes with product use during the 3rd trimester. The 3rd trimester is 180+ days following pregnancy start and until delivery. **Format:** Numeric
ANYT	Pregnancy episodes with product use during any trimester **Format:** Numeric
ONLYT1	Pregnancy episodes with product use only during the 1st trimester. The 1st trimester is 0-90 days following pregnancy start. **Format:** Numeric
ONLYT2	Pregnancy episodes with product use only during the 2nd trimester. The 2nd trimester is 91-180 days following pregnancy start. **Format:** Numeric
ONLYT3	Pregnancy episodes with product use only during the 3rd trimester. The 3rd trimester is 180+ days following pregnancy start and until delivery. **Format:** Numeric
ALLT	Pregnancy episodes with product use in all three trimesters **Format:** Numeric
SUMUSEPRE	Number of medical product use episodes in requester-specified # of days before pregnancy **Format:** Numeric
SUMRAWCNTPRE	Raw counts for medical product use episode codes during the PREPREGDAYS prior to pregnancy start date. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). “Raw” counts allow multiple codes on the same day. **Format:** Numeric

Variable	Description
SUMADJCNTPRE	Adjusted counts for medical product use episode codes during the PREPREGDAYS prior to pregnancy start date. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval).. "Adjusted" counts allow only one code per day. **Format:** Numeric
SUMANYT1	Number of medical product use episodes in the 1st trimester. The 1st trimester is 0-90 days following pregnancy start. **Format:** Numeric
SUMRAWCNTANYT1	Raw counts for medical product use episode codes during the 1st trimester. The 1st trimester is 0-90 days following pregnancy start. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Raw" counts allow multiple codes on the same day. **Format:** Numeric
SUMADJCNTANYT1	Adjusted counts for medical product use episode codes during the 1st trimester. The 1st trimester is 0-90 days following pregnancy start. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Adjusted" counts allow only one code per day. **Format:** Numeric
SUMANYT2	Number of medical product use episodes in the 2nd trimester. The 2nd trimester is 91-180 days following pregnancy start. **Format**: Numeric
SUMRAWCNTANYT2	Raw counts for medical product use episode codes during the 2nd trimester. The 2nd trimester is 91-180 days following pregnancy start. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval).. "Raw" counts allow multiple codes on the same day. **Format:** Numeric
SUMADJCNTANYT2	Adjusted counts for medical product use episode codes during the 2nd trimester. The 2nd trimester is 91-180 days following pregnancy start. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval).. "Adjusted" counts allow only one code per day.

Variable	Description
	Format: Numeric
SUMANYT3	Number of medical product use episodes in the 3rd trimester. The 3rd trimester is 180+ days following pregnancy start and until delivery. **Format:** Numeric
SUMRAWCNTANYT3	Raw counts for medical product use episode codes during the 3rd trimester. The 3rd trimester is 180+ days following pregnancy start. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Raw" counts allow multiple codes on the same day. **Format:** Numeric
SUMADJCNTANYT3	Adjusted counts for medical product use episode codes during the 3rd trimester. The 3rd trimester is 180+ days following pregnancy start. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Adjusted" counts allow only one code per day. **Format:** Numeric
SUMANYT	Number of medical product use episodes in any trimester **Format:** Numeric
SUMRAWCNTANYT	Raw counts for medical product use episode codes during any trimester. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Raw" counts allow multiple codes on the same day. **Format:** Numeric
SUMADJCNTANYT	Adjusted counts for medical product use episode codes during any trimester. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Adjusted" counts allow only one code per day. **Format:** Numeric
SUMALLT	Number of medical product use episodes in all three trimesters **Format:** Numeric
SUMRAWCNTALLT	Raw counts for medical product use episode codes that span all three trimesters. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters

Variable	Description
	based on the RxDate to RxDate + RxSup interval).. "Raw" counts allow multiple codes on the same day. **Format:** Numeric
SUMADJCNTALLT	Adjusted counts for medical product use episode codes that span all three trimesters. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval).. "Adjusted" counts allow only one code per day. **Format:** Numeric
SUMONLYT1	Number of medical product episodes with use only during the 1st trimester. The 1st trimester is 0-90 days following pregnancy start. **Format:** Numeric
SUMRAWCNTONLYT1	Raw counts for medical product use episode codes during the 1st trimester. The 1st trimester is 0-90 days following pregnancy start. Counts codes for medical product use episodes that occur in the 1st trimester only. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Raw" counts allow multiple codes on the same day. **Format:** Numeric
SUMADJCNTONLYT1	Adjusted counts for medical product use episode codes during the 1st trimester. The 1st trimester is 0-90 days following pregnancy start. Counts codes for medical product use episodes that occur in the 1st trimester only. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Adjusted" counts allow only one code per day. **Format:** Numeric
SUMONLYT2	Number of medical product episodes with use only during the 2nd trimester. The 2nd trimester is 91-180 days following pregnancy start. **Format:** Numeric
SUMRAWCNTONLYT2	Raw counts for medical product use episode codes during the 2nd trimester. The 2nd trimester is 91-180 days following pregnancy start. Counts codes for medical product use episodes that occur in the 2nd trimester only. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters

Variable	Description
	based on the RxDate to RxDate + RxSup interval). "Raw" counts allow multiple codes on the same day. **Format:** Numeric
SUMADJCNTONLYT2	Adjusted counts for medical product use episode codes during the 2nd trimester. The 2nd trimester is 91-180 days following pregnancy start. Counts codes for medical product use episodes that occur in the 2nd trimester only. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Adjusted" counts allow only one code per day. **Format:** Numeric
SUMONLYT3	Number of medical product episodes with use only during the 3rd trimester. The 3rd trimester is 180+ days following pregnancy start and until delivery. **Format:** Numeric
SUMRAWCNTONLYT3	Raw counts for medical product use episode codes during the 3rd trimester. The 3rd trimester is 180+ days following pregnancy start. Counts codes for medical product use episodes that occur in the 3rd trimester only. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Raw" counts allow multiple codes on the same day. **Format:** Numeric
SUMADJCNTONLYT3	Adjusted counts for medical product use episode codes during the 3rd trimester. The 3rd trimester is 180+ days following pregnancy start. Counts codes for medical product use episodes that occur in the 3rd trimester only. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Adjusted" counts allow only one code per day. **Format:** Numeric

r) [RUNID]_t4_CIDA_NoPreg.sas7bdat

The [*RUNID*]_t4_cidaNoPreg output table is generated for Type 4 requests only. This output captures a matched non-pregnant cohort (women were known to be non-pregnant or had a pregnancy not ending in a live birth). Table 69 contains specifications for the [*RUNID*]_t4_cida_NoPreg output table.

Table 69. [RUNID]_t4_cida_NoPreg Output

Variable	Description
GROUP	Cohort name. **Format:** Character(30)
MOINAME	Value of T4_index MP*n*, to denote the medical exposure of interest. Missing MOINAME represents all episodes. **Format:** Character(10)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character(3)
RACE	Race. Allowable values are those in the SCDM. **Format:** Character(1)
HISPANIC	A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race. Allowable values are those in the SCDM. **Format:** Character(1)
AGEGROUP	Age Groups. Categories are requester-defined. Age is maternal age at delivery. **Format:** Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
YEAR	Year of delivery date. **Format:** Numeric
PREPOSTIND	Categorical variable, with valid values PRE, POST, TERM, NONE Defined as: • PRE (0-258 days duration) • TERM (259-280 days duration) • POST (281-301 days duration) • NONE (No PregDur codes) **Format:** Character(4)
PREGDURCODE	Code used to define pregnancy duration. **Format:** Character(11)

Variable	Description
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date N: zip date occurred before the index date **Format:** Character(1)
COVAR*n*	Requester-defined covariate **Note 1:** N will equal COVARNUM specified in the Covariate file **Format**: Numeric
NPTS	Number of patients. **Format:** Numeric
EPISODES	Number of exposure episodes. **Format:** Numeric
USEPRE	Pregnancy episodes with product use requester-specified # of days before **Format:** Numeric
ANYT1	Pregnancy episodes with product use during the 1st trimester. The 1st trimester is 0-90 days following pregnancy start. **Format:** Numeric
ANYT2	Pregnancy episodes with product use during the 2nd trimester. The 2nd trimester is 91-180 days following pregnancy start. **Format:** Numeric
ANYT3	Pregnancy episodes with product use during the 3rd trimester. The 3rd trimester is 180+ days following pregnancy start and until delivery. **Format:** Numeric
ANY	Pregnancy episodes with product use during any trimester **Format:** Numeric
ONLYT1	Pregnancy episodes with product use only during the 1st trimester. The 1st trimester is 0-90 days following pregnancy start. **Format:** Numeric
ONLYT2	Pregnancy episodes with product use only during the 2nd trimester. The 2nd trimester is 91-180 days following pregnancy start. **Format:** Numeric

Variable	Description
ONLYT3	Pregnancy episodes with product use only during the 3rd trimester. The 3rd trimester is 180+ days following pregnancy start and until delivery. **Format:** Numeric
ALLT	Pregnancy episodes with product use in all three trimesters **Format:** Numeric
SUMUSEPRE	Number of medical product use episodes in 90-days before pregnancy **Format:** Numeric
SUMRAWCNTPRE	Raw counts for medical product use episode codes during the PREPREGDAYS prior to pregnancy start date. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Raw" counts allow multiple codes on the same day. **Format:** Numeric
SUMADJCNTPRE	Adjusted counts for medical product use episode codes during the PREPREGDAYS prior to pregnancy start date. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval).. "Adjusted" counts allow only one code per day. **Format:** Numeric
SUMANYT1	Number of medical product use episodes in the 1st trimester. The 1st trimester is 0-90 days following pregnancy start. **Format:** Numeric
SUMRAWCNTANYT1	Raw counts for medical product use episode codes during the 1st trimester. The 1st trimester is 0-90 days following pregnancy start. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Raw" counts allow multiple codes on the same day. **Format:** Numeric
SUMADJCNTANYT1	Adjusted counts for medical product use episode codes during the 1st trimester. The 1st trimester is 0-90 days following pregnancy start. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Adjusted" counts allow only one code per day. **Format:** Numeric
SUMANYT2	Number of medical product use episodes in the 2nd trimester. The 2nd trimester is 91-180 days following pregnancy start.

Variable	Description
	Format: Numeric
SUMRAWCNTANYT2	Raw counts for medical product use episode codes during the 2nd trimester. The 2nd trimester is 91-180 days following pregnancy start. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval).. "Raw" counts allow multiple codes on the same day. **Format:** Numeric
SUMADJCNTANYT2	Adjusted counts for medical product use episode codes during the 2nd trimester. The 2nd trimester is 91-180 days following pregnancy start. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval).. "Adjusted" counts allow only one code per day. **Format:** Numeric
SUMANYT3	Number of medical product use episodes in the 3rd trimester. The 3rd trimester is 180+ days following pregnancy start and until delivery. **Format:** Numeric
SUMRAWCNTANYT3	Raw counts for medical product use episode codes during the 3rd trimester. The 3rd trimester is 180+ days following pregnancy start. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Raw" counts allow multiple codes on the same day. **Format:** Numeric
SUMADJCNTANYT3	Adjusted counts for medical product use episode codes during the 3rd trimester. The 3rd trimester is 180+ days following pregnancy start. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Adjusted" counts allow only one code per day. **Format:** Numeric
SUMANY	Number of medical product use episodes in any trimester **Format:** Numeric
SUMRAWCNTANY	Raw counts for medical product use episode codes during any trimester. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Raw" counts allow multiple codes on the same day. **Format:** Numeric

Variable	Description
SUMADJCNTANY	Adjusted counts for medical product use episode codes during any trimester. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Adjusted" counts allow only one code per day. **Format:** Numeric
SUMALLT	Number of medical product use episodes in all three trimesters **Format:** Numeric
SUMRAWCNTALLT	Raw counts for medical product use episode codes that span all three trimesters. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval).. "Raw" counts allow multiple codes on the same day. **Format:** Numeric
SUMADJCNTALLT	Adjusted counts for medical product use episode codes that span all three trimesters. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval).. "Adjusted" counts allow only one code per day. **Format:** Numeric
SUMONLYT1	Number of medical product episodes with use only during the 1st trimester. The 1st trimester is 0-90 days following pregnancy start. **Format:** Numeric
SUMRAWCNTONLYT1	Raw counts for medical product use episode codes during the 1st trimester. The 1st trimester is 0-90 days following pregnancy start. Counts codes for medical product use episodes that occur in the 1st trimester only. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Raw" counts allow multiple codes on the same day. **Format:** Numeric
SUMADJCNTONLYT1	Adjusted counts for medical product use episode codes during the 1st trimester. The 1st trimester is 0-90 days following pregnancy start. Counts codes for medical product use episodes that occur in the 1st trimester only. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the

Variable	Description
	RxDate to RxDate + RxSup interval). "Adjusted" counts allow only one code per day. **Format:** Numeric
SUMONLYT2	Number of medical product episodes with use only during the 2nd trimester. The 2nd trimester is 91-180 days following pregnancy start. **Format:** Numeric
SUMRAWCNTONLYT2	Raw counts for medical product use episode codes during the 2nd trimester. The 2nd trimester is 91-180 days following pregnancy start. Counts codes for medical product use episodes that occur in the 2nd trimester only. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Raw" counts allow multiple codes on the same day. **Format:** Numeric
SUMADJCNTONLYT2	Adjusted counts for medical product use episode codes during the 2nd trimester. The 2nd trimester is 91-180 days following pregnancy start. Counts codes for medical product use episodes that occur in the 2nd trimester only. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Adjusted" counts allow only one code per day. **Format:** Numeric
SUMONLYT3	Number of medical product episodes with use only during the 3rd trimester. The 3rd trimester is 180+ days following pregnancy start and until delivery. **Format:** Numeric
SUMRAWCNTONLYT3	Raw counts for medical product use episode codes during the 3rd trimester. The 3rd trimester is 180+ days following pregnancy start. Counts codes for medical product use episodes that occur in the 3rd trimester only. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Raw" counts allow multiple codes on the same day. **Format:** Numeric
SUMADJCNTONLYT3	Adjusted counts for medical product use episode codes during the 3rd trimester. The 3rd trimester is 180+ days following pregnancy start. Counts codes for medical product use episodes that occur in the 3rd trimester only. Counts the total number of codes observed during a medical product use episode and attributes them to the trimester(s) that the code overlaps (i.e., if

Variable	Description
	the code is an NDC, a single code may overlap two trimesters based on the RxDate to RxDate + RxSup interval). "Adjusted" counts allow only one code per day. **Format:** Numeric

s) [RUNID]_t4_CIDA_Preg_gestwk.sas7bdat

Table 70. [RUNID]_ t4_cida_Preg_gestwk.sas7bdat

Variable	Description
GROUP	Cohort name. **Format:** Character(30)
MOINAME	Value of T4_index MP*n*, to denote the medical exposure of interest. **Format:** Character(10)
LEVEL	Stratification identifier. Each unique combination of strata receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. **Format:** Character(3)
GESTWK	Gestational week (relative to calculated pregnancy start). Determined as follows: • -2: days -14 to -8 • -1: days -7 to -1 • 1: days 0 to 6 • 2: days 7 to 13 There is no Week 0. All potential GESTWK values should be assigned by counting back or forward by 7 to encompass the entire timeframe under evaluation. **Format:** Numeric(8)
MOIEPISODES	Number of MOI episodes with an index date (i.e., episode start date) occurring during the specified gestational week. **Format:** Numeric(8)
MOIEPISODES_OVERLAP	Number of MOI episodes that overlap the gestational week, considering the length of the MOI episode.
MOIPATIENTS	Number of patients with an MOI episode index date (i.e., episode start date) occurring during the specified gestational week. **Format:** Numeric(8)
MOIPATIENT_OVERLAP	Number of patients with an MOI that overlaps the gestational week
RAWCODECOUNT	"Raw" MOI code count. Counts the total number of codes observed during an MOI episode and attributes them to the gestational week of the index date (i.e.,

Variable	Description
	episode start date). "Raw" counts allow multiple codes on the same day. **Format:** Numeric(8)
RAWCODECOUNT_OVERLAP	Number of raw codes that overlap the gestational week considering the day supply
ADJCODECOUNT	"Adjusted" MOI code count. Counts the total number of codes observed during an MOI episode and attributes them to the gestational week of the index date (i.e., episode start date). "Adjusted" counts allow only one code per day. **Format:** Numeric(8)
ADJUSTEDCODECOUNT_OVERLAP	Number of adjusted codes that overlap the gestational week considering the day supply
PREGEPISODES	Total number of pregnancy episodes in the gestational week (i.e., pregnancy episodes that lasted into the specified gestational week). **Format:** Numeric(8)
PREGPATIENTS	Total number of pregnant women in the gestational week (i.e., pregnant women whose pregnancies lasted into the specified gestational week). **Format:** Numeric(8)

t) *[RUNID]*_t4_CIDA_NoPreg_gestwk.sas7bdat

Table 71. *[RUNID]*_t4_cida_NoPreg_gestwk.sas7bdat Output

Variable	Description
GROUP	Cohort name. **Format:** Character(30)
MOINAME	Value of T4_index MP*n*, to denote the medical exposure of interest. **Format:** Character(10)
LEVEL	Stratification identifier. Each unique combination of strata receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. **Format:** Character(3)
GESTWK	Gestational week (relative to calculated pregnancy start). Determined as follows: • -2: days -14 to -8 • -1: days -7 to -1 • 1: days 0 to 6 • 2: days 7 to 13

Variable	Description
	There is no Week 0. All potential GESTWK values should be assigned by counting back or forward by 7 to encompass the entire timeframe under evaluation. **Format:** Numeric(8)
MOIEPISODES	Number of MOI episodes with an index date (i.e., episode start date) occurring during the specified gestational week. **Format:** Numeric(8)
MOIEPISODES_OVERLAP	Number of MOI episodes that overlap the gestational week, considering the length of the MOI episode.
MOIPATIENTS	Number of patients with an MOI episode index date (i.e., episode start date) occurring during the specified gestational week. **Format:** Numeric(8)
MOIPATIENT_OVERLAP	Number of patients with an MOI that overlaps the gestational week
RAWCODECOUNT	"Raw" MOI code count. Counts the total number of codes observed during an MOI episode and attributes them to the gestational week of the index date (i.e., episode start date). "Raw" counts allow multiple codes on the same day. **Format:** Numeric(8)
RAWCODECOUNT_OVERLAP	Number of raw codes that overlap the gestational week considering the day supply
ADJCODECOUNT	"Adjusted" MOI code count. Counts the total number of codes observed during an MOI episode and attributes them to the gestational week of the index date (i.e., episode start date). "Adjusted" counts allow only one code per day. **Format:** Numeric(8)
ADJUSTEDCODECOUNT_OVERLAP	Number of adjusted codes that overlap the gestational week considering the day supply
PREGEPISODES	Total number of pregnancy episodes in the gestational week (i.e., pregnancy episodes that lasted into the specified gestational week). **Format:** Numeric(8)
PREGPATIENTS	Total number of pregnant women in the gestational week (i.e., pregnant women whose pregnancies lasted into the specified gestational week). **Format:** Numeric(8)

u) *[RUNID]*_t5_CIDA_disp_by_daysupp.sas7bdat

This output table includes the number of adjusted dispensings by days supply, sex, race, hispanic, age group and geographic stratifications (optionally). Note that same-day dispensings should be handled according to specifications in the stockpiling input file. This means that:

If Two dispensings on the same day:

- Dispensing 1: RxSup = 30, RxAmount = 30
- Dispensing 2: RxSup = 60, RxAmount = 60

With default stockpiling algorithm (aa), output in [RUNID]_T5_CIDA_disp_by_daysupp will count this as 1 dispensing with RxSup = 90, RxAmt = 90.

Table 72Table 72 contains specifications for the [RUNID]_cida_disp_by_daysupp output table.

Table 72. [RUNID]_t5_cida_disp_by_daysupp Output

Variable	Description
GROUP	Cohort name. **Format:** Character(30)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character(3)
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(2)
AGEGROUP	Age Groups. Categories are requester-defined. Age is calculated at index date. **Format:** Character(variable)
RACE	Race. Allowable values are those in the SCDM. **Format:** Character(2)
HISPANIC	Hispanic. Allowable values are those in the SCDM. **Format:** Character(2)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions.

Variable	Description
	Format: Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date N: zip date occurred before the index date **Format:** Character(1)
DAYSUPP	Number of days of supply for dispensing (0+). **Format:** Numeric
COVAR*n*	Requester-defined covariate **Note 1:** N will equal COVARNUM specified in the Covariate file **Format**: Numeric
ADJUSTEDCODECOUNT	Number of dispensings used to define episode defining records. **Note 1:** When multiple dispensings occur on the same day with same STOCKGROUP, this will be counted as 1 dispensing, associated with the total day's supply. **Format:** Numeric

v) [*RUNID*]_t5_CIDA_firsteps.sas7bdat

This output table includes the number of patients, episodes, dispensings, and total days of supply by sex, race, hispanic, age group, and month of study start for the first patient episode (i.e., the index date defining episode) during the query period. Table 73 contains specifications for the [*RUNID*] _t5_cida_firsteps output table.

Table 73. [*RUNID*]_t5_cida_firsteps Output

Variable	Description
GROUP	Cohort name. **Format:** Character(30)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character(3)
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(2)
AGEGROUP	Age Groups. Categories are requester-defined. Age is calculated at index date. **Format:** Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
RACE	Race. Allowable values are those in the SCDM. **Format:** Character(2)
HISPANIC	Hispanic. Allowable values are those in the SCDM. **Format:** Character(2)
MNTSFROMSTART	Number of months from episode start date. Since this output table only includes the first episode, episode start date = index date. MNTSFROMSTART = 1 indicates the same month as the episode start date. **Format:** Numeric
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date

Variable	Description
	N: zip date occurred before the index date **Format:** Character(1)
COVAR*n*	Requester-defined covariate **Note 1:** N will equal COVARNUM specified in the Covariate file **Format**: Numeric
DAYSUPP	Total days of supply. **Format:** Numeric
EPISODES	Number of episodes initiatied. **Format:** Numeric
NPTS	Number of patients. **Format:** Numeric
ADJUSTEDCODECOUNT	Number of dispensings used to define index date defining records. **Note 1:** When multiple dispensings occur on the same day with same STOCKGROUP, this will be counted as 1 dispensing. This count will equal the count of the DISPENSINGS metric in prior QRP versions. **Format:** Numeric
RAWCODECOUNT	Number of dispensings used to define index date defining records. **Note 1:** When multiple dispensings occur on the same day with same STOCKGROUP, each dispensing will be counted. **Format:** Numeric

w) [*RUNID*]_t5_CIDA_alleps.sas7bdat

This output table includes the number of patients, episodes, dispensings, and total days of supply by sex, race, hispanic, age group, and month of study start for all observed episodes during the query period. Table 74 contains specifications for the [*RUNID*] _t5_cida_firsteps output table.

Table 74. [*RUNID*]_t5_cida_alleps Output

Variable	Description
GROUP	Cohort name. **Format:** Character(30)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character(3)
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(2)
AGEGROUP	Age Groups. Categories are requester-defined. Age is calculated at index date. **Format:** Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
RACE	Race. Allowable values are those in the SCDM. **Format:** Character(2)
HISPANIC	Hispanic. Allowable values are those in the SCDM. **Format:** Character(2)
MNTSFROMSTART	Number of months from episode start date. Since this output table includes all episodes, episode start date will not always equal the index date. MNTSFROMSTART = 1 indicates the same month as the episode start date. **Format:** Numeric
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date N: zip date occurred before the index date

Variable	Description
	Format: Character(1)
COVAR*n*	Requester-defined covariate **Note 1:** N will equal COVARNUM specified in the Covariate file Format: Numeric
DAYSUPP	Total days of supply. **Format:** Numeric
EPISODES	Number of episodes initiatied. **Format:** Numeric
NPTS	Number of patients. **Format:** Numeric
ADJUSTEDCODECOUNT	Number of dispensings used to define index date defining records. **Note 1:** When multiple dispensings occur on the same day with same STOCKGROUP, this will be counted as 1 dispensing. This count will equal the count of the DISPENSINGS metric in prior QRP versions. **Format:** Numeric
RAWCODECOUNT	Number of dispensings used to define index date defining records. **Note 1:** When multiple dispensings occur on the same day with same STOCKGROUP, each dispensing will be counted. **Format:** Numeric

x) [RUNID]_t5_CIDA_episdur.sas7bdat

This output table includes the number of episodes by episode number, episode length, sex, race, hispanic, and age group. Note that, multiple censoring criteria may be flagged as an episode is censored for multiple reasons on the same day. Table 75 contains specifications for the [*RUNID*]_t5_cida_episdur output table.

Table 75. [*RUNID*]_t5_cida_episdur Output

Variable	Description
GROUP	Cohort name. **Format:** Character(30)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character(3)
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(2)
AGEGROUP	Age Groups. Categories are requester-defined. Age is calculated at index date. **Format:** Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
RACE	Race. Allowable values are those in the SCDM. **Format:** Character(2)
HISPANIC	Hispanic. Allowable values are those in the SCDM. **Format:** Character(2)
EPISODENUM	Episode number (i.e., first episode =1, second episode=2, etc.) **Format:** Numeric
ZIP3	3-digit ZIP code. **Format**: Numeric
STATE	2-digit State code. **Format**: Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format**: Character(2)
CB_REG	Census Bureau regions. **Format**: Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date N: zip date occurred before the index date

Variable	Description
	Format: Character(1)
COVAR*n*	Requester-defined covariate **Note 1:** N will equal COVARNUM specified in the Covariate file **Format**: Numeric
EPISODELENGTH	Episode length (in days) **Format:** Numeric
CUMEPISODELENGTH	Cumulative Episode length (in days) **Format:** Numeric
EPISODES	Number of exposure episodes. **Format:** Numeric
NPTS	Number of Patients **Format:** Numeric

y) [RUNID]_T5_CIDA_episdur_censor.sas7bdat

This output table includes the number of episodes by episode number, episode length and reason(s) for censoring. Note that, multiple censoring criteria may be flagged as an episode is censored for multiple reasons on the same day. Table 76 contains specifications for the [*RUNID*] _t5_cida_episdur_censor output table.

Table 76. [RUNID]_t5_cida_ episdur_censor Output

Variable	Description
GROUP	Cohort name. **Format:** Character(30)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character(3)
EPISODENUM	Episode number (i.e., first episode =1, second episode=2, etc.) **Format:** Numeric
EPISODELENGTH	Episode length (in days) **Format:** Numeric
COVAR*n*	Requester-defined covariate **Note 1:** N will equal COVARNUM specified in the Covariate file **Format**: Numeric
CENS_ELIG	Number of episodes censored due to disenrollment **Format:** Numeric

Variable	Description
CENS_DTH	Number of episodes censored due to evidence of death **Format:** Numeric
CENS_DPEND	Number of episodes episode censored due to DP data end date **Format:** Numeric
CENS_QRYEND	Number of episodes censored due to query end date **Format:** Numeric
CENS_EPISEND	Number of episodes censored due to end of exposure **Format:** Numeric
CENS_SPEC	Number of episodes censored due to additional requester-defined criteria (e.g., censor due to occurrence of another set of clinical codes). **Format:** Numeric
EPISODES	Number of exposure episodes. **Format:** Numeric
NPTS	Number of Patients **Format:** Numeric

z) [RUNID]_t5_cida_gaps.sas7bdat

This output table includes the number of episode gaps by gap number, gap length, sex, race, hispanic, and age group. Gap number is determined as follows:

Gap 1 = gap in days between exposure episode 1 and exposure episode 2
Gap 2 = gap in days between exposure episode 2 and exposure episode 3
Gap n = gap in days between exposure episode n and exposure episode (n+1)

For the patient's last episode, the gap represents the number days from the end of the last episode to enrollment end/death/dp end date/query end date.

Table 77 contains specifications for the [*RUNID*] _t5_cida_gaps output table.

Table 77. [RUNID]_t5_cida_gaps Output

Variable	Description
GROUP	Cohort name. **Format:** Character(30)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.*, variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character(3)
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(2)

Variable	Description
AGEGROUP	Age Groups. Categories are requester-defined. Age is calculated at index date. **Format:** Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
RACE	Race. Allowable values are those in the SCDM. **Format:** Character(2)
HISPANIC	Hispanic. Allowable values are those in the SCDM. **Format:** Character(2)
GAPNUM	Gap number (i.e., first gap=1, second gap=2, etc.). For the patient's last episode, GAPNUM should be set to 999, and the gap should represent the number days from the end of the last episode to enrollment end/death/dp end date/query end date. **Format:** Numeric(8)
GAPLENGTH	Gap length (in days) **Format:** Numeric
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date N: zip date occurred before the index date **Format:** Character(1)
COVAR*n*	Requester-defined covariate **Note 1:** N will equal COVARNUM specified in the Covariate file Format: Numeric
EPISODES	Number of episodes **Format:** Numeric

aa) [RUNID]_profile_[T3OUT]_[T4OUT]_[PERIODID].sas7bdat

This output table describes patients and episodes by their unique combination of covariates observed. For any Type of analysis (1-5), requester can output a file describing all unique combinations of covariates for patients in the cohort [RUNID]_profile_[T3OUT]_[T4OUT]_[PERIODID].

Table 78 below contains specifications for the [RUNID]_profile_[T3OUT]_[T4OUT]_[PERIODID] output table, where RUNID is the request programmer-defined execution identifier.

For Type 3 analyses, T3OUT takes the value of _an and _an_censor for analysis and analysis_censor datasets; T3OUT is blank for all non-Type 3 analyses.

For Type 4 analyses, T4OUT takes the value of Preg or NoPreg_for the pregnant cohort and the non-pregnant cohort, respectively. T4OUT is blank for all non-Type 4 analyses.

Table 78. [RUNID]_profile_[T3OUT]_[T4OUT]_[PERIODID] Output

Variable	Description
GROUP	Standardized name used to differentiate cohorts. **Format:** Character(30)
COVAR_X	For each covariate specified in the Covariate Codes File, 1/0 indicator of presence/absence in the unique covariate profile. Number of columns dependent on number of binary covariates specified in the COVARIATECODES file. **Format:** Numeric(8)
PATIENT	Number of unique patients with the specified covariate profile (i.e., unique combination of covariates). **Format:** Numeric(8)
N_EPISODES	Number of unique episodes with the specified covariate profile (i.e., unique combination of covariates). **Format:** Numeric(8)

bb) [RUNID]_distindex.sas7bdat.sas7bdat

This output table describes exposure and/or HOI episodes by the unique combination of codes used to identify the exposure/HOI on the index date. The user can request this output for any Type of analysis (1-5). Table 79 below contains specifications for the [RUNID]_distindex output table, where RUNID is the request programmer-defined execution identifier

Table 79. [RUNID]_distindex Output

Variable	Description
GROUP	Cohort name. **Format:** Character(40)
DISTINDEXTYPE	Identifier for type of index event represented by row of data. Valid values include: EXP = exposure HOI = health outcome of interest **Format:** Character(5)
DISTINDEXLIST	Underscore-concatenated list of distributed index identifiers per distributed exposure type. Each list value represents a distinct combination of clinical codes associated with an index event. **Format:** Character(18)
EPISODES	Number of episodes within strata. **Format:** Numeric

cc) [RUNID]_distindexmap.sas7bdat.sas7bdat

This output table includes the mapping to the codes in the DISTINDEXLIST variable from [RUNID]_distindex output above. Table 80 below contains specifications for the [RUNID]_distindexmap output table, where RUNID is the request programmer-defined execution identifier.

Table 80. [RUNID]_distindexmap Output

Variable	Description
GROUP	Cohort name. **Format:** Character(40)
DISTINDEXTYPE	Identifier for type of index event represented by row of data. Valid values include: EXP = exposure HOI = health outcome of interest **Format:** Character(5)
STOCKGROUP	STOCKGROUP value (populated when exp or HOI is RX) **Format:** Character(18)
CODECAT	Clinical code category. (populated when exp or HOI is PX/DX) **Format:** Character(2)
CODETYPE	Clinical code type. (populated when exp or HOI is PX/DX)

Variable	Description
	Format: Character(2)
ENCTYPE	Care setting. (populated when exp or HOI is PX/DX) **Format:** Character(3)
PDX	Prinicpal diagnosis position (populated when exp or HOI is PX/DX) **Format:** Character(1)
CODE	PX or DX code (populated when exp or HOI is PX/DX) **Format:** Character(11)
DISTINDEXID	Index code combination identifier **Format:** Character

dd) [RUNID]_t6_productsdates.sas7bdat

This summary table provides overview of the calendar dates that were computed and may have been used to compute various episode durations and switching metrics.

Table 81. [RUNID]_t6_productsdates output

Variable	Description
GROUP	Cohort name. **Format:** Character(40)
PRODUCTMARKETINGDATE	Product Start Marketing Date **Format**: Date9.
PRODUCTAPPROVALDATE	Product Approval Date **Format**: Date9.
OTHERPRODUCTDATE	Other product date **Format**: Date9.
COMPUTEDSTARTMARKETINGDATE	Computed marketing Start Date **Format:** Date9.

ee) [RUNID]_t6_utilcounts.sas7bdat

This output table includes number of incident users, all users, and dispensings.

Table 82. [RUNID]_t6_ utilcounts output

Variable	Description
GROUP	Cohort name. **Format:** Character(40)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character(4)
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(2)
AGEGROUP	Age Groups. Categories are requester-defined. **Format:** Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
YEAR	Year of index date. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric
RACE	Race Indicator **Format:** Character(1)
HISPANIC	Hispanic Indicator **Format:** Character(1)
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date N: zip date occurred before the index date

Variable	Description
	Format: Character(1)
NINCUSERS	Number of incident users. **Format:** Numeric
NALLUSERS	Number of all users. **Format:** Numeric
AdjustedCodeCount	Number of dispensings used to define index date defining records. **Note 1:** When multiple dispensings occur on the same day with same STOCKGROUP, this will be counted as 1 dispensing. This count will equal the count of the DISPENSINGS metric in prior QRP versions. **Format:** Numeric
RawCodeCount	Number of dispensings used to define index date defining records. **Note 1:** When multiple dispensings occur on the same day with same STOCKGROUP, each dispensing will be counted. **Format:** Numeric

ff) [RUNID]_t6_trendcounts.sas7bdat

This output table includes prevalent counts of users and are based on use during the entire episode. Prevalent dispensings are based on index date of dispensing. Except for Year and Month, all other stratifications are based on the index date of episode. This table does not adjust AgeGroup or Geographic strata over time.

Table 83. [RUNID]_t6_trendcounts output

Variable	Description
GROUP	Cohort name. **Format:** Character(40)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character(4)
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(2)
AGEGROUP	Age Groups. Categories are requester-defined. **Format:** Character(variable)

Variable	Description
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
YEAR	Year of index date. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric
RACE	Race Indicator **Format:** Character(1)
HISPANIC	Hispanic Indicator **Format:** Character(1)
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date N: zip date occurred before the index date **Format:** Character(1)
NALLUSERS	Number of all users. **Format:** Numeric
AdjustedCodeCount	Number of dispensings used to define index date defining records. **Note 1:** When multiple dispensings occur on the same day with same STOCKGROUP, this will be counted as 1 dispensing. This count will equal the count of the DISPENSINGS metric in prior QRP versions. **Format:** Numeric
RawCodeCount	Number of dispensings used to define index date defining records. **Note 1:** When multiple dispensings occur on the same day with same STOCKGROUP, each dispensing will be counted. **Format:** Numeric

gg) [RUNID]_t6_utildispstats.sas7bdat

For overall and for each stratification level, this output table includes one row per each dispensing days supply value and a count of dispensings with that duration.

Table 84. [RUNID]_t6_utildispstats output

Variable	Description
GROUP	Cohort name. **Format:** Character(40)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character(4)
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(2)
AGEGROUP	Age Groups. Categories are requester-defined. **Format:** Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
YEAR	Year of index date. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric
RACE	Race Indicator **Format:** Character(1)
HISPANIC	Hispanic Indicator **Format:** Character(1)
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date N: zip date occurred before the index date

Variable	Description
	Format: Character(1)
DaySupp	Number of days supplied **Format:** Numeric
Count	Number of dispensings used to define index date defining records. **Format:** Numeric

hh) [RUNID]_t6_utilepis_censor.sas7bdat

Overall and for each stratification level, this output table includes number of treatment episodes and a count of episodes censored in the given category.

Table 85. [RUNID]_t6_utilepis_censor output

Variable	Description
GROUP	Cohort name. **Format:** Character(40)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character(4)
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(2)
AGEGROUP	Age Groups. Categories are requester-defined. **Format:** Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
YEAR	Year of index date. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric
RACE	Race Indicator **Format:** Character(1)
HISPANIC	Hispanic Indicator **Format:** Character(1)
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code.

Variable	Description
	Format: Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date N: zip date occurred before the index date **Format:** Character(1)
EPISODELENGTH	Treatment Episode Duration **Format**: Numeric
EndQueryCount	Count of number of treatment episodes that ended due to the end of the query period **Format**: Numeric
EndEnrollmentCount	Count of number of treatment episodes that ended due to the end of enrollment **Format**: Numeric
EndAvailDataCount	Count of number of treatment episodes that ended due to the end of the available data **Format**: Numeric
EndProductDiscontinuationCount	Count of number of treatment episodes that ended due to product discontinuation (end of treatment episode with no observed switch to next product in switch pattern) **Format**: Numeric
DeathCount	Count of number of treatment episodes that ended due to death **Format**: Numeric

ii) [RUNID]_t6_utilepisdurstats.sas7bdat

Overall and for each stratification level, this output table includes one row per each treatment episode duration and a count of unique patients in that treatment duration.

Table 86. [RUNID]_t6_utilepisdurstats output

Variable	Description
GROUP	Cohort name. **Format:** Character(40)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character(4)
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(2)
AGEGROUP	Age Groups. Categories are requester-defined. **Format:** Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
YEAR	Year of index date. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric
RACE	Race Indicator **Format:** Character(1)
HISPANIC	Hispanic Indicator **Format:** Character(1)
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date

Variable	Description
	N: zip date occurred before the index date **Format:** Character(1)
CUMEPISODELENGTH	Cumulative treatment episode duration, in number of days **Format**: Numeric
Npts	Count of number of unique patients represented in that treatment duration **Format**: Numeric

jj) [RUNID]_t6_utiluptakestats.sas7bdat

Overall and for each stratification level, this output table includes one row per each product uptake duration and a count of the number of treatment episodes with that duration.

Table 87. [RUNID]_t6_utiluptakestats output

Variable	Description
GROUP	Cohort name. **Format:** Character(40)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character(4)
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(2)
AGEGROUP	Age Groups. Categories are requester-defined. **Format:** Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
YEAR	Year of index date. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric
RACE	Race Indicator **Format:** Character(1)
HISPANIC	Hispanic Indicator **Format:** Character(1)
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code.

Variable	Description
	Format: Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date N: zip date occurred before the index date **Format:** Character(1)
UPTAKEDAYS	Number of days from [start date] until uptake **Format**: Numeric
COUNT	Count of number of treatment episodes **Format**: Numeric

kk) [RUNID]_t6_switchattrition.sas7bdat

This attrition file shows how many switch pattern episodes are originally selected for inclusion and at each operational level, how many switch pattern episodes are retained or discarded due to the reasons described in field "DESCR".

- All switch patterns episodes (the starting point)
- Effect of using switch cohort inclusion/exclusion criteria (INLUSIONCODES input file criteria)
- Effect of applying switch cohort entering criteria (SWITCHCOHORTINCLDATE and SWITCHDATEUSE)
- Effect of retaining one versus all switch pattern episode per person (SWITCHCOHORTDEF)

Table 88. [RUNID]_t6_switchattrition output

Variable	Description
ANALYSISGRP	Highest-level cohort name. **Format:** Character(30)
LEVEL	Criterion identifier. **Format:** Numeric
DESCR	Criterion description. **Format:** Character(500)
REMAINING	Number of individuals remaining after previous exclusion criterion. **Format:** Numeric
EXCLUDED	Number of individuals excluded by the exclusion criterion. **Format:** Numeric

ll) [RUNID]_t6_switchplota.sas7bdat

This output table includes information on the number of switch pattern episodes and reason for censoring for every day of follow-up.

Table 89. [RUNID]_t6_switchplota output

Variable	Description
ANALYSISGRP	Cohort name. **Format:** Character(40)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character(4)
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(2)
AGEGROUP	Age Groups. Categories are requester-defined. **Format:** Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
YEAR	Year of index date. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric
RACE	Race Indicator **Format:** Character(1)
HISPANIC	Hispanic Indicator **Format:** Character(1)
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date

Variable	Description
	N: zip date occurred before the index date **Format:** Character(1)
TTSWITCH	Duration, in number of days **Format**: Numeric
SwitchedCount	Count of number of switch pattern episodes that ended due to switch to product in switch pattern **Format**: Numeric
EndQueryCount	Count of number of switch pattern episodes that ended due to the end of the query period **Format**: Numeric
EndEnrollmentCount	Count of number of switch pattern episodes that ended due to the end of enrollment **Format**: Numeric
EndAvailDataCount	Count of number of switch pattern episodes that ended due to the end of the available data **Format**: Numeric
ProductDiscontinuationCount	Count of number of switch pattern episodes that ended due to product discontinuation (end of treatment episode with no observed switch to next product in switch pattern) **Format**: Numeric
DeathCount	Count of number of switch pattern episodes that ended due to death **Format**: Numeric
SwitchedPatCount	Count of number of switch pattern patients that ended due to switch to product in switch pattern **Format**: Numeric
EndQueryPatCount	Count of number of switch pattern episodes that ended due to the end of the query period **Format**: Numeric
EndEnrollmentPatCount	Count of number of switch pattern patients that ended due to the end of enrollment **Format**: Numeric
EndAvailDataPatCount	Count of number of switch pattern patients that ended due to the end of the available data **Format**: Numeric
ProductDiscontinuationPatCount	Count of number of switch pattern patients that ended due to product discontinuation (end of treatment episode with no observed switch to next product in switch pattern) **Format**: Numeric

Variable	Description
DeathPatCount	Count of number of switch pattern patients that ended due to death **Format**: Numeric

mm) [RUNID]_t6_switchplotb.sas7bdat

This output table includes information on the number of switch pattern episodes with a second switch and reason for censoring for every day of follow-up.

Table 90. [RUNID]_t6_switchplotb output

Variable	Description
ANALYSISGRP	Cohort name. **Format:** Character(40)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character(4)
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(2)
AGEGROUP	Age Groups. Categories are requester-defined. **Format:** Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
YEAR	Year of index date. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric
RACE	Race Indicator **Format:** Character(1)
HISPANIC	Hispanic Indicator **Format:** Character(1)
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)

Variable	Description
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date N: zip date occurred before the index date **Format:** Character(1)
TTSWITCH	Duration, in number of days **Format**: Numeric
SwitchedCount	Count of number of switch pattern episodes that ended due to switch to product in switch pattern **Format**: Numeric
EndQueryCount	Count of number of switch pattern episodes that ended due to the end of the query period **Format**: Numeric
EndEnrollmentCount	Count of number of switch pattern episodes that ended due to the end of enrollment **Format**: Numeric
EndAvailDataCount	Count of number of switch pattern episodes that ended due to the end of the available data **Format**: Numeric
ProductDiscontinuationCount	Count of number of switch pattern episodes that ended due to product discontinuation (end of treatment episode with no observed switch to next product in switch pattern) **Format**: Numeric
DeathCount	Count of number of switch pattern episodes that ended due to death **Format**: Numeric
SwitchedPatCount	Count of number of switch pattern patients that ended due to switch to product in switch pattern **Format**: Numeric
EndQueryPatCount	Count of number of switch pattern episodes that ended due to the end of the query period **Format**: Numeric
EndEnrollmentPatCount	Count of number of switch pattern patients that ended due to the end of enrollment **Format**: Numeric
EndAvailDataPatCount	Count of number of switch pattern patients that ended due to the end of the available data **Format**: Numeric

Variable	Description
ProductDiscontinuationPatCount	Count of number of switch pattern patients that ended due to product discontinuation (end of treatment episode with no observed switch to next product in switch pattern) **Format**: Numeric
DeathPatCount	Count of number of switch pattern patients that ended due to death **Format**: Numeric

nn) [RUNID]_t6_switchepisdurstats.sas7bdat

Overall and for each stratification level, this output table includes one row per each treatment episode length and a count of the number of treatment episodes with that duration.

Table 91. [RUNID]_t6_ switchepisdurstats output

Variable	Description
GROUP	Cohort name. **Format:** Character(40)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. Allowable values will contain those specified in the strata levels input file. **Format:** Character(4)
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(2)
AGEGROUP	Age Groups. Categories are requester-defined. **Format:** Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
YEAR	Year of index date. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric
RACE	Race Indicator **Format:** Character(1)
HISPANIC	Hispanic Indicator **Format:** Character(1)
ZIP3	3-digit ZIP code. **Format:** Numeric

Variable	Description
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date N: zip date occurred before the index date **Format:** Character(1)
EPISODELENGTH	Treatment episode duration, in number of days **Format:** Numeric
COUNT	Count of number of treatment episodes **Format:** Numeric

2. DPLOCAL Folder

QRP generates output to the DPLOCAL folder based on the type of analysis specified.

Background Rate Calculations (Type 1 Analyses)

This analysis generates four output tables:

- [*RUNID*]_mstr.sas7bdat
- [*RUNID*]_numcounts.sas7bdat
- [*RUNID*]_denomcounts.sas7bdat
- [*RUNID*]_ads_mstr_[*PERIODID*].sas7bdat

Where RUNID is the request programmer-defined execution identifier (defined in main macro parameters) and PERIODID is the time period under analysis specified in the Monitoring File.

Exposures and Follow-up (Type 2 Analyses)

This analysis generates four output tables:

- [*RUNID*]_mstr.sas7bdat
- [*RUNID*]_numcounts.sas7bdat
- [*RUNID*]_denomcounts.sas7bdat
- [*RUNID*]_ads_mstr_[*PERIODID*].sas7bdat

Where RUNID is the request programmer-defined execution identifier (defined in main macro parameters) and PERIODID and the time period under analysis specified in the Monitoring File.

Self-controlled Risk Interval Design (Type 3 Analyses)

This analysis generates five output tables:

- [*RUNID*]_mstr.sas7bdat
- [*RUNID*]_numcounts.sas7bdat
- [*RUNID*]_ads_mstr_[*PERIODID*].sas7bdat
- [*RUNID*]_mstr_censor.sas7bdat
- [*RUNID*]_numcounts_censor.sas7bdat

Where RUNID is the request programmer-defined execution identifier.

Pregnancy Episodes and Medical Product Use (Type 4 Analyses)

This analysis generates seven output tables:

- [*RUNID*]_mstr.sas7bdat
- [*RUNID*]_ads_ctrl_[*PERIODID*].sas7bdat
- [*RUNID*]_ctrl.sas7bdat
- [*RUNID*]_alldeliveries.sas7bdat
- [*RUNID*]_pregdurcodes.sas7bdat
- [*RUNID*]_sec.sas7bdat
- [*RUNID*]_ads_mstr_[*PERIODID*].sas7bdat

Where RUNID is the request programmer-defined execution identifier (defined in main macro parameters) and PERIODID is the time period under analysis specified in the Monitoring File.

Medical Product Utilization (Type 5 Analyses)

This analysis generates three output tables:

- [*RUNID*]_mstr.sas7bdat
- [*RUNID*]_dispensings.sas7bdat
- [*RUNID*]_ads_mstr_[*PERIODID*].sas7bdat

Where RUNID is the request programmer-defined execution identifier (defined in main macro parameters) and PERIODID is the time period under analysis specified in the Monitoring File.

Output tables for all types of analyses are described below.

Manufacturer-level Product Utilization and Switching Patterns Cohort Identification Strategy (Type 6 Analyses)

This analysis generates two output tables:

- [RUNID]_mstr.sas7bdat
- [RUNID]_t6_switchepisodes.sas7bdat

For Types 1-5, whenever MFU or HDPS analyses are performed, then the following outputs are also generated:

- [*RUNID*]_claims_icddx09
- [*RUNID*]_claims_icddx10
- [*RUNID*]_claims_icddxOT
- [*RUNID*]_claims_icdpx09
- [*RUNID*]_claims_icdpx10

- [*RUNID*]_claims_icdpxOT
- [*RUNID*]_claims_cpt
- [*RUNID*]_claims_hcpcs
- [*RUNID*]_claims_drugclass
- [*RUNID*]_claims_rx
- [*RUNID*]_claims_lab

a) [RUNID]_mstr.sas7bdat

The [*RUNID*] _mstr output table contains one record per individual per index date for every cohort (GROUP) specified in the CIDA tool execution (*i.e.,* one row per exposure episode per RUNID). The [*RUNID*] _mstr output table contents depend on the type of cohort identification strategy used. Table 92 contains specifications for the [*RUNID*] _mstr output table for Type 1 analyses.

Table 92. [RUNID]_mstr Output for Type 1 Analyses

Variable	Description
GROUP	Cohort name. **Format:** Character(40)
TYPE	Identifies the type of analysis performed in the CIDA tool (=1). **Format**: Numeric
PATID	Individual identifier. **Format:** Character(variable)
INDEXDT	Index date. **Format**: Numeric (date9.)
ENR_START	Start of enrollment span where index exposure was identified. **Format**: Numeric (date9.)
ENR_END	End of enrollment span where index exposure was identified. **Format**: Numeric (date9.)
YEAR	Year of index date. **Format**: Numeric
AdjustedDisp	Number of outpatient pharmacy dispensings used to define the index date. When 2 or more dispensings with the same RxDate are stockpiled, this metric will count as 1 dispensing **Format**: Numeric
RawDisp	Number of outpatient pharmacy dispensings used to define the index date. When 2 or more dispensings with the same RxDate are stockpiled, this metric will count as 2 (or more) dispensings **Format**: Numeric
TOTRXSUP	Days supplied (RXSUP) value associated with the dispensing that defined the index date. **Format**: Numeric

Variable	Description
TOTRXAMT	Amount supplied (RXAMT) value associated with the dispensing that defined the index date. **Format**: Numeric
EPISODEENDDT	Exposure episode end date. For Type 1 analyses, the EPISODEENDDT = INDEXDT + TORXSUP. **Format**: Numeric (date9.)
FEVENTDT	Blank for Type 1 analyses. **Format**: Numeric
AGE	Age at index date. **Format**: Numeric
BIRTH_DATE	Birth Date. **Format**: Numeric (MMDDYY10.)
SEX	Sex. Allowable values are those in the SCDM Demographic table. **Format:** Character(1)
RACE	Race. Allowable values are those in the SCDM Demographic table. **Format:** Character(1)
HISPANIC	Hispanic. Allowable values are those in the SCDM Demographic table. **Format**: Character(1)
AGEGROUP	Age Group at index date. **Format:** Character(variable)
MINAGEDT	Date on which the individual qualified to enter the cohort, based on age ranges specified. **Format**: Numeric (date9.)
MAXAGEDT	Last date the individual qualified to be in the cohort, based on age ranges specified. **Format**: Numeric (date9.)
ZIP	Zip code. **Format:** Character(5)
ZIP_DATE	Earliest date that the ZIP code is believed to be valid **Format**: Numeric (date9.)
STATE	State associated with the zip code **Format:** Character
DISTINDEXEXP	Underscore-concatenated list of distributed index identifiers associated with exposure. **Format:** Character(250)
DISTINDEXHOI	Underscore-concatenated list of distributed index identifiers associated with HOI. **Format:** Character(250)

Table 93 contains specifications for the [*RUNID*] _mstr output table for Type 2 analyses.

Table 93. [RUNID] _mstr Output for Type 2 Analyses

Variable	Description
GROUP	Cohort name. **Format**: Character(40)
TYPE	Identifies the type of analysis performed in the CIDA tool (=2). **Format**: Numeric
PATID	Individual identifier. **Format**: Character(variable)
INDEXDT	Index date. **Format**: Numeric (date9.)
ENR_START	Start of enrollment span where index exposure was identified. **Format**: Numeric (date9.)
ENR_END	End of enrollment span where index exposure was identified. **Format**: Numeric (date9.)
YEAR	Year of index date. **Format**: Numeric
EPISODEENDDT	Exposure episode end date. Important note: this date does not take into account the HOI date (if an HOI is observed). This is the exposed time censored for the earliest of 1) end of exposure period; 2) disenrollment; 3) maximum exposure episode length (MAXEPISDUR) settings; and 4) any additional censoring criteria specified but NOT the HOI date **Format**: Numeric (date9.)
AdjustedDisp	Number of outpatient pharmacy dispensings used to define the index date. When 2 or more dispensings with the same RxDate are stockpiled, this metric will count as 1 dispensing **Format**: Numeric
RawDisp	Number of outpatient pharmacy dispensings used to define the index date. When 2 or more dispensings with the same RxDate are stockpiled, this metric will count as 2 (or more) dispensings **Format**: Numeric
TOTRXSUP	Total days supplied for outpatient pharmacy dispensings used to build the exposure episode. For requester-defined follow-up time (i.e., when exposure episodes are *not* created using dispensing days supply), this value will always be populated with the RxSup value associated with the dispensing that defined the index date. Value will always be =0 for never-exposed cohort. **Format**: Numeric
TOTRXAMT	Total amount supplied for outpatient pharmacy dispensings used to build the exposure episode. For requester-defined follow-up time (i.e., when

Variable	Description
	exposure episodes are *not* created using dispensing days supply), this value will always be populated with the RxAmt value associated with the dispensing that defined the index date. Value will always be =0 for never-exposed cohort. **Format**: Numeric
NUMEVENTS	Total number of HOIs observed during the exposure episode (note: days at risk [TTE] stop accumulating after the occurrence of the first HOI). **Format**: Numeric
FEVENTDT	HOI date. **Format**: Numeric (MMDDYY10.)
TTE	Days-at-risk (*i.e.,* time to HOI in days). Days-at-risk metrics stop accumulating at the earliest of the following: 1) end of exposure period; 2) disenrollment; 3) any additional censoring criteria [defined using CIDA tool, including MAXEPISDUR settings]; and 4) the HOI date. **Format**: Numeric
IndexLook	Look # individual was identified in. **Format:** Numeric
LastLookFollowed	When an individual is lost to follow up (data is no longer being updated), this is the last look # the individual contributed data to. If an individual is still eligible to be followed in subsequent looks, this is set to 0. Value will always be blank for never-exposed cohort. **Format:** Numeric
LastLookFollowedDt	The last date of the Look specified in LastLookFollowed. Set to missing if LastLookFollowed = 0. Value will always be blank for never-exposed cohort. **Format:** Numeric (date9.)
AGE	Age at index date. **Format**: Numeric
BIRTH_DATE	Birth Date. **Format**: Numeric (MMDDYY10.)
SEX	Sex. Allowable values are those in the SCDM Demographic table. **Format**: Character(1)
RACE	Race. Allowable values are those in the SCDM Demographic table. **Format**: Character(1)
HISPANIC	Hispanic. Allowable values are those in the SCDM Demographic table. **Format**: Character(1)
AGEGROUP	Age Group at index date. **Format**: Character(variable)

Variable	Description
AGEGROUPNUM	Age Group number. **Format**: Numeric
EPISODETYPE	Indicates if exposed time is requester-defined number of days (value will be ITT) or determined using dispensing days supplied (value will be EPI). **Format**: Character(3)
DEATHDT	Date on which the individual died, if any (value is set to missing if the individual did not die). **Format**: Numeric (date9.)
MINAGEDT	Date on which the individual qualified to enter the cohort, based on age ranges specified. **Format**: Numeric (date9.)
MAXAGEDT	Last date the individual qualified to be in the cohort, based on age ranges specified. **Format**: Numeric (date9.)
ZIP3	Zip code. **Note**: will be missing if geographic information is not requested. **Format:** Character(5)
ZIP_DATE	Earliest date that the ZIP code is believed to be valid. **Format**: Numeric (date9.)
STATE	State associated with the zip code. **Note**: will be missing if geographic information is not requested. **Format:** Character
HHS_REG	Health and Human Services region associated with the zip code. **Note:** will be missing if geographic information is not requested. **Format:** Character
CB_REG	Census Bureau region associated with the zip code. **Note:** will be missing if geographic information is not requested. **Format:** Character
EVENT_FLAG	Indicates whether censoring is determined by occurrence of an event. **Format:** Character(1)
CENS_ELIG	Indicates whether the episode is censored due to disenrollment. **Format:** Numeric (8)
CENS_DTH	Indicates whether the episode is censored due to evidence of death. **Format:** Numeric (8)
CENS_DPEND	Indicates whether the episode is censored due to DP data end date (based on DP_MaxDate in common components). **Format:** Numeric (8)

Variable	Description
CENS_QRYEND	Indicates whether the episode is censored due to query end date. **Format:** Numeric (8)
CENS_EPISEND	Indicates whether the episode is censored due to episode end date. **Format:** Numeric (8)
CENS_SPEC	Indicates whether the episode is censored due to additional requester-defined criteria (e.g., censor due to occurrence of another set of clinical codes). **Format:** Numeric (8)
CENS_EVENT	Indicates whether the episode is censored due to occurrence of request-defined event. **Format:** Numeric (8)
EXACTNUMVISIT	The total number of medical encounters during the requester-defined evaluation period. **Format**: Numeric
NUMVISITS_AV	The total number of ambulatory encounters during the requester-defined evaluation period. **Format**: Numeric
NUMVISITS_OA	The total number of other ambulatory encounters during the requester-defined evaluation period. **Format**: Numeric
NUMVISITS_IP	The total number of inpatient encounters during the requester-defined evaluation period. **Format**: Numeric
NUMVISITS_IS	The total number of institutional stay encounters during the requester-defined evaluation period. **Format**: Numeric
NUMVISITS_ED	The total number of emergency department encounters during the requester-defined evaluation period. **Format**: Numeric
NUMVISITS	The requester-defined stratum that the EXACTNUMVISIT value is in. **Format**: Numeric
NUMGENERIC	The total number of generic names dispensed during the requester-defined evaluation period. **Format**: Numeric
NUMCLASS	The total number of unique class names dispensed during the requester-defined evaluation period. **Format**: Numeric
NUMRX	The total number of dispensings during the requester-defined evaluation period.

Variable	Description
	Format: Numeric
COMBINED_SCORE	The comorbidity score requester-defined category. **Format**: Character(3)
COMBINED_SCORE_NUM	The comorbidity score value. **Format**: Numeric
CCIELIXGRP	The pre-index comorbidity score requester-defined category. **Format**: Character(10)
DISTINDEXEXP	Underscore-concatenated list of distributed index identifiers associated with exposure. **Format:** Character(250)
DISTINDEXHOI	Underscore-concatenated list of distributed index identifiers associated with HOI. **Format:** Character(250)

Table 94 contains specifications for the [*RUNID*] _mstr output table for Type 3 analyses.

Table 94. [RUNID] _mstr Output for Type 3 Analyses

Variable	Description
GROUP	Cohort name. **Format**: Character(40)
TYPE	Identifies the type of analysis performed in the CIDA tool (=3). **Format**: Numeric
PATID	Individual identifier. **Format**: Character(variable)
INDEXDT	Index date. **Format**: Numeric (date9.)
ENR_START	Start of enrollment span where index exposure was identified. **Format**: Numeric (date9.)
ENR_END	End of enrollment span where index exposure was identified. **Format**: Numeric (date9.)
YEAR	Year of index date. **Format**: Numeric
FEVENTDT	HOI date. **Format**: Numeric (date9.)
AGE	Age at index date. **Format**: Numeric
BIRTH_DATE	Birth Date. **Format**: Numeric (MMDDYY10.)
SEX	Sex. Allowable values are those in the SCDM Demographic table.

Variable	Description
	Format: Character(1)
RACE	Race. Allowable values are those in the SCDM Demographic table. **Format**: Character(1)
HISPANIC	Hispanic. Allowable values are those in the SCDM Demographic table. **Format**: Character(1)
AGEGROUP	Age Group at index date. **Format**: Character(variable)
MINENROLHOIMET	Indicates if minimum enrollment requirements to determine HOI incidence are met. Will equal 1 if requirement is met; will be set to missing when enrollment is sufficient to determine exposure incidence but insufficient to determine HOI incidence (when a window start date is prior to enrollment start date). **Format**: Numeric
WINOFEVENT	Indicates whether the HOI occurred during the risk or control window. Valid values include "RISK" and "CTRL". **Format**: Character(4)
PT_ANALYSIS_EPISODE_NUM	Indicates the episode number for a patient in the analysis cohort. When T3COHORTDEF=02 a patient can have multiple episodes. This indicator can take any integer value. **Format**: Numeric
PT_EXPOSURE_EPISODE_NUM	Indicates the episode number for a patient in the exposure cohort. When T3COHORTDEF=02 a patient can have multiple episodes. This indicator can take any integer value. **Format**: Numeric
TTC	Time to censor. Represents the number of days between the date of a censoring condition (e.g. death date) and the index date **Format**: Numeric
CENSORDEATH	Indicates if exposure episode was censored due to death. CENSORDEATH =1 if censored; else 0. Note: if death and disenrollment occur on the same day, censoring will be attributed to death. **Format**: Numeric
CENSORENROL	Indicates if exposure episode was censored due to disenrollment. CENSORENROL =1 if censored; else 0. Note: if death and disenrollment occur on the same day, censoring will be attributed to death. **Format**: Numeric

Variable	Description
ANALYSISCOHORT	Indicates if patient's exposure episode contributed to the analytic cohort. Will =1 for yes, =0 for no. **Format**: Numeric
DAYS_FROM_EXPOS_TO_EVENT	Number of days from the index date to the HOI date. **Format**: Numeric
DAYS_FROM_EXPOS_TO_CTRL_EVENT	Number of days from the index date to the control interval HOI date. **Format**: Numeric
DAYS_FROM_EXPOS_TO_RISK_EVENT	Number of days from the index date to the risk interval HOI date. **Format**: Numeric
POSTENR_EXPOSURE	Number of days of continuous enrollment post-exposure. Calculated as Enr_End – IndexDt +1. **Format**: Numeric
EXACTNUMVISIT	The total number of medical encounters during the requester-defined evaluation period. **Format:** Numeric
NUMVISITS_AV	The total number of ambulatory encounters during the requester-defined evaluation period. **Format:** Numeric
NUMVISITS_OA	The total number of other ambulatory encounters during the requester-defined evaluation period. **Format:** Numeric
NUMVISITS_IP	The total number of inpatient encounters during the requester-defined evaluation period. **Format:** Numeric
NUMVISITS_IS	The total number of institutional stay encounters during the requester-defined evaluation period. **Format:** Numeric
NUMVISITS_ED	The total number of emergency department encounters during the requester-defined evaluation period. **Format**: Numeric
NUMVISITS	The requester-defined stratum that the EXACTNUMVISIT value is in. **Format:** Numeric
NUMGENERIC	The total number of generic names dispensed during the requester-defined evaluation period. **Format**: Numeric
NumClass	The total number of unique class names dispensed during the requester-defined evaluation period.

Variable	Description
	Format: Numeric
NUMRX	The total number of dispensings during the requester-defined evaluation period. **Format**: Numeric
COMBINED_SCORE	The comorbidity score requester-defined category. **Format:** Character(3)
COMBINED_SCORE_NUM	The comorbidity score value. **Format:** Numeric
CCIELIXGRP	The pre-index comorbidity score requester-defined category. **Format:** Character(10)
DISTINDEXEXP	Underscore-concatenated list of distributed index identifiers associated with exposure. **Format:** Character(250)
DISTINDEXHOI	Underscore-concatenated list of distributed index identifiers associated with HOI. **Format:** Character(250)

Table 95 contains specifications for the [*RUNID*] _mstr_[T4COHORT] output table for Type 4 analyses. T4COHORT takes the value of Preg or NoPreg for the pregnant cohort and the non-pregnant cohort, respectively.

Table 95. [RUNID] _mstr_[T4COHORT] Output for Type 4 Analyses

Variable	Description
GROUP	Cohort name. **Format**: Character(40)
PATID	Individual identifier. **Format**: Character(variable)
TYPE	Identifies the type of analysis performed in the CIDA tool (=4). **Format**: Numeric
YEAR	Year of index date. **Format**: Numeric
INDEXDT	Index date (pregnancy start date). Can be determined by a preterm/postterm code. **Format**: Numeric (date9.)
ENR_START	Start of enrollment span where index exposure was identified. **Format**: Numeric (date9.)
ENR_END	End of enrollment span where index exposure was identified. **Format**: Numeric (date9.)

Variable	Description
EPISODESTARTDT	Pregnancy episode start date. Can be determined by a preterm/postterm code. **Format**: Numeric (date9.)
EPISODEENDDT	Pregnancy episode end date (delivery). **Format**: Numeric (date9.)
AGE	Age at index date. **Format**: Numeric
BIRTH_DATE	Birth Date. **Format**: Numeric (MMDDYY10.)
SEX	Sex. Allowable values are those in the SCDM Demographic table. **Format**: Character(1)
RACE	Race. Allowable values are those in the SCDM Demographic table. **Format**: Character(1)
HISPANIC	Hispanic. Allowable values are those in the SCDM Demographic table. **Format**: Character(1)
AGEGROUP	Age Group at delivery date. **Format**: Character(variable)
MINAGEDT	Date on which the individual qualified to enter the cohort, based on age ranges specified. **Format**: Numeric (date9.)
MAXAGEDT	Last date the individual qualified to be in the cohort, based on age ranges specified. **Format**: Numeric (date9.)
WASHPER	Washout used for the pregnancy cohort **Format**: Numeric
T4FUPWASHPER	Washout used for the pregnancy cohort HOI **Format**: Numeric
FEVENTDT	Blank for Type 4 analyses. **Format**: Numeric
ENRDAYSFLOOR	Minimum number of days of continuous enrollment required prior to the delivery date. **Format**: Numeric
ADATE	Delivery date **Format**: Numeric (MMDDYY10.)
DEATHDT	Date on which the individual died, if any. (value will be set to missing if the individual did not die) **Format**: Numeric (date9.)

Variable	Description
CONCWASHPER	Washout used to determine incidence of medical product use episodes with respect to delivery episodes when assessing "trimester only" indicators **Format**: Numeric
HADPRETERM	Indicator if pregnancy episode has a preterm birth code **Format**: Numeric
HADPOSTTERM	Indicator if pregnancy episode has a postterm birth code **Format**: Numeric
HADNONE	Indicator if pregnancy episode has neither a preterm or a postterm birth code **Format**: Numeric
PREPOSTIND	Indicator for type of Preterm/Postterm code (PRETERM, POSTTERM, NONE) **Format**: Character (10)
PREDURCODE	Prioritygroup1/Prioritygroup2 code, if any **Format**: Character (11)
PSADATE	Prioritygroup1/Prioritygroup2 code date, if any **Format**: (MMDDYY10.)
HAS3TRIM	Indicates if the pregnancy episode reaches the 3rd trimester **Format**: Numeric
DELNUM	Delivery number for the patient. This is the overall delivery number for the patient and is not confined to valid deliveries **Format**: Numeric
DELYEAR	Year of delivery date. **Format**: Numeric
DURATION	Length of pregnancy episode **Format**: Numeric
EXACTNUMVISIT	The total number of medical encounters during the requester-defined evaluation period. **Format:** Numeric
NUMVISITS_AV	The total number of ambulatory encounters during the requester-defined evaluation period. **Format:** Numeric
NUMVISITS_OA	The total number of other ambulatory encounters during the requester-defined evaluation period. **Format:** Numeric
NUMVISITS_IP	The total number of inpatient encounters during the requester-defined evaluation period. **Format:** Numeric

Variable	Description
NUMVISITS_IS	The total number of institutional stay encounters during the requester-defined evaluation period. **Format:** Numeric
NUMVISITS_ED	The total number of emergency department encounters during the requester-defined evaluation period. **Format:** Numeric
NUMVISITS	The requester-defined stratum that the EXACTNUMVISIT value is in. **Format:** Numeric
NUMGENERIC	The total number of generic names dispensed during the requester-defined evaluation period. **Format**: Numeric
NumClass	The total number of unique class names dispensed during the requester-defined evaluation period. **Format**: Numeric
NUMRX	The total number of dispensings during the requester-defined evaluation period. **Format**: Numeric
COMBINED_SCORE	The comorbidity score requester-defined category. **Format:** Character(3)
COMBINED_SCORE_NUM	The comorbidity score value. **Format:** Numeric
CCIELIXGRP	The pre-index comorbidity score requester-defined category. **Format:** Character(10)
DISTINDEXEXP	Underscore-concatenated list of distributed index identifiers associated with exposure. **Format:** Character(250)
DISTINDEXHOI	Underscore-concatenated list of distributed index identifiers associated with HOI. **Format:** Character(250)

Table 96 contains specifications for the [*RUNID*] _mstr_MI output table for Type 4 analyses. This table is produced for exposed and comparator/unexposed pregnant cohorts.

Table 96. [RUNID] _mstr_MI Output for Type 4 Analyses

Variable	Description
ANALYSISGRP	Requester defined name of MI exposure-comparator group **Format**: Character(40)
GROUP	Cohort name (either &analysisgrp_exp or &analysisgrp_comp) **Format**: Character(40)
TYPE	Identifies the type of analysis performed in the CIDA tool (=4). **Format**: Numeric
PATID	Individual identifier. **Format**: Character(variable)
PREGSTARTDT	Calculated start of pregnancy period. **Format**: Numeric (date9.)
PREGENDDT	Delivery date. **Format**: Numeric (date9.)
INDEXDT	Index date. This is the requester defined index date for the cohort. **Format**: Numeric (date9.)
ENR_START	Start of enrollment span where index exposure was identified. **Format**: Numeric (date9.)
ENR_END	End of enrollment span where index exposure was identified. **Format**: Numeric (date9.)
YEAR	Year of index date. **Format**: Numeric
EPISODEENDDT2	MOI of interest exposure episode end date. **Format**: Numeric (date9.)
EXPOSUREUNIT	Type of time interval that defines the exposure time period **Format:** Character(30)
EXPOSUREFROM	Start of the evaluation period for exposure of interest and will be used in combination with INTERVAL. **Format:** Numeric
EXPOSURETO	End of the evaluation period for exposure of interest and will be used in combination with INTERVAL **Format:** Numeric
OUTCOMEPOP	Population that is defined in outcomes. **Format:** character(3)
INDEXDATE	Identifies which index date to use for baseline period and propensity score risk-set creation.

Variable	Description
	Format: character(30)
OUTCOMEFROM	Start of the evaluation period for outcome with respect to the index date. **Format:** Numeric
OUTCOMETO	End of the evaluation period for outcome with respect to the index date. **Format**: Numeric
OUTCOMETOANCHOR	Identifies which index date to use for OUTCOMETO evaluation **Format**: Numeric
NUMEVENTS	Total number of HOIs observed during the exposure episode (note: days at risk [TTE] stop accumulating after the occurrence of the first HOI). **Format**: Numeric
FEVENTDT	First HOI date. **Format**: Numeric (MMDDYY10.)
ENROLL_DIFF	Difference between birth date and start of enrollment for matched infants. Will be blank if there is no matched infant **Format:** Numeric
BIRTH_ENROLL	Length of enrollment for matched infant Will be blank if there is no matched infant **Format:** Numeric
AGE	Age at index date. **Format**: Numeric
BIRTH_DATE	Birth Date of Mother. **Format**: Numeric (MMDDYY10.)
SEX	Sex. Allowable values are those in the SCDM Demographic table. **Format**: Character(1)
RACE	Race. Allowable values are those in the SCDM Demographic table. **Format**: Character(1)
HISPANIC	Hispanic. Allowable values are those in the SCDM Demographic table. **Format**: Character(1)
AGEGROUP	Age Group at index date. **Format**: Character(variable)
AGEGROUPNUM	Age Group number. **Format**: Numeric
MINAGEDT	Date on which the individual qualified to enter the cohort, based on age ranges specified. **Format**: Numeric (date9.)
MAXAGEDT	Last date the individual qualified to be in the cohort, based on age ranges specified. **Format**: Numeric (date9.)

Variable	Description
ZIP3	3 digit Zip code. **Note**: will be missing if geographic information is not requested. **Format:** Character(5)
ZIP3	Zip code. **Note**: will be missing if geographic information is not requested. **Format:** Character(5)
ZIP_UNCERTAIN	Inidicator if zip date occurs after index date **Format**: Numeric (date9.)
ZIP_DATE	Earliest date that the ZIP code is believed to be valid. **Format**: Numeric (date9.)
STATE	State associated with the zip code. **Note**: will be missing if geographic information is not requested. **Format:** Character
HHS_REG	Health and Human Services region associated with the zip code. **Note:** will be missing if geographic information is not requested. **Format:** Character
CB_REG	Census Bureau region associated with the zip code. **Note:** will be missing if geographic information is not requested. **Format:** Character
EXACTNUMVISIT	The total number of medical encounters during the requester-defined evaluation period. **Format**: Numeric
NUMAV	The total number of ambulatory encounters during the requester-defined evaluation period. **Note:** This is updates the variable name to be consistent across all datasets **Format**: Numeric
NUMOA	The total number of other ambulatory encounters during the requester-defined evaluation period. **Note:** This is updates the variable name to be consistent across all datasets **Format**: Numeric
NUMIP	The total number of inpatient encounters during the requester-defined evaluation period. **Note:** This is updates the variable name to be consistent across all datasets **Format**: Numeric
NUMIS	The total number of institutional stay encounters during the requester-defined evaluation period.

Variable	Description
	Note: This is updates the variable name to be consistent across all datasets **Format**: Numeric
NUMED	The total number of emergency department encounters during the requester-defined evaluation period. **Note:** This is updates the variable name to be consistent across all datasets **Format**: Numeric
NUMVISITS	The requester-defined stratum that the EXACTNUMVISIT value is in. **Format**: Numeric
NUMGENERIC	The total number of generic names dispensed during the requester-defined evaluation period. **Format**: Numeric
NUMCLASS	The total number of unique class names dispensed during the requester-defined evaluation period. **Format**: Numeric
NUMRX	The total number of dispensings during the requester-defined evaluation period. **Format**: Numeric
COMBINED_SCORE	The comorbidity score requester-defined category. **Format**: Character(3)
COMBINED_SCORE_NUM	The comorbidity score value. **Format**: Numeric
CCIELIXGRP	The pre-index comorbidity score requester-defined category. **Format**: Character(10)
RAWDISP	Total number of raw code counts associated with the MOI exposure episode **Format**: Numeric
ADJUSTEDDISP	Total number of adjusted code counts associated with the MOI exposure episode **Format**: Numeric
DELEPISODEIND	Indicates delivery episodes (0/1) **Format**: Numeric
DELNUM	Delivery Number per PATID **Format**: Numeric
DRUGUTILFROM	Start of drug utilization period **Format**: Numeric
DRUGUTILTO	End of drug utilization period **Format**: Numeric
HAS3TRIM	Indicates if pregnancy episode reaches 3rd trimester

Variable	Description
	Format: Numeric
INDEXDT2	MOI episode index date. Will be blank for unexposed cohort **Format**: Numeric (MMDDYY10.)
LASTLOOKFOLLOWED	Set to missing for Type 4 **Format**: Numeric
MOINAME	Name of MOI episode for cohort **Format:** Character
MILID	ID number for exposure/comparator episode **Format:** Numeric
PREPOSTIND	Indicator for type of Preterm/Postterm code (PRETERM, POSTTERM, NONE) **Format**: Character (10)
PREDURCODE	Prioritygroup1/Prioritygroup2 code, if any **Format**: Character (11)
EXP_PRE	Any exposure during the pre-pregnancy period **Format:** Numeric
EXP_T1	Any exposure during the 1st trimester **Format:** Numeric
EXP_T2	Any exposure during the 2nd trimester **Format:** Numeric
EXP_T3	Any exposure during the 3rd trimester **Format:** Numeric
ADJUSTEDDISP_PRE	Number of adjusted code counts during the pre-pregnancy period **Format:** Numeric
ADJUSTEDDISP_T1	Number of adjusted code counts during the 1st trimester **Format:** Numeric
ADJUSTEDDISP_T2	Number of adjusted code counts during the 2nd trimester **Format:** Numeric
ADJUSTEDDISP_T3	Number of adjusted code counts during the 3rd trimester **Format:** Numeric
BIRTH_TYPE	Birth Type. **Format:** Numeric
CBIRTH_DATE	Linked child birth date **Format**: Numeric (MMDDYY10.)
CENR_END	Linked child enrollment end date **Format**: Numeric (MMDDYY10.)
CENR_START	Linked child enrollment start date **Format**: Numeric (MMDDYY10.)

Variable	Description
CSEX	Linked child sex **Format:** Character
GA_BIRTH	Gestational week at birth **Format:** Numeric
GA_FIRST	Gestational week of first exposure **Format:** Numeric
GESTWKEND	Gestational week at end of MOI episode **Format:** Numeric
INDEXDT_EXP	Date of exposure **Format**: Numeric (MMDDYY10.)
LINKEDID	Child PATID **Format:** Character
MATCHMETHOD	Match Method **Format:** Character

Table 97 contains specifications for the [*RUNID*] _mstr output table for Type 5 analyses. This output table contains one row per scenario (GROUP), patient (PATID), and episode (EPISODENUM). This table includes the index episodes, and all subsequent episodes per patient.

Table 97. [RUNID] _mstr Output for Type 5 Analyses

Variable	Description
StudyStartDate	**Details:** query period start date specified in request. **Format**: Numeric (date9.)
STUDYENDDATE	**Details:** query period end date specified in request. **Format**: Numeric (date9.)
Group	**Details:** standardized name used to differentiate cohorts. **Format**: Character(40)
Patid	**Details:** unique patient identifier in SCDM **Format**: Character (variable)
Birth_date	**Details:** patient birth date in SCDM. **Format**: Numeric (MMDDYY10.)
Sex	**Details:** patient sex in SCDM. **Format**: Character(1)
AGEGROUP	**Details:** patient age group (age calculated at index date) **Format**: Character (variable)
RACE	**Details:** patient race in SCDM. **Format**: Character (1)
HISPANIC	**Details:** patient Hispanic value in SCDM. **Format**: Character (1)

Variable	Description
MinAgeDate	**Details:** first date the patient qualifies for cohort inclusion, based on specified AGESTRAT parameters. **Format**: Numeric (date9.)
MaxAgeDate	**Details:** last date the patient qualifies for cohort inclusion, based on specified AGESTRAT parameters. **Format**: Numeric (date9.)
IndexDt	**Details**: start date of the first exposure episode. **Format**: Numeric (date9.)
YEAR	**Details:** Year of index date. **Format**: Numeric (8)
EpisodeNUM	**Details:** episode number, in sequential order (starting at 1), by PatID. EPISODENUM = 1 is the index date defining episode. **Format**: Numeric (4)
EpisodeStartDt	**Details:** start date for the episode denoted by EPISODENUM. Will equal INDEXDATE for EPISODENUM=1. **Format**: Numeric (date9.)
EpisodeEndDt	**Details:** end date for the episode denoted by EPISODENUM. **Format**: Numeric (date9.)
CUMEPISODELENGTH	**Details:** cumulative time at risk. **Format**: Numeric (8)
EPISODELENGTH	**Details:** time at risk. **Format**: Numeric (8)
AdjustedDisp	Number of outpatient pharmacy dispensings used to define the index date. When 2 or more dispensings with the same RxDate are stockpiled, this metric will count as 1 dispensing **Format**: Numeric (8)
RawDisp	Number of outpatient pharmacy dispensings used to define the index date. When 2 or more dispensings with the same RxDate are stockpiled, this metric will count as 2 (or more) dispensings **Format**: Numeric (8)
TOTRXSUP	**Details:** days supply associated with all dispensings in the episode denoted by EPISODENUM. **Format**: Numeric (8)
EPISODE_Cens_elig	**Details:** 0/1 indicator denoting if episode was censored due to disenrollment. **Format**: Numeric (4)
EPISODE_Cens_dth	**Details:** 0/1 indicator denoting if episode was censored due to death. **Format**: Numeric (4)

Variable	Description
EPISODE_Cens_dpend	**Details:** 0/1 indicator denoting if episode was censored due to DP data end date. **Format**: Numeric (4)
EPISODE_Cens_qryend	**Details:** 0/1 indicator denoting if episode was censored due to study end date. **Format**: Numeric (4)
EPISODE_Cens_episend	**Details:** 0/1 indicator denoting if episode was censored due to end of exposure. **Format**: Numeric (4)
EPISODE_Cens_spec	**Details:** 0/1 indicator denoting if episode was censored due to requester-specified censoring criteria. **Format**: Numeric (4)
MNTSFROMSTART	**Details:** Number of months from episode start date. MNTSFROMSTART = 1 indicates the same month as the episode start date. **Format**: Numeric (4)
GAPNUM	**Details:** Gap number (i.e., first gap=1, second gap=2, etc.). **Format**: Numeric (4)
GAPLENGTH	**Details:** Episode Gap. If EPISODENUM = 1, the GAPLENGTH is the number of days between the end of EPISODENUM 1 and EPISODENUM 2. If EPISODENUM = n, the GAPLENGTH is the number of days between the end of EPISODENUM n and EPISODENUM n+1. **For the patient's last episode, GAPNUM should be set to 999, and the gap should represent the number days from the end of the last episode to enrollment end/death/dp end date/query end date.** **Format**: Numeric (8)
EXACTNUMVISIT	**Details:** The total number of medical encounters during the requester-defined evaluation period. **Format**: Numeric (4)
NUMVISITS_AV	**Details:** The total number of ambulatory encounters during the requester-defined evaluation period. **Format**: Numeric (4)
NUMVISITS_OA	**Details:** The total number of other ambulatory encounters during the requester-defined evaluation period. **Format**: Numeric (4)
NUMVISITS_IP	**Details:** The total number of inpatient encounters during the requester-defined evaluation period. **Format**: Numeric (4)
NUMVISITS_IS	**Details:** The total number of institutional stay encounters during the requester-defined evaluation period. **Format**: Numeric (4)

Variable	Description
NUMVISITS_ED	**Details:** The total number of emergency department encounters during the requester-defined evaluation period. **Format**: Numeric (4)
NUMVISITS	**Details:** The requester-defined stratum that the EXACTNUMVISIT value is in. **Format**: Numeric (4)
NUMGENERIC	**Details:** The total number of generic names dispensed during the requester-defined evaluation period. **Format**: Numeric (4)
NumClass	**Details:** The total number of unique class names dispensed during the requester-defined evaluation period. **Format**: Numeric (4)
NUMRX	**Details:** The total number of dispensings during the requester-defined evaluation period. **Format**: Numeric (4)
COMBINED_SCORE	**Details:** The comorbidity score requester-defined category. **Format**: Character(3)
COMBINED_SCORE_NUM	The comorbidity score value. **Format**: Numeric (4)
CCIELIXGRP	**Details:** The pre-index comorbidity score requester-defined category. **Format**: Character(10)
DISTINDEXEXP	Underscore-concatenated list of distributed index identifiers associated with exposure. **Format:** Character(250)
DISTINDEXHOI	Underscore-concatenated list of distributed index identifiers associated with HOI. **Format:** Character(250)

Table 98 contains specifications for the [RUNID]_mstr output table for Type 6 analyses.

Table 98. [RUNID]_mstr OUTPUT for Type 6 Analyses

Variable	Description
PATID	**Details:** unique patient identifier in SCDM. **Format**: Character(variable)
INDEXDT	**Details:** Start date of first product treatment episode. **Format**: Numeric (date9.)
PRODUCTMARKETINGDATE	**Details:** user-specified start marketing date for product. **Format**: Numeric (date9.)
OTHERPRODUCTDATE	**Details:** user-specified other product-related date. **Format**: Numeric (date9.)
ENR_START	**Details:** Enrollment Start Date. **Format**: Numeric (date9.)
ENR_END	**Details:** Enrollment End Date. **Format**: Numeric (date9.)
UPTAKEDATE	**Details:** which Date field is used for production uptake duration computations. **Format**: Numeric (date9.)
PRODUCTAPPROVALDATE	**Details:** user-specified approval date for product. **Format**: Numeric (date9.)
DEATHDT	**Details**: Date on which the individual died, if any. (value will be set to missing if the individual did not die) **Format**: Numeric (date9.)
EPISODEENDDT	**Details:** end date for the product treatment episode **Format**: Numeric (date9.)
INCEXL	**Details:** 1/0 indicator for inclusion/exclusion. **Format**: Numeric (8)
EPISODELENGTH	**Details:** Length of product treatment episode **Format**: Numeric (8)
GROUP	**Details:** standardized name used to differentiate cohorts. **Format**: Character(40)
TYPE	**Details:** Indicates the type of analysis performed in the CIDA tool (type = 6) **Format**: Numeric (8)
YEAR	**Details:** year of the product treatment episode. **Format**: Numeric (8)
CENS_ELIG	**Details:** 0/1 indicator denoting if episode was censored due to disenrollment.

Variable	Description
	Format: Numeric (8)
CENS_DTH	**Details:** 0/1 indicator denoting if episode was censored due to death. **Format**: Numeric (8)
CENS_DPEND	**Details:** 0/1 indicator denoting if episode was censored due to DP data end date. **Format**: Numeric (8)
CENS_QRYEND	**Details:** 0/1 indicator denoting if episode was censored due to study end date. **Format**: Numeric (8)
CENS_EPISEND	**Details:** 0/1 indicator denoting if episode was censored due to end of product treatment episode. **Format**: Numeric (8)
EPISODEENDDT_CENSOR	**Details:** Censored Date **Format**: Numeric (date9.)
TOTRXSUP	**Details:** days supply associated with all dispensings in that episode **Format**: Numeric (8)
TOTRXAMT	**Details:** Amount supplied value associated with all dispensings in the given episode. **Format**: Numeric (8)
AdjustedDisp	**Details:** Number of outpatient pharmacy dispensings during the episode. Counts same day stockpiled dispensings as 1 **Format**: Numeric (8)
RawDisp	**Details:** Number of outpatient pharmacy dispensings during the episode. Counts each same day stockpilied dispensing **Format**: Numeric (8)
BIRTH_DATE	**Details:** patient birth date in SCDM. **Format**: Numeric (MMDDYY10.)
SEX	**Details:** patient sex in SCDM. **Format**: Character(1)
RACE	**Details:** patient race in SCDM. **Format**: Character(1)
AGE	**Details**: Age at index date. **Format**: Character(variable)
ZIP	**Details**: Zip code. **Note**: will be missing if geographic information is not requested. **Format:** Character(5)

Variable	Description
ZIP_DATE	**Details**: Earliest date that the ZIP code is believed to be valid. **Format**: Numeric (date9.)
AGEGROUP	**Details**: Age Group at index date. **Format**: Character(variable)
MINAGEDATE	**Details**: Date on which the individual qualified to enter the cohort, based on age ranges specified. **Format**: Numeric (date9.)
MAXAGEDATE	**Details**: Last date the individual qualified to be in the cohort, based on age ranges specified. **Format**: Numeric (date9.)
STATE	**Details**: State associated with the zip code. **Note**: will be missing if geographic information is not requested. **Format:** Character
PREVINC	**Details:** Indicator if the episode is prevalent or incident. **Format:** Character
COMPUTEDSTARTMARKETINGDATE	**Details:** Date associated with product start marketing date. **Format**: Numeric (date9.)
CumEpisodeLength	**Details:** Sum of all tte for each PATID. **Format**: Numeric (8)

b) [*RUNID*]_numcounts.sas7bdat

The [*RUNID*] _numcounts output table is the source dataset for cohort metrics for the T#_CIDA table in the *msoc* output folder. The [*RUNID*] _numcounts output table contents depend on the type of cohort identification strategy used. Table 99Table 99 describes this output table for Type 1 analyses.

Table 99. [RUNID] _numcounts Output for Type 1 Analyses

Variable	Description
GROUP	Cohort name. **Format:** Character(40)
TYPE	Identifies the type of analysis performed in the CIDA tool (=1). **Format:** Numeric
NPTS	Number of individuals. **Format:** Numeric
EPISODES	Number of exposure episodes. **Format:** Numeric
DISPENSINGS	Number of dispensing used to define index dates. **Format:** Numeric
DAYSUPP	Total days supply associated with dispensing used to define index dates.

Variable	Description
	Format: Numeric
AMTSUPP	Total supplied amount associated with dispensing used to define index dates. **Format:** Numeric
EPS_WEVENTS	0 for Type 1 analyses. **Format:** Numeric
ALL_EVENTS	0 for Type 1 analyses. **Format:** Numeric
TTE	0 for Type 1 analyses. **Format:** Numeric
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. **Format:** Character(3)
YEAR	Year of index date. **Format:** Numeric
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(1)
RACE	Race. Allowable values are those in the SCDM. **Format**: Character(1)
HISPANIC	Hispanic. Allowable values are those in the SCDM. **Format**: Character(1)
AGEGROUP	Age Groups. Categories are requester-defined. **Format:** Character(variable)
AGEGROUPNUM	Age Group sorting order indicator. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)

Variable	Description
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date N: zip date occurred before the index date **Format:** Character(1)
COVAR*n*	Requester-defined covariate **Note 1:** Will equal STUDYNAME specified in the Covariate file **Format**: Character (50)

Table 100 contains specifications for the [*RUNID*] _numcounts output table for Type 2 analyses.

Table 100. [RUNID]_numcounts Output for Type 2 Analyses

Variable	Description
GROUP	Cohort name. **Format:** Character(40)
TYPE	Identifies the type of analysis performed in the CIDA tool (=2). **Format:** Numeric
NPTS	Number of individuals. **Format:** Numeric
EPISODES	Number of exposure episodes. **Format:** Numeric
ADJUSTEDCODECOUNT	Number of dispensings used to define index date defining records. **Format:** Numeric
RAWCODECOUNT	Number of dispensings used to define index date defining records. **Format:** Numeric
DAYSUPP	Total days supply associated with dispensing used to create exposure episodes. **Format:** Numeric
AMTSUPP	Total supplied amount associated with dispensing used to create exposure episodes. **Format:** Numeric
EPS_WEVENTS	Number of exposure episodes with an HOI. **Format:** Numeric
ALL_EVENTS	Total number of HOIs in all exposure episodes. For characterization purposes only. Days-at-risk stop accumulating after the first HOI during an exposure episode. ALL_EVENTS/TTE should never be calculated. This variable value just reports the number of times during treatment episodes that the HOI definition was met. **Format:** Numeric
TTE	Days at-risk. **Format:** Numeric

Variable	Description
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. **Format:** Character(3)
YEAR	Year of index date. **Format:** Numeric
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(1)
RACE	Race. Allowable values are those in the SCDM. **Format**: Character(1)
HISPANIC	Hispanic. Allowable values are those in the SCDM. **Format**: Character(1)
AGEGROUP	Age Groups. Categories are requester-defined. **Format:** Character(variable)
AGEGROUPNUM	Age Group sorting order indicator. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date N: zip date occurred before the index date **Format:** Character(1)
COVAR*n*	Requester-defined covariate **Note 1:** Will equal STUDYNAME specified in the Covariate file **Format**: Character (50)

Table 101 contains specifications for the [*RUNID*] _numcounts output table for Type 3 analyses.

Table 101. [RUNID] _numcounts Output for Type 3 Analyses

Variable	Description
GROUP	Cohort name. **Format:** Character(40)
TYPE	Identifies the type of analysis performed in the CIDA tool (=3). **Format:** Numeric
LEVEL	Stratification identifier. **Format:** Character(3)
YEAR	Year of index date. **Format:** Numeric
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(1)
RACE	Race. Allowable values are those in the SCDM. **Format**: Character(1)
HISPANIC	Hispanic. Allowable values are those in the SCDM. **Format**: Character(1)
AGEGROUP	Age Groups. Categories are requester-defined. Age defined at index date. **Format:** Character (variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date N: zip date occurred before the index date **Format:** Character(1)

Variable	Description
COVAR*n*	Requester-defined covariate **Note 1:** Will equal STUDYNAME specified in the Covariate file **Format**: Character (50)
TTE_VALUE	Stratification variable. All available time to event values (e.g., -2 -1, 0, 1, 2, 3, etc.). Blank TTE_VALUE may be used to characterize patients in the exposure cohort only. If an HOI is observed on the day of exposure, TTE=0 (i.e., exposure date is day 0). **Format:** Numeric
NPTS_EXPOSURE	Number of patients identified in the exposure cohort. **Format:** Numeric
EPISODES_EXPOSURE	Number of index dates (exposure episodes) identified for all members in the exposure cohort. Relevant for requests that allow more than one exposure episode per patient. **Format:** Numeric
NPTS_CENSOR_ELIG	Number of patients excluded from the analysis cohort due to insufficient post-exposure continuous enrollment during the risk and control windows. If death and disenrollment occur on the same day, censoring will be attributed to death. **Format:** Numeric
NPTS_CENSOR_DTH	Number of patients excluded from the analysis cohort due to evidence of death during the risk or control windows. If death and disenrollment occur on the same day, censoring will be attributed to death. **Format:** Numeric
NPTS_CENSOR_NOEVENTS	Number of patients excluded from the analysis cohort due to no identified events during either the risk or control windows. **Format:** Numeric
EPISODES_CENSOR_ELIG	Number of exposure episodes excluded from the analysis cohort due to insufficient post-exposure continuous enrollment during the risk and control windows. If death and disenrollment occur on the same day, censoring will be attributed to death. **Format:** Numeric
EPISODES_CENSOR_DTH	Number of exposure episodes excluded from the analysis cohort due to evidence of death during the risk or control windows. If death and disenrollment occur on the same day, censoring will be attributed to death. **Format:** Numeric
EPISODES_CENSOR_NOEVENTS	Number of exposure episodes excluded from the analysis cohort due to no identified events during either the risk or control windows. **Format:** Numeric

Variable	Description
NPTS_ANALYSIS	Number of patients identified in the analytic cohort. **Format:** Numeric
EPISODES_ANALYSIS	Number of index dates (exposure episodes) identified for all members in the analytic cohort. Relevant for requests that allow more than one exposure episode per patient. **Format:** Numeric
EVENTS_ANALYSIS_RISK	Number of events identified in the risk window for patients in the analytic cohort. **Format:** Numeric
EVENTS_ANALYSIS_CTRL	Number of events identified in the control window for patients in the analytic cohort. **Format:** Numeric
MINDAYS_EVENT_ANALYSIS	The minimum number of days from exposure to event (event date – exposure date +1 for post-exposure events; exposure date- event date +1 for pre-exposure events). Relevant for members of the analytic cohort only. **Format:** Numeric
MAXDAYS_EVENT_ANALYSIS	The maximum number of days from exposure to event (event date – exposure date +1 for post-exposure events; exposure date- event date +1 for pre-exposure events). Relevant for members of the analytic cohort only. **Format:** Numeric
MINDAYS_POSTENR_EXPOSURE	The minimum number of days of post-exposure enrollment for the exposure cohort (enrollment end – exposure date). **Format:** Numeric
MAXDAYS_POSTENR_EXPOSURE	The maximum number of days of post-exposure enrollment for the exposure cohort (enrollment end – exposure date). **Format:** Numeric

Table 102 contains specifications for the *[RUNID]* _mstr_censor output table for Type 3 analyses. This output file is similar to *[RUNID]* _mstr output table for Type 3 analyses, except it 1) includes information for patients that do not meet post-exposure enrollment requirements and 2) includes an additional variable "TTC" for time-to-censor.

Table 102. [RUNID] _mstr_censor Output for Type 3 Analyses

Variable	Description
GROUP	Cohort name. **Format**: Character(30)
TYPE	Identifies the type of analysis performed in the CIDA tool (=3). **Format**: Numeric
PATID	Individual identifier. **Format**: Character(variable)
INDEXDT	Index date. **Format**: Numeric (date9.)
ENR_START	Start of enrollment span where index exposure was identified. **Format**: Numeric (date9.)
ENR_END	End of enrollment span where index exposure was identified. **Format**: Numeric (date9.)
YEAR	Year of index date. **Format**: Numeric
FEVENTDT	HOI date. **Format**: Numeric (date9.)
TTC	Number of days from index to censoring criterion. Patients may be censored due to disenrollment or death. **Format**: Numeric
AGE	Age at index date. **Format**: Numeric
BIRTH_DATE	Birth Date. **Format**: Numeric (MMDDYY10.)
SEX	Sex. Allowable values are those in the SCDM Demographic table. **Format**: Character(1)
RACE	Race. Allowable values are those in the SCDM Demographic table. **Format**: Character(1)
HISPANIC	Hispanic. Allowable values are those in the SCDM Demographic table. **Format**: Character(1)
AGEGROUP	Age Group at index date. **Format**: Character(variable)

Variable	Description
MINENROLHOIMET	Indicates if minimum enrollment requirements to determine HOI incidence are met. Will equal 1 if requirement is met; will be set to missing when enrollment is sufficient to determine exposure incidence but insufficient to determine HOI incidence (when a window start date is prior to enrollment start date). **Format**: Numeric
WINOFEVENT	Indicates whether the HOI occurred during the risk or control window. Valid values include "RISK" and "CTRL". **Format**: Character(4)
PT_ANALYSIS_EPISODE_NUM	Indicates the episode number for a patient in the analysis cohort. When T3COHORTDEF=02 a patient can have multiple episodes. This indicator can take any integer value. **Format**: Numeric
PT_EXPOSURE_EPISODE_NUM	Indicates the episode number for a patient in the exposure cohort. When T3COHORTDEF=02 a patient can have multiple episodes. This indicator can take any integer value. **Format**: Numeric
CENSORDEATH	Indicates if exposure episode was censored due to death. CENSORDEATH =1 if censored; else 0. Note: if death and disenrollment occur on the same day, censoring will be attributed to death. **Format**: Numeric
CENSORENROL	Indicates if exposure episode was censored due to disenrollment. CENSORENROL =1 if censored; else 0. Note: if death and disenrollment occur on the same day, censoring will be attributed to death. **Format**: Numeric
ANALYSISCOHORT	Indicates if patient's exposure episode contributed to the analytic cohort. Will =1 for yes, =0 for no. **Format**: Numeric
DAYS_FROM_EXPOS_TO_EVENT	Number of days from the index date to the HOI date. **Format**: Numeric
DAYS_FROM_EXPOS_TO_CTRL_EVENT	Number of days from the index date to the control interval HOI date. **Format**: Numeric
DAYS_FROM_EXPOS_TO_RISK_EVENT	Number of days from the index date to the risk interval HOI date. **Format**: Numeric
POSTENR_EXPOSURE	Number of days of continuous enrollment post-exposure. Calculated as Enr_End – IndexDt +1. **Format**: Numeric

Table 103 contains specifications for the [*RUNID*] _numcounts_censor output table for Type 3 analyses. This output file is similar to [*RUNID*] _numcounts output table for Type 3 analyses, except it 1) includes information for patients that do not meet post-exposure enrollment requirements and 2) includes an additional variable "TTC_VALUE" for time-to-censor information.

Table 103. [RUNID] _numcounts_censor Output for Type 3 Analyses

Variable	Description
GROUP	Cohort name. **Format:** Character(40)
TYPE	Identifies the type of analysis performed in the CIDA tool (=3). **Format:** Numeric
LEVEL	Stratification identifier. **Format:** Character(3)
YEAR	Year of index date. **Format:** Numeric
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(1)
RACE	Race. Allowable values are those in the SCDM. **Format**: Character(1)
HISPANIC	Hispanic. Allowable values are those in the SCDM. **Format**: Character(1)
AGEGROUP	Age Groups. Categories are requester-defined. Age defined at index date. **Format:** Character (variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric
TTE_VALUE	Stratification variable. All available time to event values (e.g., -2 -1, 0, 1, 2, 3, etc.). Blank TTE_VALUE may be used to characterize patients in the exposure cohort only. If an HOI is observed on the day of exposure, TTE=0 (i.e., exposure date is day 0). **Format:** Numeric
TTC_VALUE	Stratification variable. All available time to censor values (e.g., 0, 1, 2, 3, etc.). Blank TTC_VALUE may be used to characterize patients in the exposure cohort only. If the last day of follow-up is the index date, TTC_VALUE=0. **Format:** Numeric
NPTS_EXPOSURE	Number of patients identified in the exposure cohort. **Format:** Numeric

Variable	Description
EPISODES_EXPOSURE	Number of index dates (exposure episodes) identified for all members in the exposure cohort. Relevant for requests that allow more than one exposure episode per patient. **Format:** Numeric
NPTS_CENSOR_ELIG	Number of patients excluded from the analysis cohort due to insufficient post-exposure continuous enrollment during the risk and control windows. If death and disenrollment occur on the same day, censoring will be attributed to death. **Format:** Numeric
NPTS_CENSOR_DTH	Number of patients excluded from the analysis cohort due to evidence of death during the risk or control windows. If death and disenrollment occur on the same day, censoring will be attributed to death. **Format:** Numeric
NPTS_CENSOR_NOEVENTS	Number of patients excluded from the analysis cohort due to no identified events during either the risk or control windows. **Format:** Numeric
EPISODES_CENSOR_ELIG	Number of exposure episodes excluded from the analysis cohort due to insufficient post-exposure continuous enrollment during the risk and control windows. If death and disenrollment occur on the same day, censoring will be attributed to death. **Format:** Numeric
EPISODES_CENSOR_DTH	Number of exposure episodes excluded from the analysis cohort due to evidence of death during the risk or control windows. If death and disenrollment occur on the same day, censoring will be attributed to death. **Format:** Numeric
EPISODES_CENSOR_NOEVENTS	Number of exposure episodes excluded from the analysis cohort due to no identified events during either the risk or control windows. **Format:** Numeric
NPTS_ANALYSIS	Number of patients identified in the analytic cohort. **Format:** Numeric
EPISODES_ANALYSIS	Number of index dates (exposure episodes) identified for all members in the analytic cohort. Relevant for requests that allow more than one exposure episode per patient. **Format:** Numeric
EVENTS_ANALYSIS_RISK	Number of events identified in the risk window for patients in the analytic cohort. **Format:** Numeric
EVENTS_ANALYSIS_CTRL	Number of events identified in the control window for patients in the analytic cohort.

Variable	Description
	Format: Numeric
MINDAYS_EVENT_ANALYSIS	The minimum number of days from exposure to event (event date – exposure date +1 for post-exposure events; exposure date- event date +1 for pre-exposure events). Relevant for members of the analytic cohort only. **Format:** Numeric
MAXDAYS_EVENT_ANALYSIS	The maximum number of days from exposure to event (event date – exposure date +1 for post-exposure events; exposure date- event date +1 for pre-exposure events). Relevant for members of the analytic cohort only. **Format:** Numeric
MINDAYS_POSTENR_EXPOSURE	The minimum number of days of post-exposure enrollment for the exposure cohort (enrollment end – exposure date). **Format:** Numeric
MAXDAYS_POSTENR_EXPOSURE	The maximum number of days of post-exposure enrollment for the exposure cohort (enrollment end – exposure date). **Format:** Numeric

c) [*RUNID*]_denomcounts.sas7bdat

The [*RUNID*] _denomcounts output table is the source dataset for eligible members and member-days metrics for the T1_CIDA and T2_CIDA tables in the *msoc* output folder. Table 104 describes this output.

Table 104. [RUNID]_denomcounts Output

Variable	Description
GROUP	Cohort name. **Format:** Character(40)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. **Format:** Character(3)
YEAR	Year of index date. **Format:** Numeric
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(1)
RACE	Race. Allowable values are those in the SCDM. **Format**: Character(1)
HISPANIC	Hispanic. Allowable values are those in the SCDM. **Format**: Character(1)

Variable	Description
AGEGROUPNUM	Age Group sorting order indicator. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric
DENNUMPTS	Number of eligible individuals. **Note:** For requests that will use the prospective surveillance with propensity score matched design options 1 or 2, this will be blank. **Format:** Numeric
DENNUMMEMDAYS	Number of eligible days. **Note:** For requests that will use the prospective surveillance with propensity score matched design options 1 or 2, this will be blank. **Format:** Numeric
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date N: zip date occurred before the index date **Format:** Character(1)

d) [*RUNID*]_ads_mstr_[*PERIODID*].sas7bdat

The [*RUNID*] _ads_mstr_ [*PERIODID*] output table will be generated for Types 1-5. The output file contains one record per PATID per GROUP when the T#COHORTDEF parameter is set to "01". Note that if T#COHORTDEF is set to "02" or "03" or "04" ot "05", PATIDs may appear multiple times in this output file. Table 105 describes this output table.

Table 105. [RUNID] _ads_mstr_ [PERIODID] Output

Variable	Description
GROUP	Cohort Name. **Format:** Character(40)
PATID	Individual Identifier (PatID in SDD). **Format:** Character(30)
YEAR	Year of index date. **Note 1**: For Type 4 requests, this is the year of delivery date. **Format:** Numeric
INDEXDT	Index date. **Format:** Date
LastLookFollowed	When an individual is lost to follow up (data is no longer being updated), this is the last look # the individual contributed data to. If an individual is still eligible to be followed in subsequent looks, this is set to 0. **Note 1:** Type 2 requests only. **Format:** Numeric
LastLookFollowedDt	The last date of the Look specified in LastLookFollowed. Set to missing if LastLookFollowed = 0. **Note 1:** Type 2 requests only. **Format:** Numeric (date9.)
EVENTDT	Date of HOI, if it occurred during the time period (PERIODID). This field is only populated if the HOI occurs prior to, or on the same day as, censoring (i.e. if the patient is censored prior to the HOI date, the HOI date is set to missing), and if the HOI occurred during the time period. **Note 1:** blank for Type 1 requests. **Format:** Date
AGE	Age as of index date, calculated as (INDEXDT – BIRTH_DATE)/365.25. It is not rounded to the nearest integer, but rather kept in decimal form. **Format:** Numeric
SEX	Sex as defined in SDD Demographic table. **Format:** Character(1)

Variable	Description
RACE	Race as defined in SDD Demographic table. **Format:** Character(1)
HISPANIC	Hispanic as defined in SDD Demographic table. **Format:** Character(1)
AGE_CAT	Requester-defined age category. Determined by AGESTRAT macro parameter values. **Format:** Character(9)
NUMAV	Visit count for AV encounter type. One visit allowed per encounter, per day. **Note 1**: this metric is automatically calculated using the medical utilization module (Medical Utilization File). **Format:** Numeric
NUMOA	Visit count for OA encounter type. One visit allowed per encounter, per day. **Note 1**: this metric is automatically calculated using the medical utilization module (Medical Utilization File). **Format:** Numeric
NUMIP	Visit count for IP encounter type. One visit allowed per encounter, per day. **Note 1**: this metric is automatically calculated using the medical utilization module (Medical Utilization File). **Format:** Numeric
NUMIS	Visit count for IS encounter type. One visit allowed per encounter, per day. **Note 1**: this metric is automatically calculated using the medical utilization module (Medical Utilization File). **Format:** Numeric
NUMED	Visit count for ED encounter type. One visit allowed per encounter, per day. **Note 1**: this metric is automatically calculated using the medical utilization module (Medical Utilization File). **Format:** Numeric
COMORBIDSCORE	Combined Charlson-Elixhauser Comorbidity Score (exact value). **Note 1:** this metric is calculated using the combined comorbidity score module (Comorbidity Score File). **Format:** Numeric
COVAR1-COVARN	Boolean indicating whether or not patient had covariate in the covariate evaluation window. Number of COVARN variables

Variable	Description
	determined from number of distinct STUDYNAME values in the Covariate Codes File. **Format:** Numeric
NUMGENERIC	Count of unique generic names dispensed during covariate evaluation window. **Format:** Numeric
NUMCLASS	Count of unique class names dispensed during covariate evaluation window. **Format:** Numeric
NUMRX	Count of dispensings during the covariate evaluation window. **Format:** Numeric
FOLLOWUPTIME	Number of days that the patient is followed for HOI or censoring. Variable is calculated using the earliest of 1) end of exposure period; 2) disenrollment; 3) any additional censoring criteria [defined using CIDA tool]; and 4) the HOI date. **Note 1:** Type 2 requests only. **Format:** Numeric
EVENT	Boolean indicator of whether or not patient had HOI during the time period (PERIODID). **Note 1:** Type 2 and Type 4 requests only. **Format:** Numeric
TTE	Days at-risk. **Note 1:** Type 2 requests only. **Format:** Numeric
NUMEVENTS	Total number of HOIs observed during the exposure episode. **Note 1:** Type 2 and Type 4 requests only. **Format**: Numeric
EPISODETYPE	Indicates if exposed time is requester-defined number of days (value will be ITT) or determined using dispensing days supplied (value will be EPI). **Note 1:** Type 2 requests only. **Format**: Character(3)
TIME	Variable indicating during which time period the patient was selected (*e.g.,* TIME = 1 indicates the patient was selected in PERIODID 1, even if the current PERIODID being executed is not the first). **Format:** Numeric
CENSORDEATH	Indicates if exposure episode was censored due to death. CENSORDEATH =1 if censored; else 0. Note: if death and

Variable	Description
	disenrollment occur on the same day, censoring will be attributed to death. **Note 1:** Type 3 requests only. **Format**: Numeric
CENSORENROL	Indicates if exposure episode was censored due to disenrollment. CENSORENROL =1 if censored; else 0. Note: if death and disenrollment occur on the same day, censoring will be attributed to death. **Note 1:** Type 3 requests only. **Format**: Numeric
ANALYSISCOHORT	Indicates if patient's exposure episode contributed to the analytic cohort. Will =1 for yes, =0 for no. **Note 1:** Type 3 requests only. **Format**: Numeric
DAYS_FROM_EXPOS_TO_EVENT	Number of days from the index date to the HOI date. **Note 1:** Type 3 requests only. **Format**: Numeric
DAYS_FROM_EXPOS_TO_CTRL_EVENT	Number of days from the index date to the control interval HOI date. **Note 1:** Type 3 requests only. **Format**: Numeric
DAYS_FROM_EXPOS_TO_RISK_EVENT	Number of days from the index date to the risk interval HOI date. **Note 1:** Type 3 requests only. **Format**: Numeric
DELNUM	Delivery number for the patient. This is the overall delivery number for the patient and is not confined to valid deliveries **Note 1:** Type 4 requests only. **Format**: Numeric
EPISODENUM	Episode number, in sequential order (starting at 1), by PatID. **Note 1:** Type 5 requests only. **Format**: Numeric
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)

Variable	Description
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date N: zip date occurred before the index date **Format:** Character(1)

e) [*RUNID*]_ads_mstr_MI_[*PERIODID*].sas7bdat

The [*RUNID*] _ads_mstr_ MI_[*PERIODID*] output table will be generated for Type 4 anlayses that create cohorts for further processing with the PSA tool. The output file contains one record per PATID per GROUP. Table 106 describes this output table.

Table 106. [RUNID] _ads_mstr_ [PERIODID] Output

Variable	Description
GROUP	Cohort Name. **Format:** Character(30)
PATID	Individual Identifier (PatID in SDD). **Format:** Character(30)
INDEXDT	Requester defined index date **Format:** Date
YEAR	Year of delivery date. **Format:** Numeric
EVENTDT	Date of HOI, if it occurred during the time period (PERIODID). **Format:** Date
NUMEVENTS	Total number of HOIs observed during the exposure episode. **Format**: Numeric
AGE	Age as of index date, calculated as (INDEXDT – BIRTH_DATE)/365.25. It is not rounded to the nearest integer, but rather kept in decimal form. **Format:** Numeric
SEX	Sex as defined in SDD Demographic table. **Format:** Character(1)
RACE	Race as defined in SDD Demographic table. **Format:** Character(1)
HISPANIC	Hispanic as defined in SDD Demographic table. **Format:** Character(1)
AGE_CAT	Requester-defined age category. Determined by AGESTRAT macro parameter values. **Format:** Character(9)
ENROLL_DIFF	Difference between birth date and start of enrollment for matched infants.

Variable	Description
	Note 1: Blank if there is no matched infant **Format:** Numeric
BIRTH_ENROLL	Length of enrollment for matched infant. **Note 1:** Blank if there is no matched infant **Format:** Numeric
AGEGROUPNUM	Age Group number. **Format**: Numeric
NUMAV	Visit count for AV encounter type. One visit allowed per encounter, per day. **Note 1**: this metric is automatically calculated using the medical utilization module (Medical Utilization File). **Format:** Numeric
NUMOA	Visit count for OA encounter type. One visit allowed per encounter, per day. **Note 1**: this metric is automatically calculated using the medical utilization module (Medical Utilization File). **Format:** Numeric
NUMIP	Visit count for IP encounter type. One visit allowed per encounter, per day. **Note 1**: this metric is automatically calculated using the medical utilization module (Medical Utilization File). **Format:** Numeric
NUMIS	Visit count for IS encounter type. One visit allowed per encounter, per day. **Note 1**: this metric is automatically calculated using the medical utilization module (Medical Utilization File). **Format:** Numeric
NUMED	Visit count for ED encounter type. One visit allowed per encounter, per day. **Note 1**: this metric is automatically calculated using the medical utilization module (Medical Utilization File). **Format:** Numeric
COMORBIDSCORE	Combined Charlson-Elixhauser Comorbidity Score (exact value). **Note 1:** this metric is calculated using the combined comorbidity score module (Comorbidity Score File). **Format:** Numeric
COVAR1-COVARN	Boolean indicating whether or not patient had covariate in the covariate evaluation window. Number of COVARN variables determined from number of distinct STUDYNAME values in the Covariate Codes File. **Format:** Numeric
NUMGENERIC	Count of unique generic names dispensed during covariate evaluation window.

Variable	Description
	Format: Numeric
NUMCLASS	Count of unique class names dispensed during covariate evaluation window. **Format:** Numeric
NUMRX	Count of dispensings during the covariate evaluation window. **Format:** Numeric
FOLLOWUPTIME	Will be blank. **Format:** Numeric
EVENT	Boolean indicator of whether or not patient had HOI during the time period (PERIODID). **Format:** Numeric
TIME	Variable indicating during which time period the patient was selected (*e.g.,* TIME = 1 indicates the patient was selected in PERIODID 1, even if the current PERIODID being executed is not the first). **Format:** Numeric
DELNUM	Delivery number for the patient. This is the overall delivery number for the patient and is not confined to valid deliveries **Format**: Numeric
PREPOSTIND	Indicator for type of Preterm/Postterm code (PRETERM, POSTTERM, TERM, NONE) **Format**: Character (10)
EXP_T1	Patient exposed in the first trimester. The first trimester is 0-90 days following pregnancy start. **Note 1:** Is not computed for unexposed cohort **Format**: Numeric(8)
EXP_T2	Patient exposed in the second trimester. The second trimester is 91-180 days following pregnancy start. **Note 1:** Is not computed for unexposed cohort **Format**: Numeric(8)
EXP_T3	Patient exposed in the third trimester. The third trimester is 180+ days following pregnancy start and until delivery or child birth date. **Note 1:** Is not computed for unexposed cohort **Format**: Numeric(8)
EXP_PREGPRE	The number of pregnancy episodes with exposure episodes within the pre-pregnancy period. **Note 1:** Is not computed for unexposed cohort **Format**: Numeric(8)
GA_BIRTH	Gestational age at birth (in days).

Variable	Description
	Format: Numeric(8)
GA_FIRST	Gestational age of first exposure (in weeks). **Format**: Numeric(8)
ADJUSTEDDISP_PRE	Number of adjusted code counts in pre-pregnancy period. **Format**: Numeric(8)
ADJUSTEDDISP_T1	Number of adjusted code counts in first trimester. **Format**: Numeric(8)
ADJUSTEDDISP_T2	Number of adjusted code counts in second trimester. **Format**: Numeric(8)
ADJUSTEDDISP_T3	Number of adjusted code counts in third trimester. **Format**: Numeric(8)

f) [RUNID]_mstr_concomitance.sas7bdat for concomitant use event/outcome

This output table is produced when using the concomitant use tool with a type 2 analysis. It contains one row per patient (PATID) per concomitant episode.

Table 107. [RUNID]_mstr_concomitance.sas7bdat

Variable	Description
ANALYSISGRP	**Details:** GROUP name to differentiate primary/secondary pairs. **Format:** Character (40)
PATID	Individual identifier. **Format:** Character(variable)
INDEXDT	Index date for concomitant episode. **Format**: Numeric (date9.)
EPISODEENDDT	End date for concomitant episode **Format**: Numeric (date9.)
ENR_START	Start of enrollment span where index exposure was identified. **Format**: Numeric (date9.)
ENR_END	End of enrollment span where index exposure was identified. **Format**: Numeric (date9.)
YEAR	Year of index date. **Format**: Numeric
RawDisp	Number of index defining codes **Format**: Numeric
AdjustedDisp	Number of index defining codes adjusted for codes incurred on same date. **Format**: Numeric
TOTRXSUP	Days supplied (RXSUP) value associated with the dispensing that defined the index date.

Variable	Description
	Format: Numeric
TOTRXAMT	Amount supplied (RXAMT) value associated with the dispensing that defined the index date. **Format**: Numeric
TTE	Days at risk **Format**: Numeric
AGE	Age at index date. **Format**: Numeric
BIRTH_DATE	Birth Date. **Format**: Numeric (MMDDYY10.)
SEX	Sex. Allowable values are those in the SCDM Demographic table. **Format:** Character(2)
RACE	Race. Allowable values are those in the SCDM. **Format:** Character(1)
HISPANIC	A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race. Allowable values are those in the SCDM. **Format:** Character(1)
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. **Format:** Character(1)
AGEGROUP	Age Group at index date. **Format:** Character(variable)
MINAGEDATE	Date on which the individual qualified to enter the cohort, based on age ranges specified. **Format**: Numeric (date9.)
MAXAGEDATE	Last date the individual qualified to be in the cohort, based on age ranges specified. **Format**: Numeric (date9.)
FEVENTDT	HOI date. **Format**: Numeric (MMDDYY10.)

Variable	Description
NUMEVENTS	Total number of HOIs observed during the exposure episode (note: days at risk [TTE] stop accumulating after the occurrence of the first HOI). **Format**: Numeric
EXACTNUMVISIT	The total number of medical encounters during the requester-defined evaluation period. **Format**: Numeric
NUMVISITS_AV	The total number of ambulatory encounters during the requester-defined evaluation period. **Format**: Numeric
NUMVISITS_OA	The total number of other ambulatory encounters during the requester-defined evaluation period. **Format**: Numeric
NUMVISITS_IP	The total number of inpatient encounters during the requester-defined evaluation period. **Format**: Numeric
NUMVISITS_IS	The total number of institutional stay encounters during the requester-defined evaluation period. **Format**: Numeric
NUMVISITS_ED	The total number of emergency department encounters during the requester-defined evaluation period. **Format**: Numeric
NUMVISITS	The requester-defined stratum that the EXACTNUMVISIT value is in. **Format**: Numeric
NUMGENERIC	The total number of generic names dispensed during the requester-defined evaluation period. **Format**: Numeric
NUMCLASS	The total number of unique class names dispensed during the requester-defined evaluation period. **Format**: Numeric
NUMRX	The total number of dispensings during the requester-defined evaluation period. **Format**: Numeric
COMBINED_SCORE	The comorbidity score requester-defined category. **Format**: Character(3)
COMBINED_SCORE_NUM	The comorbidity score value. **Format**: Numeric
CCIELIXGRP	The pre-index comorbidity score requester-defined category. **Format**: Character(10)

g) [RUNID]_mstr_multevent.sas7bdat for Multiple Events

This output table is produced when using the multiple events tool with a type 2 analysis. It contains one row per individual [patid] per primary episode.

Table 108. [RUNID]_mstr_multevent.sas7bdat

Variable	Description
ANALYSISGRP	**Details:** GROUP name to differentiate primary/secondary pairs. **Format:** Character (40)
PATID	Individual identifier. **Format:** Character(variable)
PRIMARYEP_NUM	Primary episode number **Format**: Numeric
INDEXDT	Index date for Primary Episode. **Format**: Numeric (date9.)
EPISODEENDDT	Primary episode end date. **Format**: Numeric (date9.)
ENR_START	Start of enrollment span where index exposure was identified. **Format**: Numeric (date9.)
ENR_END	End of enrollment span where index exposure was identified. **Format**: Numeric (date9.)
EPISODETYPE	Type of episode created (EPI, ITT, PIT) **Format:** Character(3)
YEAR	Year of index date. **Format**: Numeric
RAWCODECOUNT	Number of index defining codes. **Format**: Numeric
ADJUSTEDCODECOUNT	Number of index defining codes adjusted for codes incurred on the same date. **Format**: Numeric
TOTRXSUP	Days supplied (RXSUP) value associated with the dispensing that defined the index date. **Format**: Numeric
TOTRXAMT	Amount supplied (RXAMT) value associated with the dispensing that defined the index date. **Format**: Numeric
TTE	Length of primary episode **Format**: Numeric

Variable	Description
AGE	Age at index date. **Format**: Numeric
BIRTH_DATE	Birth Date. **Format**: Numeric (MMDDYY10.)
SEX	Sex. Allowable values are those in the SCDM Demographic table. **Format:** Character(2)
RACE	Race. Allowable values are those in the SCDM. **Format:** Character(1)
HISPANIC	A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race. Allowable values are those in the SCDM. **Format:** Character(1)
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. **Format:** Character(1)
AGEGROUP	Age Group at index date. **Format:** Character(variable)
MINAGEDATE	Date on which the individual qualified to enter the cohort, based on age ranges specified. **Format**: Numeric (date9.)
MAXAGEDATE	Last date the individual qualified to be in the cohort, based on age ranges specified. **Format**: Numeric (date9.)
MINEPISDUR	Minimum Episode duration parameter set for primary episode **Format**: Numeric
EPI_COUNT	Number of secondary episodes **Format**: Numeric
TIME_TO_EPI	Time to secondary episode (1st,2nd,3rd)

Variable	Description
	Format: Numeric
MIN_EPIGAP	Minimum episode gap among all episode gaps **Format**: Numeric
MAX_EPIGAP	Maximum episode gap among all episode gaps **Format**: Numeric
ADHERENCE	Indicator if primary episode meets adherence **Format**: Numeric
ADHERENCE_#	Indicator if patient meets ADHERENCE_ID adherence (one variable per adherence criteria) **Format**: Numeric
ADHERENCE_PAT	Indicator if patient meets adherence **Format**: Numeric
INDENOMINATOR	For ITS analysis, if primary episode meets all criteria after truncation **Format**: Numeric
ITS_ADH_TRUNCATED	For ITS analysis, indicator if primary episode meets adherence, after truncation **Format**: Numeric
ADHERENCE_#_TRUNCATED	Indicator if primary episode meets ADHERENCE_ID adherence, after truncation **Format**: Character (1)
UTILIZATION VARS, COMORBIDSCORE, COVARS	For ITS analysis, utilization variables, combined comorbidity score, and covariate indicators **Format**: Numeric
INTERVENTIONPERIOD	Indicator for ITS period **Format**: Numeric

h) [RUNID]_mstr_overlap.sas7bdat

This output table is produced when using the Overlap Tool with a Type 2 analysis and contains one row per patient (PATID) per primary episode.

Table 109. [RUNID]_mstr_overlap.sas7bdat

Variable	Description
ANALYSISGRP	**Details:** GROUP name to differentiate primary/secondary pairs. **Format:** Character (40)
PATID	Individual identifier. **Format:** Character(variable)
PRIMARYEP_NUM	Primary episode number **Format**: Numeric
INDEXDT	Index date for Primary Episode. **Format**: Numeric (date9.)
EPISODEENDDT	Primary episode end date. **Format**: Numeric (date9.)
ENR_START	Start of enrollment span where index exposure was identified. **Format**: Numeric (date9.)
ENR_END	End of enrollment span where index exposure was identified. **Format**: Numeric (date9.)
EPISODETYPE	Type of episode created (EPI, ITT, PIT) **Format:** Character(3)
YEAR	Year of index date. **Format**: Numeric
RAWCODECOUNT1	Number of index defining codes for primary episodes. **Format**: Numeric
ADJUSTEDCODECOUNT1	Number of index defining codes for primary episodes adjusted for codes incurred on same date. **Format**: Numeric
TOTRXSUP1	Days supplied (RXSUP) value associated with the dispensing that defined the index date. **Format**: Numeric
TOTRXAMT1	Amount supplied (RXAMT) value associated with the dispensing that defined the index date. **Format**: Numeric
TTE_CAT	Primary episode length categories **Format**: Numeric
TTE	Length of primary episode **Format**: Numeric

Variable	Description
RAWCODECOUNT2	Number of index defining codes for secondary episodes. **Format**: Numeric
ADJUSTEDCODECOUNT2	Number of index defining codes for secondary episodes adjusted for codes incurred on same date. **Format**: Numeric
TOTRXSUP2	Days supplied (RXSUP) value associated with the dispensing that defined the index date for the secondary episode. **Format**: Numeric
TOTRXAMT2	Amount supplied (RXAMT) value associated with the dispensing that defined the index date for the secondary episode. **Format**: Numeric
AGE	Age at index date. **Format**: Numeric
BIRTH_DATE	Birth Date. **Format**: Numeric (MMDDYY10.)
SEX	Sex. Allowable values are those in the SCDM Demographic table. **Format:** Character(2)
RACE	Race. Allowable values are those in the SCDM. **Format:** Character(1)
HISPANIC	A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race. Allowable values are those in the SCDM. **Format:** Character(1)
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. **Format:** Character(1)
AGEGROUP	Age Group at index date. **Format:** Character(variable)
MINAGEDATE	Date on which the individual qualified to enter the cohort, based on age ranges specified. **Format**: Numeric (date9.)

Variable	Description
MAXAGEDATE	Last date the individual qualified to be in the cohort, based on age ranges specified. **Format**: Numeric (date9.)
MINEPISDUR	Minimum Episode duration parameter set for primary episode **Format**: Numeric
EPI_COUNT	Number of secondary episodes **Format**: Numeric
OBS_START	Start of observation window for evaluating overlap **Format**: Numeric (date9.)
OBS_END	End of observation window for evaluating overlap **Format**: Numeric (date9.)
TOTAL_DAYS_OVERLAP	Total number of days of overlap for the primary episode **Format**: Numeric
CUTOFF_CAT (_XX_XX)	Indicator for whether the overlap meets the requester specified range **Format**: Numeric
ADHERENCE	Indicator if primary episode meets adherence **Format**: Numeric
ADHERENCE_#	Indicator if patient meets ADHERENCE_ID adherence (one variable per adherence criteria) **Format**: Numeric
ADHERENCE_PAT	Indicator if patient meets adherence **Format**: Numeric
INDENOMINATOR	For ITS analysis, if primary episode meets all criteria after truncation **Format**: Numeric
ITS_ADHERENCE_TRUNCATED	For ITS analysis, indicator if primary episode meets adherence, after truncation **Format**: Numeric
UTILIZATION VARS, COMORBIDSCORE, COVARS	For ITS analysis, utilization variables, combined comorbidity score, and covariate indicators **Format**: Numeric
ADHERENCE_#_TRUNCATED	Indicator if primary episode meets ADHERENCE_ID adherence, after truncation **Format**: Character (1)
INTERVENTIONPERIOD	Indicator for ITS period **Format**: Numeric

i) [RUNID]_concepi_mstr_[T4COHORT]_tri.sas7bdat

This output table is produced when using a type 4 analysis. T4COHORT takes the value of Preg or NoPreg for the pregnant cohort and the non-pregnant cohort, respectively. This table contains trimester statisitics for each delivery and medical product of interest. It contains one row per individual [patid] per MOI episode.

Variable	Description
PATID	Individual Identifier (PatID in SDD). **Format:** Character(30)
INDEXDT	Calculated start of pregnancy **Format:** Date
EPISODEENDDT	End of Pregnancy **Format:** Date
DELNUM	Delivery number for the patient. This is the overall delivery number for the patient and is not confined to valid deliveries **Format**: Numeric
GROUP	Cohort Name. **Format:** Character(30)
MOINAME	MOI Name **Format:** Character(30)
INDEXDT2	MOI Index date **Format:** Date
ADATE	Delivery date **Format:** Date
PREGDURCODE	Prioritygroup1/Prioritygroup2 code, if any **Format**: Character (11)
PSADATE	Preterm/Postterm code date, if any **Format:** Date (MMDDYY10.)
HADPRIORITY1	Indicator if pregnancy episode has a priority 1 birth code **Format**: Numeric
HADPRIORITY2	Indicator if pregnancy episode has a priority 2 birth code **Format**: Numeric
HADNONE	Indicator if pregnancy episode has neither a priority 1 or a priority 2 birth code **Format**: Numeric
HAD3TRIM	Pregnancy episode has 3 trimesters **Format**: Numeric
AGEGROUP	Age Group at delivery date. **Format**: Character(variable)

Variable	Description
YEAR	Year of delivery date. **Format**: Numeric
MONTH	Month of delivery date. **Format**: Numeric
AGE	Age at index date. **Format**: Numeric
AGEGROUPNUM	Age group indicator **Format**: Numeric
EPISODESTARTDATE	Pregnancy episode start date **Format**: Date (date9.)
PREPOSTIND	Indicator for type of Preterm/Postterm code (PRETERM, POSTTERM, TERM, NONE) **Format**: Character (10)
GESTAGE	Gestational age reached **Format**: Numeric
EPISODEENDDT2	End date for MOI episode **Format**: Date
ADJUSTEDDISP	Adjusted Dispensings for MOI episode **Format**: Numeric
RAWDISP	Raw Dispensings for MOI episode **Format**: Numeric
ENR_START	Start of enrollment span where index exposure was identified. **Format**: Date (date9.)
ENR_END	End of enrollment span where index exposure was identified. **Format**: Date (date9.)
ENCTYPE	Encounter Type **Format**: Character(1)
SEX	Sex **Format**: Character(1)
RACE	Race **Format**: Character(1)
HISPANIC	Hispanic **Format**: Character(1)
ZIP3	3 digit zip code **Format**: Numeric
ZIP_Uncertain	Y / N variable. **Format:** Character(1)

Variable	Description
STATE	State **Format:** Character(2)
CB_REG	CB region **Format:** Character(7)
HHS_REG	HHS region **Format:** Character(2)
PDX	PDX **Format:** Character(1)
USEPREADJ	Adjusted dispensing in pre-period indicator **Format**: Numeric
ANYT1ADJ	Adjusted dispensing in 1st trimester indicator **Format**: Numeric
ANYT2ADJ	Adjusted dispensing in 2nd trimester indicator **Format**: Numeric
ANYT3ADJ	Adjusted dispensing in 3rd trimester indicator **Format**: Numeric
ANYTADJ	Adjusted dispensing in any trimester indicator **Format**: Numeric
ONLYT1ADJ	Adjusted dispensing in only 1st trimester indicator **Format**: Numeric
ONLYT2ADJ	Adjusted dispensing in only 2nd trimester indicator **Format**: Numeric
ONLYT3ADJ	Adjusted dispensing in only 3rd trimester indicator **Format**: Numeric
USEPRERAW	Raw dispensing in pre-period indicator **Format**: Numeric
ANYT1RAW	Raw dispensing in 1st trimester indicator **Format**: Numeric
ANYT2RAW	Raw dispensing in 2nd trimester indicator **Format**: Numeric
ANYT3RAW	Raw dispensing in 3rd trimester indicator **Format**: Numeric
ANYTRAW	Raw dispensing in any trimester indicator **Format**: Numeric
ONLYT1RAW	Raw dispensing in only 1st trimester indicator **Format**: Numeric
ONLYT2RAW	Raw dispensing in only 2nd trimester indicator

Variable	Description
	Format: Numeric
ONLYT3RAW	Raw dispensing in only 3rd trimester indicator **Format**: Numeric
SUMADJCNTPRE	Total number of adjusted dispensings in pre-period **Format**: Numeric
SUMADJCNTANYT1	Total number of adjusted dispensings in 1st trimester **Format**: Numeric
SUMADJCNTANYT2	Total number of adjusted dispensings in 2nd trimester **Format**: Numeric
SUMADJCNTANYT3	Total number of adjusted dispensings in 3rd trimester **Format**: Numeric
SUMADJCNTANYT	Total number of adjusted dispensings in any trimester **Format**: Numeric
SUMADJCNTONLYT1	Total number of adjusted dispensings in only 1st trimester **Format**: Numeric
SUMADJCNTONLYT2	Total number of adjusted dispensings in only 2nd trimester **Format**: Numeric
SUMADJCNTONLYT3	Total number of adjusted dispensings in only 3rd trimester **Format**: Numeric
SUMRAWCNTPRE	Total number of raw dispensings in pre-period **Format**: Numeric
SUMRAWCNTANYT1	Total number of raw dispensings in 1st trimester **Format**: Numeric
SUMRAWCNTANYT2	Total number of raw dispensings in 2nd trimester **Format**: Numeric
SUMRAWCNTANYT3	Total number of raw dispensings in 3rd trimester **Format**: Numeric
SUMRAWCNTANYT	Total number of raw dispensings in any trimester **Format**: Numeric
SUMRAWCNTONLYT1	Total number of raw dispensings in only 1st trimester **Format**: Numeric
SUMRAWCNTONLYT2	Total number of raw dispensings in only 2nd trimester **Format**: Numeric
SUMRAWCNTONLYT3	Total number of raw dispensings in only 3rd trimester **Format**: Numeric
USEPRE	Indicator for MOI use in pre-pregnancy period **Format**: Numeric

Variable	Description
ANYT1	Indicator for MOI use in 1st trimester **Format**: Numeric
ANYT2	Indicator for MOI use in 2nd trimester **Format**: Numeric
ANYT3	Indicator for MOI use in 3rd trimester **Format**: Numeric
ANY	Indicator for MOI use in any trimester **Format**: Numeric
ONLYT1	Indicator for MOI use in only 1st trimester **Format**: Numeric
ONLYT2	Indicator for MOI use in only 2nd trimester **Format**: Numeric
ONLYT3	Indicator for MOI use in only 3rd trimester **Format**: Numeric
ALLT	Indicator for Moi use in all 3 trimesters **Format**: Numeric
TRIMONLYWASHMET	Trimester washout criteria met (set to 0) **Format**: Numeric
DELEPISODEIND	Delivery episode indicator **Format**: Numeric
ANYT	Indicator for MOI use in any trimester **Format**: Numeric
GESTWK	Gestational week MOI begins **Format**: Numeric
GESTWKEND	Gestational week MOI ends **Format**: Numeric

j) [RUNID]_numcounts.sas7bdat for overlap

This output table is produced when using the overlap tool with a type 2 analysis.

Table 110. [RUNID]_numcounts.sas7bdat

Variable	Description
ANALYSISGRP	**Details:** GROUP name to differentiate primary/secondary pairs. **Format:** Character (40)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. **Format:** Character(4)
SEX	Sex. Allowable values are those in the SCDM. **Format:** Character(2)
AGEGROUP	Age Groups. Categories are requester-defined. **Format:** Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format:** Numeric
YEAR	Year of index date. **Format:** Numeric
MONTH	Month of index date. **Format:** Numeric
TTE_CAT	Primary episode length categories **Format:** Numeric
TOTAL_DAYS_OVERLAP	Number of days of overlap **Format:** Numeric
ADHERENCE	Meets adherence **Format:** Numeric
NPTS	Number of patients. **Format:** Numeric
EPISODES	Number of Episodes **Format:** Numeric
RawCodeCount1	Number of index defining codes for primary episode. **Format:** Numeric
RawCodeCount2	Number of index defining codes for secondary episode. **Format:** Numeric
AdjustedCodeCount1	Number of index defining codes for primary episode adjusted for codes incurred on the same date. **Format:** Numeric

Variable	Description
AdjustedCodeCount2	Number of index defining codes for secondary episode adjusted for codes incurred on the same date. **Format:** Numeric
DAYSUPP1	Days supply associated with dispensing for primary episode. **Format:** Numeric
DAYSUPP2	Days supply associated with dispensing for secondary episode. **Format:** Numeric
EPS_WSecEpi	Number of primary episodes with at least one secondary episode **Format:** Numeric
EPS_WOSecEp	Number of primary episodes with no secondary episode **Format:** Numeric
NPTS_WSecEp	Number of users with at least one secondary episode **Format:** Numeric
NPTS_WOSecEp	Number of users with no secondary episode **Format:** Numeric
TTE	Total duration of primary episode **Format:** Numeric
TOTAL_OVERLAP	Total number of days overlap between primary and secondary episodes **Format:** Numeric
EPI_XX_XX (CUTOFFCAT)	Output for each CUTOFFCAT. Number of episodes that where CUTOFFCAT value = 1 **Format:** Numeric
NPTS_XX_XX (CUTOFFCAT)	Output for each CUTOFFCAT. Number of users where at least one episode had a CUTOFFCAT value = 1 **Format:** Numeric
EPISODES	Number of primary episodes that meet adherence **Format:** Numeric
NPTS	Number of patients that meet adherence **Format:** Numeric

k) [RUNID]_denomcounts.sas7bdat

This output table is produced when using the Overlap and Multiple Events tool in a type 2 analysis. For each [GROUP] and Stratification Level, this table will output the number of eligible members and the number of eligible member days.

Table 111. [RUNID]_denomcounts.sas7bdat

Variable	Description
GROUP	**Details:** Cohort group **Format:** Character (40)
LEVEL	Stratification identifier. Each unique combination of strata (*i.e.,* variables bolded in the Variable column) receives a unique level value that remains consistent across requests. This allows for simpler development of reusable report generation tools. **Level** **Stratification** 000 Overall 002 Sex 003 Age Group 011 Year*Month **Format:** Character(3)
SEX	Sex. Allowable values are those in the SCDM. **Format**: Character(2)
AGEGROUP	Age Groups. Categories are requester-defined. **Format**: Character(variable)
AGEGROUPNUM	Numeric identifier of each AGEGROUP value. **Format**: Numeric
YEAR	Year of index date. **Format**: Numeric
MONTH	Month of index date. **Format**: Numeric
DENNUMPTS	Number of patients eligible to have at least one index date. Only calculated for overall, age group, sex, and year*month stratified analysis. **Format**: Numeric
DENNUMMEMDAYS	Number of days that patients are eligible to have an index date. Only calculated for overall, age group, sex, and year*month stratified analysis. **Format**: Numeric

l) Output Files for High-Dimensional Propensity Score and for Most Frequent Utilization

Eleven output tables are created for each execution of the CIDA tool if the most frequent utilization file is requested (MFUFILE is assigned) or high-dimensional propensity score matched analyses are planned (ANALYSIS=PS or ADS; HDPS=Y):

- [*RUNID*]_claims_drugclass.sas7bdat
- [*RUNID*]_claims_cpt.sas7bdat
- [*RUNID*]_claims_hcpcs.sas7bdat
- [*RUNID*]_claims_icdpx09.sas7bdat
- [*RUNID*]_claims_icddx09.sas7bdat
- [*RUNID*]_claims_icdpx10.sas7bdat
- [*RUNID*]_claims_icddx10.sas7bdat
- [*RUNID*]_claims_dxot.sas7bdat (specific to Most Frequent Utilization analysis only)
- [*RUNID*]_claims_pxot.sas7bdat (specific to Most Frequent Utilization analysis only)
- [*RUNID*]_claims_lab.sas7bdat (specific to Most Frequent Utilization analysis only)
- [*RUNID*]_claims_rx.sas7bdat (specific to Most Frequent Utilization analysis only)

These eleven datasets represent all codes for a given data dimension during the widest of covariate evaluation window and most frequent use window, and are only generated for the last PERIODID specified by the CIDA tool (i.e., the longest query period that encompasses the date ranges of all prior PERIODIDs). A dataset will be created for each of the following data dimensions: drug class, dispensings, labs, ICD-9-CM diagnosis, ICD-10-CM diagnosis, ICD-9-CM procedure, ICD-10-CM procedure, CPT, and HCPCS. These datasets are not de-duped and are only created for the last PERIODID value processed by the CIDA tool. Table 112 represents the output for all seven data dimension files, with the dimension denoted as *typeid*.

Table 112. claims_ [typeid] Output

Variable	Description
PatID	Patient Identifier.
GROUP	Cohort name.
Code	Code (or drug class).
IndexDt	Index Date.
ADate	Claim Date.
RxSup	Days supplied. **Note:** Applies to RX only.
RxAmt	Amount supplied. **Note:** Applies to RX only.
DX_CodeType	Diagnosis code type. **Note:** Applies to ICD-9-CM diagnosis, ICD-10-CM diagnosis, and Other diagnosis only.
EncType	Encounter type. **Note:** Applies to ICD-9-CM diagnosis, ICD-10-CM diagnosis, Other diagnosis, CPT, HCPCS, ICD-9-CM procedure, ICD-10-CM procedure, and Other procedure only.

Variable	Description
PDX	Principal discharge diagnosis flag. **Note:** Applies to ICD-9-CM diagnosis, ICD-10-CM diagnosis, and Other diagnosis only.
PX_CodeType	Procedure code type. **Note:** Applies to labs, CPT, HCPCS, ICD-9-CM procedure, ICD-10-CM procedure, and Other procedure only.
RESULT_TYPE	Type of result (qualitative or quantitative). **Note:** Applies to labs only.
LOINC	Logical Observation Identifiers, Names, and Codes. **Note:** Applies to labs only.
PX	Procedure code **Note:** Applies to labs only.
MS_TEST_NAME	Abbreviation for the type of lab test. **Note:** Applies to labs only.
SPECIMEN_SOURCE	Specimen source for MS_Test_Name. **Note:** Applies to labs only.
MS_TEST_SUB_CATEGORY	Sub-category for MS_Test_Name. **Note:** Applies to labs only.
MS_RESULT_UNIT	Converted/standardized units for the result **Note:** Applies to labs only.
FAST_IND	Fasting Indicator **Note:** Applies to labs only.
PT_LOC	Patient location where the lab specimen was obtained. **Note:** Applies to labs only.

m) [*RUNID*]_ctrl.sas7bdat

The [*RUNID*] _ctrl output table is generated for Type 4 requests. [RUNID]_ctrl describes the comparator cohort for the pregnancy episodes cohort identification strategy. Variables specific to pregnancy episodes are the values that correspond to the matched pregnancy episode (Table 113).

Table 113. [RUNID] _ctrl Output

Variable	Description
GROUP	Cohort name. **Format**: Character(40)
PATID	Individual identifier. **Format**: Character(variable)
INDEXDT	Index date (pregnancy start date). Can be determined by a preterm/postterm code. **Format**: Numeric (date9.)
ENR_START	Start of enrollment span where index exposure was identified. **Format**: Numeric (date9.)
ENR_END	End of enrollment span where index exposure was identified. **Format**: Numeric (date9.)
EPISODEENDDT	Pregnancy episode end date (delivery). **Format**: Numeric (date9.)
AGE	Age at index date. **Format**: Numeric
BIRTH_DATE	Birth Date. **Format**: Numeric (MMDDYY10.)
SEX	Sex. Allowable values are those in the SCDM Demographic table. **Format**: Character(1)
RACE	Race. Allowable values are those in the SCDM Demographic table. **Format**: Character(1)
HISPANIC	Hispanic. Allowable values are those in the SCDM Demographic table. **Format**: Character(1)
AGEGROUP	Age Group at delivery date. **Format**: Character(variable)
FEVENTDT	Blank for Type 4 analyses. **Format**: Numeric
CONCWASHPER	Washout used to determine incidence of medical product use episodes with respect to delivery episodes when assessing "trimester only" indicators **Format**: Numeric
HADPRETERM	Indicator if pregnancy episode has a preterm birth code **Format**: Numeric

Variable	Description
HADPOSTTERM	Indicator if pregnancy episode has a postterm birth code **Format**: Numeric
HADNONE	Indicator if pregnancy episode has neither a preterm or a postterm birth code **Format**: Numeric
PREPOSTIND	Indicator for type of Preterm/Postterm code (PRETERM, POSTTERM, TERM, NONE) **Format**: Character (10)
PREGDURCODE	Prioritygroup1/Prioritygroup2 code, if any **Format**: Character (11)
HAS3TRIM	Indicates if the pregnancy episode reaches the 3rd trimester **Format**: Numeric
DELNUM	Delivery number for the patient. This is the overall delivery number for the patient and is not confined to valid deliveries **Format**: Numeric
YEAR	Year of delivery date. **Format**: Numeric
EXACTNUMVISIT	The total number of medical encounters during the requester-defined evaluation period. **Format:** Numeric
NUMVISITS_AV	The total number of ambulatory encounters during the requester-defined evaluation period. **Format:** Numeric
NUMVISITS_OA	The total number of other ambulatory encounters during the requester-defined evaluation period. **Format:** Numeric
NUMVISITS_IP	The total number of inpatient encounters during the requester-defined evaluation period. **Format:** Numeric
NUMVISITS_IS	The total number of institutional stay encounters during the requester-defined evaluation period. **Format:** Numeric
NUMVISITS_ED	The total number of emergency department encounters during the requester-defined evaluation period. **Format:** Numeric
NUMVISITS	The requester-defined stratum that the EXACTNUMVISIT value is in. **Format**: Numeric
NUMGENERIC	The total number of generic names dispensed during the requester-defined evaluation period.

Variable	Description
	Format: Numeric
NumClass	The total number of unique class names dispensed during the requester-defined evaluation period. **Format**: Numeric
NUMRX	The total number of dispensings during the requester-defined evaluation period. **Format**: Numeric
COMBINED_SCORE	The comorbidity score requester-defined category. **Format:** Character(3)
COMBINED_SCORE_NUM	Thecomorbidity score value. **Format:** Numeric
CCIELIXGRP	The pre-index comorbidity score requester-defined category. **Format:** Character(10)

n) [*RUNID*]_alldeliveries.sas7bdat

The [*RUNID*] _alldeliveries output table is generated for Type 4 requests. [RUNID]_alldeliveries contains all deliveries contained in [RUNID]_mstr, plus those that do not meet the enrollment eligibility criterion. It is used to compute eligible pregnancy episodes by days of continuous enrollment prior to delivery date (Table 114Table 114).

Table 114. [RUNID] _alldeliveries Output

Variable	Description
GROUP	Cohort name. **Format**: Character(40)
PATID	Individual identifier. **Format**: Character(variable)
ENR_START	Start of enrollment span where index exposure was identified. **Format:** Numeric (date9.)
ENR_END	End of enrollment span where index exposure was identified. **Format:** Numeric (date9.)
EPISODESTARTDT	Pregnancy episode start date. Can be determined by a preterm/postterm code. **Format:** Numeric (MMDDYY10.)
EPISODEENDDT	Pregnancy episode end date (delivery). **Format:** Numeric (date9.)
AGEGROUP	Age Group. **Format**: Character(variable)
PSADATE	Preterm/Postterm code date, if any **Format:** Numeric (MMDDYY10.)

Variable	Description
DELNUM	Delivery number for the patient. **Format:** Numeric
PREPOSTIND	Indicator for type of Preterm/Postterm code (PRETERM, POSTTERM, NONE). **Format**: Character (10)
HADPRETERM	Indicator if pregnancy episode has a preterm birth code **Format**: Numeric
HADPOSTTERM	Indicator if pregnancy episode has a postterm birth code **Format**: Numeric
HADNONE	Indicator if pregnancy episode has neither a preterm or a postterm birth code **Format**: Numeric
DEATHDT	Date on which the individual died, if any. (value will be set to missing if the individual did not die) **Format**: Numeric (date9.)
ENRDAYSFLOOR	Minimum number of days of continuous enrollment required prior to the delivery date. **Format**: Numeric
ELIGDAYS	Number of days of continuous enrollment prior to EpisodeEndDate (delivery date). **Format:** Numeric

o) [*RUNID*]_sec.sas7bdat

The [*RUNID*] _sec output table is generated for Type 4 requests. [RUNID]_sec contains each medical product of interest episode for both the pregnancy and comparator cohorts (Table 115).

Table 115. [RUNID] _sec Output

Variable	Description
GROUP	Cohort name. **Format**: Character(40)
PATID	Individual identifier. **Format**: Character(variable)
ENR_START	Start of enrollment span where index exposure was identified. **Format:** Numeric (date9.)
ENR_END	End of enrollment span where index exposure was identified. **Format:** Numeric (date9.)
INDEXDT	Index date for medical product episode **Format:** Numeric (date9.)
EPISODEENDDT	End date for medical product episode **Format:** Numeric (date9.)

Variable	Description
BIRTH_DATE	Birth Date. **Format:** Numeric (MMDDYY10.)
EPISODE	Episode number indicator. Counts the number of episodes within a categorization **Format:** Numeric
SEX	Sex. Allowable values are those in the SCDM Demographic table. **Format:** Character(1)
RACE	Race. Allowable values are those in the SCDM Demographic table. **Format:** Character(1)
HISPANIC	Hispanic. Allowable values are those in the SCDM Demographic table. **Format**: Character(1)
MOINAME	Medical product category identifier. **Format:** Character(9)
ENRDAYSFLOOR	Minimum number of days of continuous enrollment required prior to the delivery date. **Format**: Numeric
DEATHDT	Date on which the individual died, if any. (value will be set to missing if the individual did not die) **Format**: Numeric (date9.)

p) [*RUNID*]_pregdurcodes.sas7bdat

The [*RUNID*] _pregdurcodes output table is generated for Type 4 requests. [RUNID]_pregdurcodes contains gestational age codes for pregnancy episodes in [RUNID]_mstr (Table 116). In the case when an individual has both Prioritygroup1 and Prioritygroup2 codes observed during an episode, only Prioritygroup1 codes are reported.

Table 116. [RUNID] _pregdurcodes Output

Variable	Description
GROUP	Cohort name. **Format**: Character(40)
PATID	Individual identifier. **Format:** Character(variable)
PRETERM	Indicator if pregnancy episode has a preterm birth code **Format:** Numeric
POSTTERM	Indicator if pregnancy episode has a postterm birth code **Format:** Numeric
CODE	Preterm/Postterm code **Format:** Character (11)
DELNUM	Delivery number for the patient. **Format:** Numeric

q) [RUNID]_dispensings

The [*RUNID*] _dispensings output table is generated for Type 5 requests. This output table contains one row per scenario (GROUP), patient (PATID), and and dispensing (RXDATE) and only contains valid claims that occur during valid enrollment spans and episodes.

Table 117. [RUNID]_dispensings Output Table

Field Name	Description
STUDYSTARTDATE	**Details:** query period start date specified in request.
STUDYENDDATE	**Details:** query period end date specified in request.
GROUP	**Details:** standardized name used to differentiate cohorts.
PATID	**Details:** unique patient identifier in SCDM
BIRTH_DATE	**Details:** patient birth date in SCDM.
SEX	**Details:** patient sex in SCDM.
AGEGROUP	**Details:** patient age group (age calculated at index date)
RACE	**Details:** patient race in SCDM.
HISPANIC	**Details:** patient Hispanic value in SCDM.
MINAGEDATE	**Details:** first date the patient qualifies for cohort inclusion, based on specified AGESTRAT parameters.
MAXAGEDATE	**Details:** last date the patient qualifies for cohort inclusion, based on specified AGESTRAT parameters.
INDEXDT	**Details**: start date of the first exposure episode.
EPISODENUM	**Details:** episode number, in sequential order (starting at 1), by PatID.
EPISODEDISPENSING	**Details:** number of dispensings in the episode.
EPISODERXSUP	**Details:** days supply associated with all dispensings in the episode denoted by EPISODENUM.
DISP_RXDATE	**Details:** date of the dispensing. This output table will have one row per GROUP, PATID, and RXDATE.
DISP_RXSUP	**Details:** days supply associated with a dispensing.

r) [RUNID]_t6_switchepisodes.sas7bdat

This output file includes patient-switch pattern episode level information.

Table 118. [RUNID]_t6_switchepisodes OUTPUT

Variable	Description
ANALYSISGRP	**Details:** standardized name used to differentiate switch patterns. **Format**: Character(40)
PATID	Individual identifier. **Format:** Character(variable)
SEX	Sex. Allowable values are those in the SCDM Demographic table. **Format:** Character(1)
AGEGROUP	**Details:** patient age group (age calculated at index date) Age Group at index date. **Format:** Character(variable)
ZIP3	3-digit ZIP code. **Format:** Numeric
STATE	2-digit State code. **Format:** Character(2)
HHS_REG	2-digit Health and Human Services region code. **Format:** Character(2)
CB_REG	Census Bureau regions. **Format:** Character(7)
ZIP_UNCERTAIN	Y / N variable. Y: zip date occurred after the index date N: zip date occurred before the index date **Format:** Character(1)
INDEXDT0	**Details:** Start date of treatment episode. **Format**: Numeric (date9.)
EPISODEENDDT0	**Details:** Episode End date of treatment episode **Format**: Numeric (date9.)
ENR_START0	**Details:** Enrollment Start Date **Format**: Numeric (date9.)
ENR_END0	**Details:** Enrollment End Date **Format**: Numeric (date9.)
SWITCHINDEXDT1	**Details:** Index Date of switch **Format**: Numeric (date9.)
SWITCHEPISODEENDDT1	**Details:** Episode End date of switch

Variable	Description
	Format: Numeric (date9.)
SWITCH1	**Details:** Indicator (Y/N) for Switching **Format:** Character(1)
CENS_ELIG0	**Details:** 0/1 indicating if episode was censored due to disenrollment. **Format:** Numeric
CENS_ELIG1	**Details:** 0/1 indicating if episode was censored due to disenrollment. **Format:** Numeric
CENS_DTH0	**Details:** 0/1 indicator denoting if episode was censored due to death. **Format:** Numeric
CENS_DTH1	**Details:** 0/1 indicator denoting if episode was censored due to death. **Format:** Numeric
CENS_DPEND0	**Details:** 0/1 indicator denoting if episode was censored due to DP data end date. **Format:** Numeric
CENS_DPEND1	**Details:** 0/1 indicator denoting if episode was censored due to DP data end date. **Format:** Numeric
CENS_QRYEND0	**Details:** 0/1 indicator denoting if episode was censored due to study end date. **Format:** Numeric
CENS_QRYEND1	**Details:** 0/1 indicator denoting if episode was censored due to study end date. **Format:** Numeric
CENS_EPISEND0	**Details:** 0/1 indicator denoting if episode was censored due to end of exposure. **Format:** Numeric
CENS_EPISEND1	**Details:** 0/1 indicator denoting if episode was censored due to end of exposure. **Format:** Numeric
CENSOR_DATE0	**Details:** Date of Censoring. **Format**: Numeric (date9.)
CENSOR_DATE1	**Details:** Date of Censoring. **Format**: Numeric (date9.)
SWITCHPATTERNENDDT	**Details:** Switch Pattern End Date **Format**: Numeric (date9.)

Variable	Description
COMPUTEDSTARTMARKETINGDATE	**Details:** Date associated with product start marketing date. **Format**: Numeric (date9.)
EXCL_SWITCHDATE	**Details:** 1/0 indicator denoting exclusion **Format:** Numeric
METCOHORTDEF	**Details:** indicator 1/0 denoting if the patient met cohort definition. **Format:** Numeric
TIME_TO_FIRST_SWITCH	**Details:** Number of days to first switch. **Format:** Numeric
TIME_TO_SECOND_SWITCH	**Details:** Number of days to second switch. **Format:** Numeric
YEAR	**Details:** Year of the Index Date. **Format**: Numeric
CENS_SWITCH0	**Details:** 0/1 indicator denoting if episode was censored due to switching. **Format:** Numeric

X. APPENDIX C: PSA TOOL TECHNICAL DOCUMENTATION

The PSA tool is designed to be executed following the execution of the CIDA tool. The CIDA tool identifies and extracts cohorts of interest based on requester-defined parameters, and generates output tables in the *msoc* and *dplocal* folders that are required for subsequent processing with the PSA tool. Specifically, the output tables generated by the CIDA tool that are required for the PSA tool are:

For predefined and high-dimensional propensity score matched analyses:

- [*RUNID*]_ads_mstr_[*PERIODID*].sas7bdat

For high-dimensional propensity score matched analyses only:

- [*RUNID*]_claims_drugclass.sas7bdat
- [*RUNID*]_claims_cpt.sas7bdat
- [*RUNID*]_claims_hcpcs.sas7bdat
- [*RUNID*]_claims_icdpx09.sas7bdat
- [*RUNID*]_claims_icddx09.sas7bdat
- [*RUNID*]_claims_icdpx10.sas7bdat
- [*RUNID*]_claims_icddx10.sas7bdat

This technical specification document details the lookup tables, program parameters and input files that must be specified to execute the PSA tool. These parameters and files should be included in the program package distributed *in addition* to those necessary for CIDA tool execution, if a propensity score matched analysis is requested.

A. LOOKUP TABLES, PROGRAM PARAMETERS, AND INPUT FILES

1. Lookup Tables

There are no lookup tables required for the execution of the PSA tool.

2. Main Program Parameters

There are several main program parameters that must be specified. These include defining the covariate selection strategy and number of covariates considered and selected for the hdPS estimation and specifying the name of all input files. These parameter values should be set in a program called run_programs.sas, located in the *inputfiles* folder. Note that all main programs parameters specified are fixed for a single execution of the program. Table 119 contains detailed specifications for main program parameters.

Table 119. PSA Tool Main Program Parameters

Parameter	Field Name	Description
Comparison File	COMPARISON	**Details:** name of the SAS dataset describing all exposure/comparator pairs that should be evaluated using the PSA tool. **Note 1:** If COMPARISON is specified, MFMFILE input parameter should be blank **Named by:** Request programmer **Input type:** Optional **Format:** .sas7bdat file format **Example:** COMPARISON= comparison
Number of Covariates Considered for hdPS Model by Data Dimension	COVARIATES_CONSIDERED	**Details:** relevant for requests calculating a high-dimensional propensity score. Specifies the number of variables from each data dimension (diagnoses, procedures and drug classes) that are considered for selection as a covariate. If not specified, default value is 100. **Note 1:** there are seven data dimensions: drug class, ICD-9-CM diagnosis, ICD-10-CM diagnosis, ICD-9-CM procedure, ICD-10-CM procedure, CPT, HCPCS. **Note 2:** leave blank for requests not calculating a high-dimensional propensity score. **Defined by:** Requester **Input type:** Optional (default values will be used if missing) **Format:** Numeric (positive whole numbers) **Example:** COVARIATES_CONSIDERED=150
Number of Covariates Selected for hdPS Model	COVARIATES_SELECTED	**Details:** relevant for requests calculating a high-dimensional propensity score. Specifies the total number of empirically identified covariates selected from all the data dimensions combined. These selected covariates are used for the high-dimensional propensity score. If not specified, default value is either 200 or number of new users of study drug, whichever value is smallest. **Note 1:** leave blank for requests not calculating a high-dimensional propensity score. **Defined by:** Requester **Input type:** Optional (default values will be used if missing) **Format:** Numeric (positive whole numbers) **Example:** COVARIATES_SELECTED=150

Parameter	Field Name	Description
Method for Selecting Covariates in hdPS Model	RANKING	**Details:** relevant for requests calculating a high-dimensional propensity score. Indicates one of three models for selecting variables. Allowable values: • **exp_assoc**: yields a variable list in which the variables are selected as ranked by the strength of the relationship between confounder and exposure. This is most suitable for cases where there are fewer than 150 exposed outcomes. (default) • **outcome_assoc**: yields a variable list in which the top k variables are selected as ranked by the strength of the relationship between the confounder and the outcome. This is most suitable for disease risk scores. • **bias**: yields a variable list in which the top k variables are selected as ranked by the Bross bias formula (here k is the number entered for COVARIATES_SELECTED) **Defined by:** Requester **Input type:** Optional (default values will be used if missing) **Format:** Character **Example:** RANKING= exp_assoc
Zero Cell Correction for Association with Exposure	ZERO_CELL_CORR	**Details:** indicates whether to screen variables with a zero correction added to each cell in the confounder/outcome 2x2 table. Recommended when the number of exposed outcomes is fewer than 150. Allowable values: • **0:** No • **1:** Yes (default) **Defined by:** Requester **Input type:** Optional (default values will be used if missing) **Format:** Binary **Example:** ZERO_CELL_CORR =1
Analytic Subgroups File	ANALYTICSUBGROUPS	**Details:** name of the SAS dataset describing all subgroup analyses requested for each comparison that should be evaluated using the PSA tool. **Named by:** Request programmer **Input type:** Optional

Parameter	Field Name	Description
		Format: .sas7bdat file format **Example:** ANALYTICSUBGROUPS= analyticsubgroups
Indicator for Individual Level Output Return	INDLEVEL	**Details:** Specifies whether to write individual level [*RUNID*]_matched_[*COMP*]_[*Look*] files to the SOC subfolder. Allowable values: • **Y:** Yes • **N:** No **Defined by:** Requester **Input type:** Required **Format:** Binary **Example:** IndLevel = Y
Percentile Stratification	PERCENTILES	**Details:** Specifies the number of percentiles to group propensity scores for percentile analysis. For example, 10 = deciles. **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** Percentiles = 10
Return unconditional analysis	UNCONDITIONAL	**Details:** Specifies whether to create risk sets not stratified by matchID. This allows SOC to perform unconditional analyses. Allowable values: • **Y:** Yes • **N:** No **Defined by:** Requester **Input type:** Required **Format:** Binary **Example:** UNCONDITIONAL = Y
Return PS estimation model diagnostics	DIAGNOSTICS	**Details:** Specifies if program should return three diagnostic output files from the proc logistic for each PS model: 1. Pair-wise correlation coefficients between each variable in the PS model 2. Output from proc reg model, including parameter estimates, and intercept adjusted collinearity diagnostics tables 3. Output from proc logistic model, including descriptive statistics for continuous variables, frequency distribution of class variables, the maximum likelihood iteration history, odds ratio estimates,

Parameter	Field Name	Description
		and the association of predicted probabilities and observed responses tables Allowable values: • **1:** Yes, return diagnostics • **0:** No, do not return diagnostics **Defined by:** Requester **Input type:** Required **Format:** Numeric **Example:** DIAGNOSTICS = 1

3. Input Files

This section describes all input files that must be included in the *inputfiles* folder in the program package to execute the PSA tool. These files are *in addition* to those required for execution of the CIDA tool (including the Medical Utilization, Comorbidity Score, and Covariate Codes files, which are optional for the CIDA tool but required if propensity score matched analyses are requested).

a) Comparison File

The Comparison File is required. The file allows requesters to specify all exposure/comparator pairs that should be evaluated in the propensity score matched analysis. Each exposure/comparator pair is assigned a unique COMP_ORDER value, to differentiate pairs in output tables. Table 120 contains specifications for this file.

Table 120. COMPARISON File Specification

Parameter	Field Name	Description
Comparison Identifier	COMP_ORDER	**Details:** numeric identifier to differentiate exposure/comparator pairs and parameter settings in output tables. **Note 1:** COMP_ORDER (&COMP value) is used as a suffix in output tables to identify each comparison. **Defined by:** Request programmer **Input type:** Required (cannot be left blank) **Format:** Numeric **Example:** 2
Exposure of Interest	COMP	**Details:** COHORTGRP name of the exposure of interest. **Defined by:** Requester **Input type:** Required (cannot be left blank) **Format:** Character(40) **Example:** drug_a
Comparator of Interest	CONTROL	**Details:** COHORTGRP name of the comparator of interest.

Parameter	Field Name	Description
		Note 1: When the never-exposed cohort serves as the comparator of interest, the group name should be <COMP>_nvrexp **Defined by:** Requester **Input type:** Required (cannot be left blank) **Format:** Character(40) **Example:** drug_b
Matching Caliper	CALIPER	**Details:** identifies caliper that should be used for propensity score matching. Any value between 0-1 is allowed. **Defined by:** Requester **Input type:** Required (cannot be left blank) **Format:** Numeric **Example:** 0.01
Matching Ratio	RATIO	**Details:** identifies the matching ratio. Allowable values: • F: fixed 1:1 matching • V: variable 1:10 matching **Defined by:** Requester **Input type:** Required (cannot be left blank) **Format:** Char (1) **Example:** F
Categorical variables to include in Propensity score model	CLASS	**Details:** used to specify variables to include in the 'CLASS' statement in the PS logistic regression model. All variables should be separated by a space. Allowable values include: • Sex • Race • Hispanic • Year • Time • Covar1 … CovarN For Type 4 Analyses only: • PrePostInd • Exp_t1 • Exp_t2 • Exp_t3 • Exp_pre **Named by:** Requester **Input type:** Optional **Format:** Character **Example:** Sex Time Year Covar1 Covar2 Covar3

Parameter	Field Name	Description
Continuous variables to include in Propensity score model	NOCLASS	**Details:** used to specify variables to include as continuous in the PS logistic regression model. All variables should be separated by a space. Allowable values include:Age • Comorbidscore • NumIP • NumIS • NumED • NumAV • NumOA • NumRx • NumGeneric • Covar1 … CovarN For Type 4 Analyses only: • GA_birth • GA_first • AdjustedDisp_Pre • AdjustedDisp_T1 • AdjustedDisp_T2 • AdjustedDisp_T3 **Named by:** Requester **Input type:** Optional **Format:** Character **Example:** Age comorbidscore numIP numED
High-dimensional Propensity Score Indicator	HDPS	**Details:** indicates whether a high-dimensional propensity score will be estimated for the specified COMP_ORDER. Allowable values are “Y” and “N”. **Defined by:** Requester **Input type:** Required **Format:** SAS character $1. **Example:** N (default)
HDPS Evaluation Period Start	HDPSWINFROM	**Details:** used in combination with HDPSWINTO (below). HDPSWINFROM defines the start of the HDPS evaluation period when HDPS=Y, expressed in terms of “days from Index Date”. For example, if Index Date=01/08/2009 and HDPSWINFROM is set to -7, the MP algorithm will start evaluating HDPS variables on 01/01/2009. **Note 1**: the index date is “day zero”. Therefore, if zero is included in the HDPSWINFROM-HDPSWINTO interval for a given COMP_ORDER, the index date is included in the evaluation period.

Parameter	Field Name	Description
		Note 2: Required when HDPS = Y. Should be left blank when HDPS = N. **Note 3**: if HDPSWINFROM > 0 then the evaluation period will start after the index date. **Note 4: special case**: when HDPSWINFROM = missing the program considers all codes in their entire available history before the index date. In this case, continuous enrollment is *not required* for the duration of the evaluation period (only explicitly defined enrollment criteria, e.g., specified using the ENRDAYS value, are required). **Note 5:** leave blank if HDPS=N **Named by:** Requester **Input type:** Optional **Format:** Numeric **Example:** -180
HDPS Evaluation Period End	HDPSWINTO	**Details:** used in combination with HDPSWINFROM (above). HDPSWINTO defines the end of the HDPS evaluation period when HDPS=Y, expressed in terms of "days from Index Date". For example, if Index Date=01/08/2009 and HDPSWINTO is set to -1, the MP algorithm will evaluate HDPS variables between the HDPSWINFROM date through 01/07/2009. **Note 1**: the index date is "day zero". Therefore if zero is included in the HDPSWINFROM-HDPSWINTO interval for a given COMP_ORDER value the index date is included in the evaluation period. **Note 2: special case**: when HDPSWINTO = missing the program considers all codes in their entire available history after the index date. In this case, continuous enrollment is *not required* for the duration of the evaluation period (only explicitly defined enrollment criteria, e.g., specified using the ENRDAYS value, are required). **Note 3:** leave blank if HDPS=N **Named by:** Requester **Input type:** Required **Format:** Numeric **Example:** -1

b) Analytic Subgroups File

The Analytic Subgroups file is optional. The file allows requesters to specify all subgroups and subgroup levels for each comparison (i.e., exposure/comparator/ratio/caliper/model combination under investigation). The file also allows the user to select the base population for subgroup analyses (i.e., total population or successfully matched population from the primary analysis). Table 121 Table 121 contains specifications for this file.

Table 121. Analytic Subgroups File Specification

Parameter	Field Name	Description
Comparison Identifier	COMP_ORDER	**Details:** numeric identifier to differentiate exposure/comparator pairs and parameter settings in output tables. **Note 1:** COMP_ORDER (&COMP value) is used as a suffix in output tables to identify each comparison. **Defined by:** Request programmer **Input type:** Required (cannot be left blank) **Format:** Numeric **Example:** 2
Covariate Number	COVARNUM	**Details:** numeric identifier to indicate covariates needed for subgroup analyses. For binary covariates specified in the COVARIATECODES input file, enter the COVARNUM value from COVARIATECODES here. For covariates not explicitly named in the COVARIATECODES file, use the following values: • **Sex**: 1000 • **Age**: 1001 • **Year**: 1002 • **Time**: 1003 • **Comorbidscore**: 1004 • **NumIP:** 1005 • **NumIS:** 1006 • **NumED:** 1007 • **NumAV:** 1008 • **NumOA:** 1009 • **NumRx:** 1010 • **NumGeneric**: 1011 • **Race:** 1012 • **Hispanic:** 1013 For Type 4 Analyses only: • MATCHMETHOD (2000) • Birth_Type (2001) **Input type:** Required (cannot be left blank) **Format:** Numeric

Parameter	Field Name	Description
		Example: 1000
Covariate Level Specification	CATEGORIZATION	**Details:** specifies how subgroup variables should be categorized for subgroup analysis. **Note 1:** for COVARNUM=1-999, CATEGORIZATION should be left blank as these are dichotomous variables. **Note 2:** for COVARNUM=1000 (sex), this should be left blank. Sex values of F, M, U, and A will automatically be categorized **Note 3:** for COVARNUM=1001 (Age), valid CATEGORIZATION values should follow the same rules as the AGESTRAT macro parameter in the CIDA tool (e.g., 40-59 60-79 80-99). However, the CATEGORIZATION value does not have to be the same as the AGESTRAT value used in CIDA (i.e., subgroup analyses can specify different age groups). **Note 4:** for • Year (COVARNUM=1002) • Time (COVARNUM=1003) • Comorbidscore (COVARNUM=1004) • NumIP (COVARNUM=1005) • NumIS (COVARNUM=1006) • NumED (COVARNUM=1007) • NumAV (COVARNUM=1008) • NumOA (COVARNUM=1009) • NumRx (COVARNUM=1010) • NumGeneric (COVARNUM=1011) Groups must be separated by a space. To leave the first group open-ended, use "low-". In the output "low-", will be replaced with "<=". If a negative is desired as the upper bound of a group, do not include a space in the group (*e.g.,* use low--1 for low to -1). To leave the last group open-ended, use "-high". In the output "-high", will be replaced with ">=". Note that groups should not have overlapping values. In the event that overlapping values are entered, the value will be mapped to the first group in the list. For example, if age groups are specified "45-85 80-99", patients 80-85 will be included in the 45-85 category (and re-matched within that group). **Note 5:** the theoretical range of comorbidity scores is -2 -26. **Defined by:** Requester **Input type:** Required **Format:** Alphanumeric; SAS character $100.

Parameter	Field Name	Description
		Example: low-0 1 2-3 4-7 8+
Subgroup Matching Base Population Specification	MATCHEDINFULLONLY	**Details:** specifies if the base population for subgroup re-matching should be restricted to the matched population. **Y:** subgroup re-matching should be restricted to the matched population **N:** subgroup re-matching should be done in the total population (i.e., patients who were matched and patients who were not matched in the main analysis) **Input type:** Required **Format:** Character **Example:** N

B. OUTPUT

1. MSOC Folder

The following output files are created and output to the *msoc* folder for each comparison in the Comparison File, time period, and execution of the PSA tool:

- [*RUNID*]_matched_[*COMP_ORDER*]_[*PERIODID*].sas7bdat
- [*RUNID*]_risksetdata_[*COMP_ORDER*]_[*PERIODID*].sas7bdat
- [*RUNID*]_riskdiffdata_[*COMP_ORDER*]_[*PERIODID*].sas7bdat
- [*RUNID*]_survivaldata_[*COMP_ORDER*]_[*PERIODID*].sas7bdat
- *[RUNID]_matched*_tables_[*COMP_ORDER*]_[*PERIODID*].sas7bdat
- *[RUNID]_estimates*_[*COMP_ORDER*]_[*PERIODID*].sas7bdat
- *[RUNID]*_corr_[*COMP_ORDER*]_[*PERIODID*].sas7bdat (if requested in run_programs.sas)
- *[RUNID]*_runreg_[*COMP_ORDER*]_[*PERIODID*].rtf (if requested in run_programs.sas*)*
- *[RUNID]*_runmodel_[*COMP_ORDER*]_[*PERIODID*].rtf (if requested in run_programs.sas)
- [*RUNID*]_psdistribution_[*COMP_ORDER*]_[*PERIODID*].sas7bdat
- [*RUNID*]_signature_ps_[*COMP_ORDER*]_[*PERIODID*].sas7bdat

Two additional output files are created and output to the msoc folder for each execution of the PSA tool:

- [*RUNID*]_timing.sas7bdat
- [RUNID]_varinfo_[COMP_ORDER]_[PERIODID].sas7bdat

Where RUNID is the request programmer-defined execution identifier, COMP_ORDER is the unique ps estimation and matching strategy identifier, and PERIODID is the time period.

a) [*RUNID*]_matched_[*COMP_ORDER*]_[*PERIODID*].sas7bdat

This output table contains the final matched sample, with identifiers for all matches performed (predefined covariates only, empirically identified covariates only, and predefined + empirically identified covariates). This table is only produced if INDLEVEL = Y in the PSA Tool Main Program Parameters.
Table 122Table 122 contains specifications for this output table.

Table 122. [RUNID] _matched_ [COMP_ORDER] _ [PERIODID].sas7bdat

Variable	Description
COVARNUM	The level of analysis that is performed (=0 when full analysis [overall population] and >0 when conducting a subgroup analysis). **Format:** Character(4)
STRATUMNAME	Indicates name of each strata for subgroup analyses. Should indicate "Overall" for overall analyses; will be populated with CATEGORIZATION values from the new ANALYTICSUBGROUPS input file for subgroup analyses, "Percentiles" for percentile analysis, and "Overall Unconditional" for unconditional analysis **Format:** Character(50)
STUDYCLASS	Analogous to the GROUP value on ADS_MSTR_[*PERIODID*] input file. **Format:** Character(30)
AGE_CAT	Requester-defined age category. Determined by AGESTRAT macro parameter values defined in the CIDA tool. **Format:** Character(9)
SEX	Patient Sex as defined in SDD Demographic table. **Format:** Character(1)
RACE	Patient Race as defined in SDD Demographic table. **Format:** Character(1)
HISPANIC	Hispanic as defined in SDD Demographic table. **Format:** Character(1)
EVENT	Boolean indicator of whether or not patient had the HOI during the time period (PERIODID). **Format:** Numeric
FOLLOWUPTIME	Number of days that the patient is followed for either HOI or censoring. Variable is calculated using the earliest of 1) end of exposure period; 2) disenrollment; 3) any additional censoring criteria [defined using CIDA tool]; and 4) HOI date. **Note 1:** This wil be blank for Type 4 analyses. **Format:** Numeric
MATCHID	MatchID for the analysis requested.

Variable	Description
LastLookFollowed	When an individual is lost to follow up (data is no longer being updated), this is the last look # the individual contributed data to. If an individual is still eligible to be followed in subsequent looks, this is set to 0. **Format:** Numeric
EXPOSURE	Boolean indicator of whether or not the STUDYCLASS value is the exposure of interest (COMP in the COMPARISON file). **Format:** Numeric
TIME	Variable indicating during which time period the patient was selected (*e.g.,* TIME = 1 indicates the patient was selected in PERIODID 1, even if the current PERIODID being executed is not the first). **Format:** Numeric
NUMIP	Visit count for IP encounter type. One visit allowed per encounter, per day. **Note 1**: this metric is automatically calculated using the medical utilization module in the CIDA tool (Medical Utilization File). **Format:** Numeric
NUMED	Visit count for ED encounter type. One visit allowed per encounter, per day. **Note 1**: this metric is automatically calculated using the medical utilization module in the CIDA tool (Medical Utilization File). **Format:** Numeric
NUMIS	Visit count for IS encounter type. One visit allowed per encounter, per day. **Note 1**: this metric is automatically calculated using the medical utilization module in the CIDA tool (Medical Utilization File). **Format:** Numeric
NUMAV	Visit count for AV encounter type. One visit allowed per encounter, per day. **Note 1**: this metric is automatically calculated using the medical utilization module in the CIDA tool (Medical Utilization File). **Format:** Numeric
NUMOA	Visit count for OA encounter type. One visit allowed per encounter, per day. **Note 1**: this metric is automatically calculated using the medical utilization module in the CIDA tool (Medical Utilization File). **Format:** Numeric
NUMGENERIC	Count of unique generic names dispensed during covariate evaluation window. **Format:** Numeric

Variable	Description
NUMCLASS	Count of unique class names dispensed during covariate evaluation window. **Format:** Numeric
NUMRX	Count of dispensings during the covariate evaluation window. **Format:** Numeric
COMORBIDSCORE	Combined Charlson-Elixhauser Comorbidity Score (exact value). **Note 1:** this metric is calculated using the combined comorbidity score module in the CIDA tool (Comorbidity Score File). **Format:** Numeric
YEAR	Year of index date. **Format:** Numeric
PSCORE/MFMS	Propensity score estimate. Name of the variable is determined based on the analysis requested. Allowable variable names include : • PSCORE • MFMS
COVAR1-COVAR*N*	Boolean indicators for each covariate (denoted by the value of the variable COVARNUM in the Covariate Codes File). **Format:** Numeric
PERCENTILE	Percentile number for propensity score. Calculated using the propensity score distribution from the Look the propensity score is derived in. **Format:** Character

b) [*RUNID*]_risksetdata_[*COMP_ORDER*]_[*PERIODID*].sas7bdat

This output table contains the risk sets for the entire sample, the matched sample, the entire stratified by percentiles, each subgroup, and the matched sample not stratified on matchid (when requested). This output table is not produced for Type 4 analyses, as time to event is not currently available in the pregnancy tool. Table 123 contains specifications for this output table.

Table 123. [RUNID]_risksetdata_[COMP_ORDER]_[PERIODID].sas7bdat

Variable	Description
RISKSETID	Unique risk-set identifier **Format:** Numeric
TYPE	Indicates if the records are for the matched cohort or for all patients. **Format:** Character(20)
COVARNUM	The level of analysis that is performed (=0 when full analysis [overall population] and >0 when conducting a subgroup analysis). **Format:** Numeric
STRATUMNAME	Indicates name of each strata for subgroup analyses. Should indicate "Overall" for overall analyses; will be populated with CATEGORIZATION values from the new ANALYTICSUBGROUPS input file for subgroup

Variable	Description
	analyses, "Percentiles" for percentile analysis, and "Overall Unconditional" for unconditional analysis **Format:** Character(50)
MATCHID	Unique match identifier. Will be set to percentile for percentile risk sets and missing for risk sets using the overall population **Format**: Character(20)
CASE_EXPOSURE	Exposure status of the case **Format:** Numeric
EXPOSUREPROBABILITY	Probability of exposure among patients in the risk set (including the case) **Format:** Numeric
[FOLLOWUPTIME_ITT (OR FOLLOWUPTIME_ASTREATED)	Number of days between start of exposure and [EVENT_ITT or EVENT_ASTREATED] for the risk set **Format:** Numeric

c) [*RUNID*]_riskdiffdata_[*COMP_ORDER*] _ [*PERIODID*].sas7bdat

This output table contains de-identified, aggregated data that will be used to calculate unadjusted and adjusted risk differences for the overall sample , the matched sample, the overall sample stratified by percentiles, each pre-specified subgroup, and the matched sample not stratified on matchid (when requested). Table 124 contains specifications for this output table.

Table 124. [RUNID] _ riskdiffdata_ [COMP_ORDER] _ [PERIODID].sas7bdat

Variable	Description
TYPE	Indicates if the records are for the matched cohort or for all patients. **Format:** Character(20)
COVARNUM	The level of analysis that is performed (=0 when full analysis [overall population] and >0 when conducting a subgroup analysis). **Format:** Numeric
STRATUMNAME	Indicates name of each strata for subgroup analyses. Should indicate “Overall” for overall analyses; will be populated with CATEGORIZATION values from the new ANALYTICSUBGROUPS input file for subgroup analyses, "Percentiles" for percentile analysis, and "Overall Unconditional" for unconditional analysis **Format:** Character(50)
PERCENTILEVALUE	Set to Percentile for Percentile analysis, 0 for all other analyses **Format:** Numeric
EXP	The number of patients in the exposed group **Format:** Numeric
UNEXP	The number of patients in the unexposed group **Format:** Numeric
EVEXP	The number of events in the exposed group **Format:** Numeric

Variable	Description
EVUNEXP	The number of event in the unexposed group **Format:** Numeric
FUTIMEEXP	The total number of follow-up days for the exposed group **Format:** Numeric
FUTIMEUNEXP	The total number of follow-up days for the unexposed group **Format:** Numeric
WEIGHT	The sum of the weights across strata **Format:** Numeric
WEIGHT_DIFF	The sum of the weighted differences across strata **Format:** Numeric

d) [*RUNID*]_survivaldata_[*COMP_ORDER*]_[*PERIODID*].sas7bdat

This output table contains de-identified, aggregated data that can be used for the creation of a Kaplan Meier plot for the entire sample, the matched sample, the entire sample stratified by percentiles, each subgroup, and the matched sample not stratified on matchid (when requested). This output table is not produced for Type 4 analyses, as time to event is not currently available in the pregnancy tool. Table 125 contains specifications for this output table.

Table 125. [RUNID]_survivaldata_[COMP_ORDER]_[PERIODID].sas7bdat

Variable	Description
TYPE	Indicates if the records are for the matched cohort or for all patients. Values include: • AllPts: all patients in the cohort • Matched: all matched patients in the cohort **Format:** Character(20)
COVARNUM	The level of analysis that is performed (=0 when full analysis [overall population] and >0 when conducting a subgroup analysis). **Format:** Numeric
STRATUMNAME	Indicates name of each strata for subgroup analyses. Values include: • Overall: indicates analyses within the entire base population, and conditional analysis within entire base population for the matched cohort (depends on value of TYPE) • Overall Unconditional: indicates unconditional analysis within entire base population for the matched cohort • [*Stratum*]: indicates analyses within levels of the specified subgroup for the entire base population, and conditional analysis within levels of the specified subgroup ([*Stratum*] for the matched cohort (depends on value of TYPE). *Stratum* values will represent CATEGORIZATION values from the ANALYTICSUBGROUPS input file) • [*Stratum*] Unconditional: indicates unconditional analysis within levels of the specified subgroup ([*Stratum*] will be CATEGORIZATION values from the ANALYTICSUBGROUPS input file)

Variable	Description
	• Percentiles: indicates percentile stratification analysis **Format:** Character(50)
FOLLOWUPDAY	Number of days after exposure **Format:** Numeric
PERCENTILEVALUE	Set to Percentile for Percentile analysis, 0 for all other analyses **Format:** Numeric
NEXP	Number of exposed **Format:** Numeric
EVEXP	Number of events among the exposed **Format:** Numeric
NUNEXP	Number of unexposed **Format:** Numeric
EVUNEXP	Number of events among the unexposed **Format:** Numeric

e) [*RUNID*]_matched_tables_[*COMP_ORDER*]_[PERIODID].sas7bdat

A SAS dataset is output to include information on the unmatched and matched cohort comparing demographic, predefined covariate, and utilization measures and distributions between exposure and comparison groups. The intent of these tables is to compare the results of the unmatched to the 1:1 and/or 1:*n* variable matched cohorts to see how well matching balanced the exposure and comparator groups. Note that for 1:*n* variable matched cohorts, the values included are weighted. Table 126 contains specifications for this output table.

Table 126. [RUNID]_matched_tables_[COMP_ORDER]_[PERIODID]

Variable	Description
COMP_ORDER	Numeric identifier to differentiate exposure/comparator pairs and parameter settings. Defined in Comparison File. **Format:** Character(30)
GROUP1	Indicates the exposure group. Same as COMP parameter value in Comparison File. **Format:** Character(30)
GROUP2	Indicates the comparator group. Same as CONTROL parameter value in Comparison File. **Format:** Character(30)
TABLE	Indicates whether the table 1 statistics represent the "Unmatched" or "Matched" population **Format:** Character(30)
METVAR	Name for each characteristic. • Total patients: TOTAL

Variable	Description
	• Event count: EVENT_ASTREATED or EVENT_ITT (if ITTDAYS is populated in the Type 2 File) • Follow-up time in days: FOLLOWUPTIME_ASTREATED or FOLLOWUPTIME_ITT (if ITTDAYS is populated in the Type 2 File) • Age, categorical: AGE###### (where ###### describes each age grouping specified by AGESTRAT in Main Program Parameters, without dashes, e.g., AGESTRAT= 00-19 in AGESTRAT will appear as AGE0019 in this file). • Age, continuous: AGE • Sex, female: FEMALE • Sex, male: MALE • Sex, unknown: SEX_UNKNOWN • Sex, ambiguous: AMBIGUOUS • Race, Unknown: R_UNKNOWN • Race, American Indian or Alaska Native: R_AIAN • Race, Asian: R_ASIAN • Race, Black or African American: R_BLACK • Race, Native Hawaiian or Other Pacific Islander: R_NHPI • Race, White: R_WHITE • Hispanic: HISPANIC_YES • Hispanic: HISPANIC_NO • Hispanic: HISPANIC_UNKNOWN • Year, *n*: YEAR_*n* (for ex: Year, 2005 would be YEAR_2005) • Comorbidity score: COMORBIDSCORE • Number of AV encounters: NUMAV • Number of ED encounters: NUMED • Number of IP encounters: NUMIP • Number of IS encounters: NUMIS • Number of OA encounters: NUMOA • Number of unique dispensings: NUMRX • Number of unique generics dispensed: NUMGENERIC • Number of unique drug classes dispensed: NUMCLASS • Requester-defined covariate indicator, as defined by COVARNUM in Covariate Codes File: COVAR### • Mahalanobis distance: MAHALANOBIS Characteristics output for Type 4 analyses only: • Gestational age group at birth (PRE): PREPOSTIND_PRETERM • Gestational age group at birth (Term): PREPOSTIND_TERM • Gestational age group at birth (Post): PREPOSTIND_POSTTERM • Gestational age group at birth (None): PREPOSTIND_NONE • Mean gestational age at birth, continuous: GA_BIRTH • Mean enrollment time after birth: BIRTH_ENROLL

Variable	Description
	• Mean difference between date of birth and date of enrollment of infant: ENROLL_DIFF *The following is produced when an exposed control cohort is the comparator:* • Mean gestational age of first exposure, weeks, continuous: GA_FIRST • Mean number of dispensings overlapping the pre-pregnancy period, continuous: Adjusteddisp_PRE • Mean number of adjusted code counts overlapping the first trimester, continuous: AdjustedDisp_T1 • Mean number of adjusted code counts overlapping the second trimester, continuous: AdjustedDISP_T2 • Mean number of adjusted code counts overlapping the third trimester, continuous: AdjustedDISP_T3 • Exposed during first trimester: EXP_T1 • Exposed during second trimester: EXP_T2 • Exposed during third trimester: EXP_T3 • Exposed in pre-pregnancy days before pregnancy start: EXP_PRE **Format:** Character(30)
VARTYPE	Indicates whether the METVAR is "dichotomous" or "continuous." **Format:** Character(30)
EXP_MEAN	Mean value for continuous variables, count for dichotomous variables among the exposure group **Format:** Numeric(8)
EXP_STD	Standard deviation for continuous variables, percent for dichotomous variables among the exposure group **Format:** Numeric(8)
EXP_S2	Weighted standard deviation used in standardized difference calculation for exposure group **Format:** Numeric(8)
COMP_MEAN	Mean value for continuous variables, count for dichotomous variables among the comparison group **Format:** Numeric(8)
COMP_STD	Standard deviation for continuous variables, percent for dichotomous variables among the comparison group **Format:** Numeric(8)
COMP_S2	Weighted standard deviation used in standardized difference calculation for comparison group **Format:** Numeric(8)

Variable	Description
AD	Absolute difference between exposure and comparison group means (for continuous metrics) and proportions (for dichotomous metrics). **Format:** Character(30)
SD	Standardized difference between exposure and comparison group means (for continuous metrics) and proportions (for dichotomous metrics). **Format:** Numeric(8)
EXP_W	Sum of weights among the exposure group. Will equal the number of exposed subjects. Note on weights: With variable ratio matching , each exposed subject is matched to a variable number of comparator subjects. The weight for each treated subject equals 1 ($w_i = 1$ for $i = 1, \ldots, n_t$), the weight for each control subject equals the inverse of the matching ratio for that specific matched set. Suppose there exists a matched set with 4 subjects ("A" from the treatment group and "B", "C", and "D" from the control group), then the x value for subject A is weighted by 1 and the x values for subjects B, C, and D are all weighted by 1/3. For exposure group, and for both exposure and comparison groups in 1:1 match, this will equal the number of individuals (since each individual has a weight of 1). **Format:** Numeric(8)
EXP_W2	Sum of weights squared among the exposure group **Format:** Numeric(8)
COMP_W	Sum of weights among the comparison group. Note on weights: With variable ratio matching , each exposed subject is matched to a variable number of comparator subjects. The weight for each treated subject equals 1 ($w_i = 1$ for $i = 1, \ldots, n_t$), the weight for each control subject equals the inverse of the matching ratio for that specific matched set. Suppose there exists a matched set with 4 subjects ("A" from the treatment group and "B", "C", and "D" from the control group), then the x value for subject A is weighted by 1 and the x values for subjects B, C, and D are all weighted by 1/3. **Format:** Numeric(8)
COMP_W2	Sum of weights squared among the comparison group **Format:** Numeric(8)

f) [*RUNID*]_timing.sas7bdat

A SAS dataset containing metrics on execution time for each comparison and time period (*e.g.,* the time the program to took to run with respect to HD variable selection (HDPSRUNTIME), matching (MATCHRUNTIME), total adjustment time (ADJUSTMENTRUNTIME), and time to create tables and figures (TABLETIME)).

g) [*RUNID*]_varinfo_[*COMP_ORDER*]_[*PERIODID*].sas7bdat

This dataset contains all covariates considered for high-dimensional variable selection. It contains a description of the variables (*i.e.,* what code or what utilization metric it represents) as well as an indicator for which data dimension it came from (drug class, ICD-9-CM diagnosis, ICD-10-CM diagnosis, ICD-9-CM procedure, ICD-10-CM procedure, CPT, HCPCS). It also contains statistics used to determine variable selection and an indicator for whether or not the variable was selected into the hdPS model.

h) [*RUNID*]_estimates_[*COMP_ORDER*]_[PERIODID].sas7bdat

The dataset contains the c-statistic and an indicator for whether or not the model converged for the model run (predefined, empirically identified, or predefined + empirically identified).

i) [*RUNID*]_corr_[*COMP_ORDER*]_[PERIODID].sas7bdat

This dataset is output by request (Diagnostics = 1 in run_programs.sas). The dataset includes the pair-wise correlation coefficients between each variable in the PS estimation model (Table 127).

Table 127. [RUNID]_corr_[COMP_ORDER]_[PERIODID]

Variable	Description
VAR1	Name of 1st variable **Format:** Character(13)
VAR2	Name of 2nd variable **Format:** Character(13)
CORR	Correlation coefficient between VAR1 and VAR2 **Format:** Numeric
COMP	Indicates the exposure group. **Format:** Character(30)
CONTROL	Indicates the comparator group. **Format:** Character(30)
COMP_ORDER	Numeric identifier to differentiate exposure/comparator pairs and parameter settings. Defined in Comparison File. **Format:** Numeric

j) [*RUNID*]_runreg_[*COMP_ORDER*]_[PERIODID].rtf

This file is output by request (Diagnostics = 1 in run_programs.sas). The dataset includes output from the PS estimation model, including parameter estimates, and intercept adjusted collinearity diagnostics tables.

k) [*RUNID*]_runmodel_out_[*COMP_ORDER*]_[PERIODID].rtf

This file is output by request (Diagnostics = 1 in run_programs.sas). The dataset includes output from the PS estimation model, including descriptive statistics for continuous variables, frequency distributions of class variables, the maximum likelihood iteration history, odds ratio estimates, and the association of predicted probabilities and observed responses tables.

l) [*RUNID*]_psdistribution_out_[*COMP_ORDER*]_[PERIODID].rtf

When ANALYSIS=PS, this output file should be created and output to the *msoc* folder for each comparison in the Comparison File, time period, and execution of the PSA tool. Table 123 contains specifications for this output table.

Table 128. [*RUNID*]_psdistribution_[*COMP_ORDER*]_[PERIODID]

Parameter	Field Name	Description
Cohort Name	GROUP	**Details:** cohort name. **Format:** Character(40)
Population (unmatched and matched)	TYPE	Indicates if the records are for the matched cohort or for all patients. **Note:** For PS distribution – type = Matched will only be returned for 1:1 analysis **Format:** Character(20)
PS category	PS_CAT	**Details**: propensity score category. Values are in 40 bins of equal size (0.025). **Format:** Numeric (8)
Number of Patients	NPTS	**Details**: number of patients in each PS_CAT. **Note**: if there are no patients within a value of PS_CAT, NPTS=0 (instead of having a missing PS_CAT value). **Format:** Numeric (8)

m) [*RUNID*]_signature_[analysis]_[*COMP_ORDER*]_[PERIODID].sas7bdat

The [*RUNID*]_signature_[analysis]_[*COMP_ORDER*]_[PERIODID] output table contains metadata associated with the request, including request identifiers, program identifiers, database version, and run time metrics. Table 129 contains specifications for the [*RUNID*]_signature_[analysis]_[*COMP_ORDER*]_[PERIODID] output table.

Table 129. [*RUNID*]_signature_[analysis]_[*COMP_ORDER*]_[PERIODID] Output

Variable	Description
VAR	Metric name. **Format:** Character(15)
VALUE	Metric value. **Format:** Character(200)

2. DPLOCAL Folder

a) [*RUNID*]_matched_[*COMP_ORDER*]_[PERIODID].sas7bdat

A SAS dataset is output for ps estimation and matching strategy and time period. This is the same matched dataset as the dataset of the same name in *msoc* folder output, except it contains additional variables that may be useful for follow-up queries. The additional variables in this dataset are PATID, INDEXDT, EVENTDT, (patient identifier) and each of the high-dimensional variables selected. This file is always produced regardless of the value of the INDLEVEL parameter.

b) [*RUNID*]_scores_[*COMP_ORDER*]_[PERIODID].sas7bdat

A SAS dataset is output for every propensity score estimation and matching strategy and time period. This is the same matched dataset as [*RUNID*]_matched_[*COMP_ORDER*]_[PERIODID], except it does not contain match identifiers.

c) [RUNID]_estimates_[COMP_ORDER]_[PERIODID].sas7bdat

The dataset contains the c-statistic and an indicator for whether or not the model converged for the model run (predefined, empirically identified, or predefined + empirically identified). This is the same dataset as the dataset of the same name in *msoc* folder output.

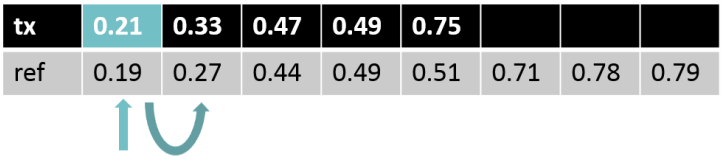
C. MATCHING ALGORITHM FUNCTION FOR 1:1 MATCHING

The propensity score matching algorithm creates matched pairs between patients in the treatment group and a comparator group (here referred to as the "reference group"). The matching algorithm works by locating patients from the treatment and reference groups with similar propensity scores, which appear close to each other on an ordered list. The distance between the propensity scores of the patients in the treatment group and reference group must fall within the specified caliper. In the example that follows, the caliper is 0.05.

1. Creating Potential Matched Pairs

Once the propensity scores from both groups have been ordered, the algorithm creates potential matched pairs. For each patient in the treatment group (**tx**) the algorithm searches the reference group (**ref**) for the first patient with a score greater than or equal to the **tx** patient.

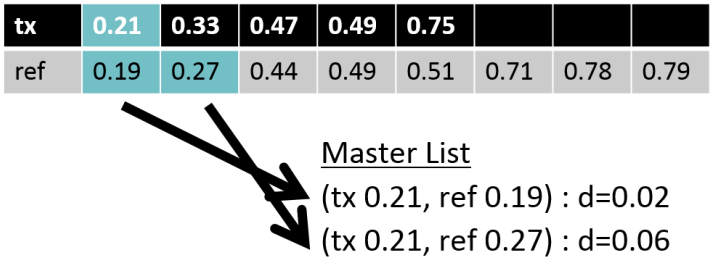
Figure 41. Algorithm seeks potential matched pairs in reference group



In Figure 41, the first patient from the **tx** group has a propensity score of 0.21. When the algorithm searches through the patients from the **ref** group, it looks for the first patient with a propensity score greater than or equal to 0.21. It first finds a patient with a propensity score of 0.19. This is less than the **tx** propensity score of 0.21, so the algorithm continues to search. The next-closest propensity score in the **ref** group is 0.27. This is larger than the **tx** propensity score, so the algorithm stops at this patient. The algorithm takes this **ref** score and the one immediately preceding it and creates two potential pairs, **tx** 0.21, **ref** 0.19 and **tx** 0.21, **ref** 0.27, which are moved to a master list of potential matched pairs.

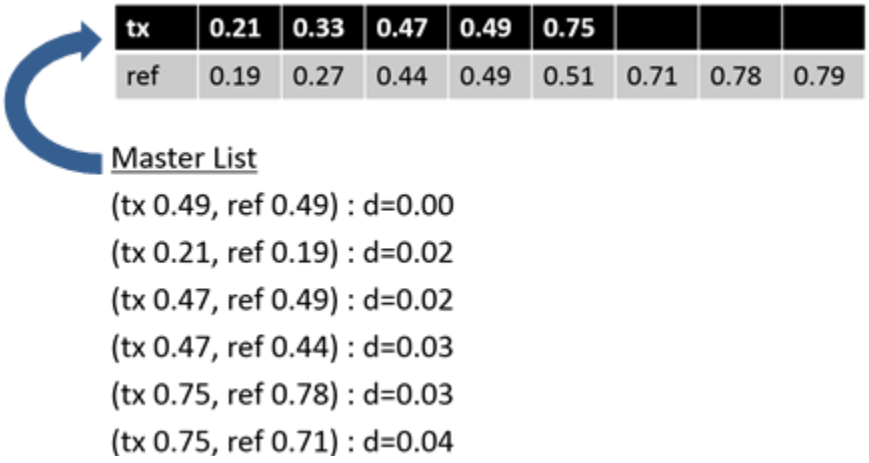
The difference between the **tx** and **ref** scores in a potential matched pair cannot be larger than the specified caliper (0.05). The difference between the first pair (**tx** 0.21, **ref** 0.19) is 0.02, and the difference between the second pair (**tx** 0.21, **ref** 0.27) is 0.06. The difference of the first pair falls within the limits of the caliper, and is added to the master list of potential matched pairs. The difference of the second pair is larger than the caliper, so it is not added to the master list (Figure 42).

Figure 42. Algorithm includes within caliper matches in the Master List



The algorithm repeats this process for every score from the **tx** group, adding all pairs within the caliper to the master list of potential matched pairs. The pairs in the master list are then ordered by distance.

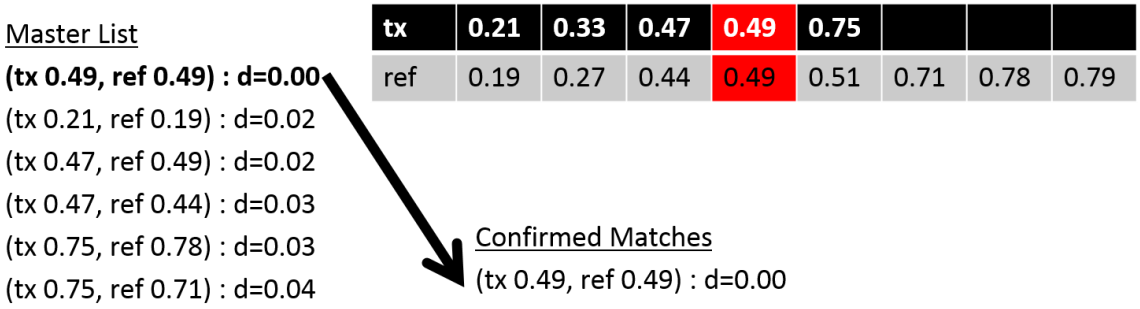
Figure 43. Algorithm generates a list of potential matched pairs, ordered by distance



2. Matching Propensity Scores

After the master list of potential matched pairs is ordered by distance, the algorithm looks at the pair at the top of the list. If both the **tx** and **ref** scores in the potential pair are unmatched, it is moved to the list of confirmed matches. These **tx** and the **ref** patients cannot be matched again.

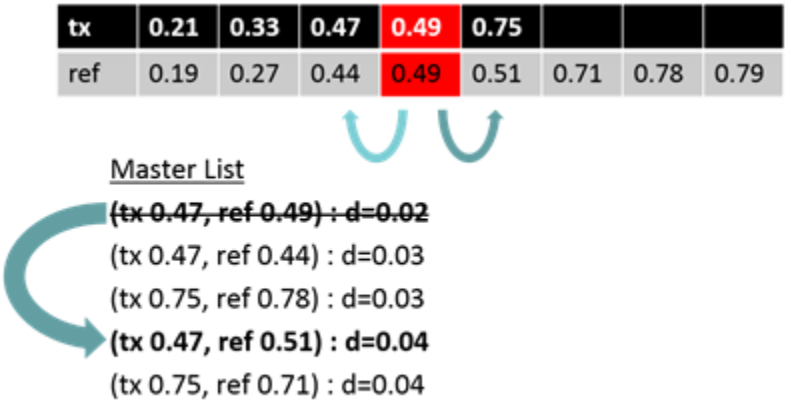
Figure 44. Confirmed matches are created from potential matched pairs



In Figure 44, the potential pair at the top of the master list is **tx** 0.49, **ref** 0.49. Neither the **tx** or the **ref** score in this pair been matched, so it is moved into the list of confirmed matches, and the algorithm returns to the top of the master list.

If the potential pair contains a **tx** score that is unmatched, but a **ref** score that has already been matched, the algorithm finds the nearest **ref** neighbor, to either the left or the right, that is still available, and this pair is added to the master list. In the case that the **ref** score is unmatched and the **tx** score is no longer available, the algorithm seeks the nearest **tx** neighbor that is still available (Figure 45).

Figure 45. Algorithm creates a new potential pair with nearest unmatched neighbor

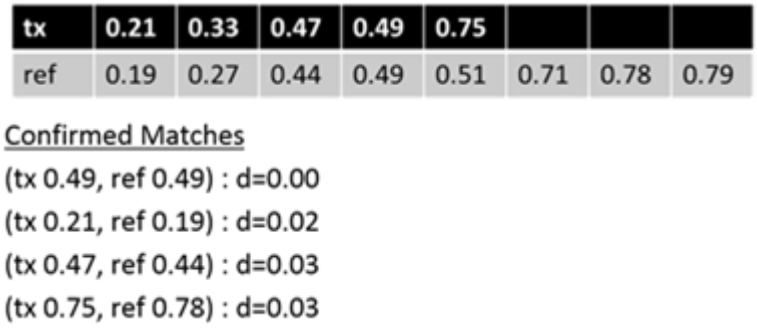


In Figure 45, the pair at the top of the list is **tx** 0.47, **ref** 0.49. Here, **tx** 0.47 is unmatched, but **ref** 0.49 has already been matched. The algorithm finds the next closest propensity score from the **ref** group, **ref** 0.51, and the new potential pair of **tx** 0.47, **ref** 0.51 is added to the master list.

3. Algorithm Output

The algorithm continues to return to the top of the master list until all of the **tx** scores have been matched. The algorithm then ends, leaving a final list of confirmed matches (Figure 46). Not all patients will be matched by this algorithm. Once several patients have been matched to others whose propensity scores have a very small distance, the propensity scores of remaining patients might not have a distance close enough to each other to fall within the caliper. In Figure 46, the patients from **tx** with the propensity scores 0.21, 0.47, 0.49, and 0.75 were all matched with patients from **ref**. The score of the remaining **tx** patient is 0.33, which is not close enough in distance to any of the remaining **ref** group scores to be matched.

Figure 46. Final list of confirmed matches



4. Matching Algorithm Function for 1:*n* Matching

The process described above for 1:1 matching is repeated multiple times until there are no more potential matches left, or until all treatment patients are matched to ten comparator patients. The number of 1:1 matches will be ≥ 1:2 matches ≥ 1:3 matches ≥ 1:4 matches … ≥ 1:*n* matches.

XI. APPENDIX D: MFM TOOL TECHNICAL DOCUMENTATION

The MFM tool is designed to be executed following the execution of the CIDA tool. The CIDA tool identifies and extracts cohorts of interest based on requester-defined parameters, and generates output tables in the *msoc* and *dplocal* folders that are required for subsequent processing with the MFM tool.

This technical specification document details the lookup tables, program parameters and input files that must be specified to execute the MFM tool. These parameters and files should be included in the program package distributed *in addition* to those necessary for CIDA tool execution, if a multiple factor matching analysis is requested.

A. LOOKUP TABLES, PROGRAM PARAMETERS, AND INPUT FILES

1. Lookup Tables

There are no lookup tables required for the execution of the MFM tool.

2. Main Program Parameters

There are five main program parameters that must be specified. These include defining matching variables, matching ratio, and specifying the name of all input files. These parameter values should be set in a program called run_programs.sas, located in the *inputfiles* folder. Note that all main programs parameters specified are fixed for a single execution of the program. Table 125 contains detailed specifications for main program parameters.

Table 130. MFM Tool Main Program Parameters

Parameter	Field Name	Description
Multiple Factor Matching File	MFMFILE	**Details:** Contains parameters for perfoming multiple factor matching **Note 1:** If MFMFILE is specified, COMPARISON input parameter should be blank **Defined by:** Request programmer **Input type:** Optional **Format:** .sas7bdat file **Example:** MFMFILE= wp014_MFM
Indicator for Individual Level Output Return	INDLEVEL	**Details:** Specifies whether to write individual level [*RUNID*]_matched_[*COMP*]_[*Look*] files to the SOC subfolder. Allowable values: • **Y:** Yes • **N:** No **Defined by:** Requester **Input type:** Required **Format:** Binary **Example:** IndLevel = Y

Parameter	Field Name	Description
Return unconditional analysis	UNCONDITIONAL	**Details:** Specifies whether to create risk sets not stratified by matchID. This allows SOC to perform unconditional analyses. Allowable values: • **Y:** Yes • **N:** No **Defined by:** Requester **Input type:** Required **Format:** Binary **Example:** UNCONDITIONAL = Y

3. Input Files

This section describes all input files that must be included in the *inputfiles* folder in the program package to execute the MFM tool. These files are *in addition* to those required for execution of the CIDA tool (including the Medical Utilization, Comorbidity Score, and Covariate Codes files, which are optional for the CIDA tool but required if baseline characteristics tables are requested for a multiple factor matched analysis).

a) Multiple Factor Matching File

The Multiple Factor Matching File is required. The file allows requesters to specify all exposure/comparator pairs that should be evaluated in the multiple factor matching analysis. Each exposure/comparator pair is assigned a unique COMP_ORDER value, to differentiate pairs in output tables. Table 131Table 131 contains specifications for this file.

Table 131. Multiple Factor Matching File Specification

Parameter	Field Name	Description
Comparison Identifier	COMP_ORDER	**Details:** numeric identifier to differentiate exposure/comparator pairs and parameter settings in output tables. **Note 1:** COMP_ORDER (&COMP value) is used as a suffix in output tables to identify each comparison. **Defined by:** Request programmer **Input type:** Required (cannot be left blank) **Format:** Numeric **Example:** 2
Exposure of Interest	COMP	**Details:** GROUP name of the exposure of interest. **Defined by:** Requester **Input type:** Required (cannot be left blank) **Format:** Character(30) **Example:** drug_a

Parameter	Field Name	Description
Comparator of Interest	CONTROL	**Details:** GROUP name of the comparator of interest. **Note 1:** When the never-exposed cohort serves as the comparator of interest, the group name should be <COMP>_nvrexp **Defined by:** Requester **Input type:** Required (cannot be left blank) **Format:** Character(30) **Example:** drug_b
Matching Ratio	RATIO	**Details:** identifies the matching ratio. Allowable values: • F: fixed 1:1 matching • V: variable 1:10 matching **Defined by:** Requester **Input type:** Required (cannot be left blank) **Format:** Char (1) **Example:** F
Categorical variables to match on	MATCHVARS	**Details:** used to specify variables to match the COMP and CONTROL groups on. Separate multiple factors by a space. Allowable values include: • Sex • AgeGroup • Year **Named by:** Requester **Input type:** Required **Format:** Character **Example:** Sex AgeGroup Year

b) Analytic Subgroups File

The Analytic Subgroups file is optional. The file allows requesters to specify all subgroups and subgroup levels for each comparison (i.e., exposure/comparator/ratio/caliper/model combination under investigation). The file also allows the user to select the base population for subgroup analyses (i.e., total population or successfully matched population from the primary analysis). Table 132 contains specifications for this file.

Table 132. Analytic Subgroups File Specification

Parameter	Field Name	Description
Comparison Identifier	COMP_ORDER	**Details:** numeric identifier to differentiate exposure/comparator pairs and parameter settings in output tables. **Note 1:** COMP_ORDER (&COMP value) is used as a suffix in output tables to identify each comparison. **Defined by:** Request programmer **Input type:** Required (cannot be left blank) **Format:** Numeric **Example:** 2
Covariate Number	COVARNUM	**Details:** numeric identifier to indicate covariates needed for subgroup analyses. For binary covariates specified in the COVARIATECODES input file, enter the COVARNUM value from COVARIATECODES here. For covariates not explicitly named in the COVARIATECODES file, use the following values: • **Sex**: 1000 • **Age**: 1001 • **Year**: 1002 • **Time**: 1003 • **Comorbidscore**: 1004 • **NumIP:** 1005 • **NumIS:** 1006 • **NumED:** 1007 • **NumAV:** 1008 • **NumOA:** 1009 • **NumRx:** 1010 • **NumGeneric**: 1011 • **Race:** 1012 • **Hispanic:** 1013 **Input type:** Required (cannot be left blank) **Format:** Numeric **Example:** 1000
Covariate Level Specification	CATEGORIZATION	**Details:** specifies how subgroup variables should be categorized for subgroup analysis.

Parameter	Field Name	Description
		Note 1: for COVARNUM=1-999, CATEGORIZATION should be left blank as these are dichotomous variables. **Note 2:** for COVARNUM=1000 (sex), this should be left blank. Sex values of F, M, U, and A will automatically be categorized **Note 3:** for COVARNUM=1001 (Age), valid CATEGORIZATION values should follow the same rules as the AGESTRAT macro parameter in the CIDA tool (e.g., 40-59 60-79 80-99). However, the CATEGORIZATION value does not have to be the same as the AGESTRAT value used in CIDA (i.e., subgroup analyses can specify different age groups). **Note 4:** for • Year (COVARNUM=1002) • Time (COVARNUM=1003) • Comorbidscore (COVARNUM=1004) • NumIP (COVARNUM=1005) • NumIS (COVARNUM=1006) • NumED (COVARNUM=1007) • NumAV (COVARNUM=1008) • NumOA (COVARNUM=1009) • NumRx (COVARNUM=1010) • NumGeneric (COVARNUM=1011) Groups must be separated by a space. To leave the first group open-ended, use "low-". In the output "low-", will be replaced with "<=". If a negative is desired as the upper bound of a group, do not include a space in the group (*e.g.,* use low--1 for low to -1). To leave the last group open-ended, use "-high". In the output "-high", will be replaced with ">=". Note that groups should not have overlapping values. In the event that overlapping values are entered, the value will be mapped to the first group in the list. For example, if age groups are specified "45-85 80-99", patients 80-85 will be included in the 45-85 category (and re-matched within that group). **Note 5:** the theoretical range of comorbidity scores is -2 -26. **Defined by:** Requester **Input type:** Required **Format:** Alphanumeric; SAS character $100. **Example:** low-0 1 2-3 4-7 8+

Parameter	Field Name	Description
Subgroup Matching Base Population Specification	MATCHEDINFULLONLY	**Details:** specifies if the base population for subgroup re-matching should be restricted to the matched population. **Y:** subgroup re-matching should be restricted to the matched population **N:** subgroup re-matching should be done in the total population (i.e., patients who were matched and patients who were not matched in the main analysis) **Input type:** Required **Format:** Character **Example:** N

B. OUTPUT

1. MSOC Folder

The following output files are created and output to the *msoc* folder for each comparison in the Multiple Factor Matching File, time period, and execution of the MFM tool:

- [*RUNID*]_matched_[*COMP_ORDER*]_[*PERIODID*].sas7bdat
- [*RUNID*]_risksetdata_[*COMP_ORDER*]_[*PERIODID*].sas7bdat
- [*RUNID*]_riskdiffdata_[*COMP_ORDER*]_[*PERIODID*].sas7bdat
- [*RUNID*]_survivaldata_[*COMP_ORDER*]_[*PERIODID*].sas7bdat
- *[RUNID]_matche*d_tables_[*COMP_ORDER*]_[*PERIODID*].sas7bdat
- [*RUNID*]_signature_ps_[*COMP_ORDER*]_[*PERIODID*].sas7bdat

Where RUNID is the request programmer-defined execution identifier, COMP_ORDER is the unique ps estimation and matching strategy identifier, and PERIODID is the time period.

a) [*RUNID*]_matched_[*COMP_ORDER*]_[*PERIODID*].sas7bdat

This output table contains the final matched sample, with identifiers for all matches performed. This table is only produced if INDLEVEL = Y in the MFM Tool Main Program Parameters. Table 133 contains specifications for this output table.

Table 133. [RUNID] _matched_ [COMP_ORDER] _ [PERIODID].sas7bdat

Variable	Description
COVARNUM	The level of analysis that is performed (=0 when full analysis [overall population] and >0 when conducting a subgroup analysis). **Format:** Character(4)
STRATUMNAME	Indicates name of each strata for subgroup analyses. Should indicate "Overall" for overall analyses; will be populated with CATEGORIZATION values from the new ANALYTICSUBGROUPS input file for subgroup analyses, "Percentiles" for percentile analysis, and "Overall Unconditional" for unconditional analysis **Format:** Character(50)
STUDYCLASS	Analogous to the GROUP value on ADS_MSTR_ [*PERIODID*] input file. **Format:** Character(30)
AGE_CAT	Requester-defined age category. Determined by AGESTRAT macro parameter values defined in the CIDA tool. **Format:** Character(9)
SEX	Patient Sex as defined in SDD Demographic table. **Format:** Character(1)
RACE	Patient Race as defined in SDD Demographic table. **Format:** Character(1)
HISPANIC	Hispanic as defined in SDD Demographic table. **Format:** Character(1)
EVENT_ITT (OR EVENT_ASTREATED)	Boolean indicator of whether or not patient had the HOI during the time period (PERIODID). This is set to 1 when the EVENTDT is the same as the end of follow-up *for* the patient (i.e. when the follow-up is ended by an HOI). **Note 1:** variable name is dependent on whether exposed time was defined using dispensings' days supply (EVENT_ASTREATED; ITTDAYS parameter set to missing in CIDA tool Type 2 File) or a requested-defined number of days (EVENT_ITT; ITTDAYS parameter populated in CIDA tool Type 2 File). **Format:** Numeric
FOLLOWUPTIME_ITT (OR FOLLOWUPTIME_ASTREATED)	Number of days that the patient is followed for either HOI or censoring. Variable is calculated using the earliest of 1) end of exposure period; 2) disenrollment; 3) any additional censoring criteria [defined using CIDA tool]; and 4) HOI date. **Note 1:** variable name is dependent on whether exposed time was defined using dispensings' days supply (FOLLOWUPTIME_ASTREATED; ITTDAYS parameter set to missing in CIDA tool Type 2 File) or a requested-defined number of days (FOLLOWUPTIME_ITT; ITTDAYS parameter populated in CIDA tool Type 2 File). **Format:** Numeric

Variable	Description
[MATCHID VALUE]	MatchID for the analysis requested. Name of the variable is determined based on the analysis requested (i.e., specifications in the MULTIPLE FACTOR MATCHING file). Allowable values include : • **MFMSMATCHID1_1:** Multi-factor matching algorithm: 1:1 matching strategy • **MFMSMATCHID100_1:** Multi-factor matching algorithm: 1:10 matching strategy **Format:** Character(50)
LastLookFollowed	When an individual is lost to follow up (data is no longer being updated), this is the last look # the individual contributed data to. If an individual is still eligible to be followed in subsequent looks, this is set to 0. **Format:** Numeric
EXPOSURE	Boolean indicator of whether or not the STUDYCLASS value is the exposure of interest (COMP in the MULTIPLE FACTOR MATCHING file). **Format:** Numeric
TIME	Variable indicating during which time period the patient was selected (*e.g.,* TIME = 1 indicates the patient was selected in PERIODID 1, even if the current PERIODID being executed is not the first). **Format:** Numeric
NUMIP	Visit count for IP encounter type. One visit allowed per encounter, per day. **Note 1**: this metric is automatically calculated using the medical utilization module in the CIDA tool (Medical Utilization File). **Format:** Numeric
NUMED	Visit count for ED encounter type. One visit allowed per encounter, per day. **Note 1**: this metric is automatically calculated using the medical utilization module in the CIDA tool (Medical Utilization File). **Format:** Numeric
NUMIS	Visit count for IS encounter type. One visit allowed per encounter, per day. **Note 1**: this metric is automatically calculated using the medical utilization module in the CIDA tool (Medical Utilization File). **Format:** Numeric
NUMAV	Visit count for AV encounter type. One visit allowed per encounter, per day. **Note 1**: this metric is automatically calculated using the medical utilization module in the CIDA tool (Medical Utilization File). **Format:** Numeric
NUMOA	Visit count for OA encounter type. One visit allowed per encounter, per day.

Variable	Description
	Note 1: this metric is automatically calculated using the medical utilization module in the CIDA tool (Medical Utilization File). **Format:** Numeric
NUMGENERIC	Count of unique generic names dispensed during covariate evaluation window. **Format:** Numeric
NUMCLASS	Count of unique class names dispensed during covariate evaluation window. **Format:** Numeric
NUMRX	Count of dispensings during the covariate evaluation window. **Format:** Numeric
COMORBIDSCORE	Combined Charlson-Elixhauser Comorbidity Score (exact value). **Note 1:** this metric is calculated using the combined comorbidity score module in the CIDA tool (Comorbidity Score File). **Format:** Numeric
YEAR	Year of index date. **Format:** Numeric
COVAR1-COVAR*N*	Boolean indicators for each covariate (denoted by the value of the variable COVARNUM in the Covariate Codes File). **Format:** Numeric
MFMS	Indicates matching strata from 1 to *n* based matching groups in MATCHVARS parameter. **Format:** Numeric

b) [*RUNID*]_risksetdata_[*COMP_ORDER*]_[*PERIODID*].sas7bdat

This output table contains the risk sets for the entire sample, the matched sample, the entire stratified by percentiles, each subgroup, and the matched sample not stratified on matchid (when requested). Table 134contains specifications for this output table.

Table 134. [RUNID]_risksetdata_[COMP_ORDER]_[PERIODID].sas7bdat

Variable	Description
RISKSETID	Unique risk-set identifier **Format:** Numeric
TYPE	Indicates if the records are for the matched cohort or for all patients. **Format:** Character(20)
COVARNUM	The level of analysis that is performed (=0 when full analysis [overall population] and >0 when conducting a subgroup analysis). **Format:** Numeric
STRATUMNAME	Indicates name of each strata for subgroup analyses. Should indicate "Overall" for overall analyses; will be populated with CATEGORIZATION values from the new ANALYTICSUBGROUPS input file for subgroup analyses, and "Overall Unconditional" for unconditional analysis **Format:** Character(50)
MATCHID	Unique match identifier. Will be set to missing for risk sets using the overall population **Format**: Character(20)
CASE_EXPOSURE	Exposure status of the case **Format:** Numeric
EXPOSUREPROBABILITY	Probability of exposure among patients in the risk set (including the case) **Format:** Numeric
[FOLLOWUPTIME_ITT (OR FOLLOWUPTIME_ASTREATED)	Number of days between start of exposure and [EVENT_ITT or EVENT_ASTREATED] for the risk set **Format:** Numeric

c) [*RUNID*]_riskdiffdata_[*COMP_ORDER*] _ [*PERIODID*].sas7bdat

This output table contains de-identified, aggregated data that will be used to calculate unadjusted and adjusted risk differences for the overall sample , the matched sample, each pre-specified subgroup, and the matched sample not stratified on matchid (when requested). Table 135 contains specifications for this output table.

Table 135. [RUNID] _ riskdiffdata_ [COMP_ORDER] _ [PERIODID].sas7bdat

Variable	Description
TYPE	Indicates if the records are for the matched cohort or for all patients. **Format:** Character(20)
COVARNUM	The level of analysis that is performed (=0 when full analysis [overall population] and >0 when conducting a subgroup analysis). **Format:** Numeric
STRATUMNAME	Indicates name of each strata for subgroup analyses. Should indicate "Overall" for overall analyses; will be populated with CATEGORIZATION values from the new ANALYTICSUBGROUPS input file for subgroup analyses, and "Overall Unconditional" for unconditional analysis **Format:** Character(50)
PERCENTILEVALUE	Set to Percentile for Percentile analysis, 0 for all other analyses **Note 1:** Unable to do a percentile analysis for an MFM analysis. PERCENTILEVALUE will always = 0. **Format:** Numeric
EXP	The number of patients in the exposed group **Format:** Numeric
UNEXP	The number of patients in the unexposed group **Format:** Numeric
EVEXP	The number of events in the exposed group **Format:** Numeric
EVUNEXP	The number of event in the unexposed group **Format:** Numeric
FUTIMEEXP	The total number of follow-up days for the exposed group **Format:** Numeric
FUTIMEUNEXP	The total number of follow-up days for the unexposed group **Format:** Numeric
WEIGHT	The sum of the weights across strata **Format:** Numeric
WEIGHT_DIFF	The sum of the weighted differences across strata **Format:** Numeric

d) [*RUNID*]_survivaldata_[*COMP_ORDER*]_[*PERIODID*].sas7bdat

This output table contains de-identified, aggregated data that can be used for the creation of a Kaplan Meier plot for the entire sample, the matched sample, each subgroup, and the matched sample not stratified on matchid (when requested). Table 136This output table contains de-identified, aggregated data that can be used for the creation of a Kaplan Meier plot for the entire sample, the matched sample, the entire sample stratified by percentiles, each subgroup, and the matched sample not stratified on matchid (when requested). This output table is not produced for Type 4 analyses, as time to event is not currently available in the pregnancy tool. Table 125 contains specifications for this output table.
Table 125 contains specifications for this output table.

Table 136. [RUNID]_survivaldata_[COMP_ORDER]_[PERIODID].sas7bdat

Variable	Description
TYPE	Indicates if the records are for the matched cohort or for all patients. Values include: • AllPts: all patients in the cohort • Matched: all matched patients in the cohort **Format:** Character(20)
COVARNUM	The level of analysis that is performed (=0 when full analysis [overall population] and >0 when conducting a subgroup analysis). **Format:** Numeric
STRATUMNAME	Indicates name of each strata for subgroup analyses. Values include: • Overall: indicates analyses within the entire base population, and conditional analysis within entire base population for the matched cohort (depends on value of TYPE) • Overall Unconditional: indicates unconditional analysis within entire base population for the matched cohort • [*Stratum*]: indicates analyses within levels of the specified subgroup for the entire base population, and conditional analysis within levels of the specified subgroup ([*Stratum*] for the matched cohort (depends on value of TYPE). *Stratum* values will represent CATEGORIZATION values from the ANALYTICSUBGROUPS input file) • [*Stratum*] Unconditional: indicates unconditional analysis within levels of the specified subgroup ([*Stratum*] will be CATEGORIZATION values from the ANALYTICSUBGROUPS input file) **Format:** Character(50)
FOLLOWUPDAY	Number of days after exposure **Format:** Numeric
PERCENTILEVALUE	Set to Percentile for Percentile analysis, 0 for all other analyses **Note 1:** Unable to do a percentile analysis for an MFM analysis. PERCENTILEVALUE will always = 0. **Format:** Numeric

Variable	Description
NEXP	Number of exposed **Format:** Numeric
EVEXP	Number of events among the exposed **Format:** Numeric
NUNEXP	Number of unexposed **Format:** Numeric
EVUNEXP	Number of events among the unexposed **Format:** Numeric

e) [*RUNID*]_matched_tables_[*COMP_ORDER*]_[PERIODID].sas7bdat

A SAS dataset is output to include information on the unmatched and matched cohort comparing demographic, predefined covariate, and utilization measures and distributions between exposure and comparison groups. The intent of these tables is to compare the results of the unmatched to the 1:1 and/or 1:*n* variable matched cohorts to see how well matching balanced the exposure and comparator groups. Note that for 1:*n* variable matched cohorts, the values included are weighted. Table 137 contains specifications for this output table.

Table 137. [RUNID]_matched_tables_[COMP_ORDER]_[PERIODID]

Variable	Description
COMP_ORDER	Numeric identifier to differentiate exposure/comparator pairs and parameter settings. Defined in Multiple Factor Matching File. **Format:** Character(30)
GROUP1	Indicates the exposure group. Same as COMP parameter value in Multiple Factor Matching File. **Format:** Character(30)
GROUP2	Indicates the comparator group. Same as CONTROL parameter value in Multiple Factor Matching File. **Format:** Character(30)
TABLE	Indicates whether the table 1 statistics represent the "Unmatched" or "Matched" population **Format:** Character(30)
METVAR	Name for each characteristic. • Total patients: TOTAL • Event count: EVENT_ASTREATED or EVENT_ITT (if ITTDAYS is populated in the Type 2 File) • Follow-up time in days: FOLLOWUPTIME_ASTREATED or FOLLOWUPTIME_ITT (if ITTDAYS is populated in the Type 2 File) • Age, categorical: AGE###### (where ###### describes each age grouping specified by AGESTRAT in Main Program Parameters, without dashes, e.g., AGESTRAT= 00-19 in AGESTRAT will appear as AGE0019 in this file).

Variable	Description
	• Age, continuous: AGE • Sex, female: FEMALE • Sex, male: MALE • Sex, unknown: SEX_UNKNOWN • Sex, ambiguous: AMBIGUOUS • Race, Unknown: R_UNKNOWN • Race, American Indian or Alaska Native: R_AIAN • Race, Asian: R_ASIAN • Race, Black or African American: R_BLACK • Race, Native Hawaiian or Other Pacific Islander: R_NHPI • Race, White: R_WHITE • Hispanic: HISPANIC_YES • Hispanic: HISPANIC_NO • Hispanic: HISPANIC_UNKNOWN • Year, *n*: YEAR_*n* (for ex: Year, 2005 would be YEAR_2005) • Comorbidity score: COMORBIDSCORE • Number of AV encounters: NUMAV • Number of ED encounters: NUMED • Number of IP encounters: NUMIP • Number of IS encounters: NUMIS • Number of OA encounters: NUMOA • Number of unique dispensings: NUMRX • Number of unique generics dispensed: NUMGENERIC • Number of unique drug classes dispensed: NUMCLASS • Requester-defined covariate indicator, as defined by COVARNUM in Covariate Codes File: COVAR### **Format:** Character(30)
VARTYPE	Indicates whether the METVAR is "dichotomous" or "continuous." **Format:** Character(30)
EXP_MEAN	Mean value for continuous variables, count for dichotomous variables among the exposure group **Format:** Numeric(8)
EXP_STD	Standard deviation for continuous variables, percent for dichotomous variables among the exposure group **Format:** Numeric(8)
EXP_S2	Weighted standard deviation used in standardized difference calculation for exposure group **Format:** Numeric(8)
COMP_MEAN	Mean value for continuous variables, count for dichotomous variables among the comparison group **Format:** Numeric(8)

Variable	Description
COMP_STD	Standard deviation for continuous variables, percent for dichotomous variables among the comparison group **Format:** Numeric(8)
COMP_S2	Weighted standard deviation used in standardized difference calculation for comparison group **Format:** Numeric(8)
AD	Absolute difference between exposure and comparison group means (for continuous metrics) and proportions (for dichotomous metrics). **Format:** Character(30)
SD	Standardized difference between exposure and comparison group means (for continuous metrics) and proportions (for dichotomous metrics). **Format:** Numeric(8)
EXP_W	Sum of weights among the exposure group. Will equal the number of exposed subjects. Note on weights: With variable ratio matching , each exposed subject is matched to a variable number of comparator subjects. The weight for each treated subject equals 1 ($w_i = 1$ for $i = 1, \ldots, n_t$), the weight for each control subject equals the inverse of the matching ratio for that specific matched set. Suppose there exists a matched set with 4 subjects ("A" from the treatment group and "B", "C", and "D" from the control group), then the x value for subject A is weighted by 1 and the x values for subjects B, C, and D are all weighted by 1/3. For exposure group, and for both exposure and comparison groups in 1:1 match, this will equal the number of individuals (since each individual has a weight of 1). **Format:** Numeric(8)
EXP_W2	Sum of weights squared among the exposure group **Format:** Numeric(8)
COMP_W	Sum of weights among the comparison group. Note on weights: With variable ratio matching , each exposed subject is matched to a variable number of comparator subjects. The weight for each treated subject equals 1 ($w_i = 1$ for $i = 1, \ldots, n_t$), the weight for each control subject equals the inverse of the matching ratio for that specific matched set. Suppose there exists a matched set with 4 subjects ("A" from the treatment group and "B", "C", and "D" from the control group), then the x value for subject A is weighted by 1 and the x values for subjects B, C, and D are all weighted by 1/3. **Format:** Numeric(8)

Variable	Description
COMP_W2	Sum of weights squared among the comparison group **Format:** Numeric(8)

f) [*RUNID*]_signature_ps_[*COMP_ORDER*]_[PERIODID].sas7bdat

The [*RUNID*]_signature_ps_[*COMP_ORDER*]_[PERIODID] output table contains metadata associated with the request, including request identifiers, program identifiers, database version, and run time metrics. Table 138 contains specifications for the [*RUNID*]_signature_ps_[*COMP_ORDER*]_[PERIODID] output table.

Table 138. [*RUNID*]_signature_ps_[*COMP_ORDER*]_[PERIODID] Output

Variable	Description
VAR	Metric name. **Format:** Character(15)
VALUE	Metric value. **Format:** Character(200)

2. DPLOCAL Folder

a) [*RUNID*]_matched_[*COMP_ORDER*]_[PERIODID].sas7bdat

A SAS dataset is output for multiple factor matching estimation and matching strategy and time period. This is the same matched dataset as the dataset of the same name in *msoc* folder output, except it contains additional variables that may be useful for follow-up queries. The additional variables in this dataset are PATID, INDEXDT, EVENTDT, (patient identifier) and each of the high-dimensional variables selected. This file is always produced regardless of the value of the INDLEVEL parameter.

b) [*RUNID*]_scores_[*COMP_ORDER*]_[PERIODID].sas7bdat

A SAS dataset is output for every multiple factor matching estimation and matching strategy and time period. This is the same matched dataset as [*RUNID*]_matched_[*COMP_ORDER*]_[PERIODID], except it does not contain match identifiers.

XII. TABLE OF TABLES