

# **MINI-SENTINEL MODULAR PROGRAM 5: BACKGROUND RATES FOR HEALTH OUTCOMES OF INTEREST**

**Documentation version 2.1**

**Prepared by the Mini-Sentinel Operations Center**

**For use with Modular Program 5 version 2.1**

**February 06, 2013**

Mini-Sentinel is a pilot project sponsored by the [U.S. Food and Drug Administration \(FDA\)](#) to inform and facilitate development of a fully operational active surveillance system, the Sentinel System, for monitoring the safety of FDA-regulated medical products. Mini-Sentinel is one piece of the [Sentinel Initiative](#), a multi-faceted effort by the FDA to develop a national electronic system that will complement existing methods of safety surveillance. Mini-Sentinel Collaborators include Data and Academic Partners that provide access to health care data and ongoing scientific, technical, methodological, and organizational expertise. The Mini-Sentinel Coordinating Center is funded by the FDA through the Department of Health and Human Services (HHS) Contract number HHSF223200910006I.

## Table of Contents

1	Overview .....	- 1 -
2	Terminology .....	- 1 -
3	Program Summary .....	- 1 -
4	Program Parameter and Input File Specifications .....	- 2 -
4.1	Program Parameter Specifications .....	- 2 -
4.2	Input File Specifications .....	- 5 -
4.2.1	Query File .....	- 5 -
4.2.2	Incident Query File .....	- 9 -
5	Key Definitions .....	- 11 -
5.1	Enrollment Type Requirements .....	- 11 -
5.2	Prevalence-based Cohort.....	- 12 -
5.3	Incidence-based Cohort .....	- 12 -
5.3.1	Minimum Incidence (MIN) .....	- 12 -
5.3.2	Multiple Incidence (MULT).....	- 13 -
5.3.3	Examples .....	- 13 -
6	Denominators and Background Rates .....	- 14 -
7	Program Steps.....	- 15 -
8	Program Execution.....	- 15 -
9	Output Tables.....	- 16 -
10	Example.....	- 19 -

## Modification History

Version	Date	Modification	By
2.1	02/06/2013	<ul style="list-style-type: none"><li>Version incremented due to bug fix in 2.0 program</li></ul>	Mini-Sentinel Operations Center
2.0	10/17/2012	<ul style="list-style-type: none"><li>Original published version</li></ul>	Mini-Sentinel Operations Center

## 1 Overview

Mini-Sentinel modular programs (MPs) allow rapid implementation of standard queries across the Mini-Sentinel Distributed Database (MSDD). MPs are designed to run against the Mini-Sentinel Common Data Model (MSCDM).<sup>1</sup> They are written in SAS and can be customized using various parameter settings that define exposures, outcomes, date ranges, age ranges, and other implementation details. This document describes the key program specifications and main assumptions underlying each of the query parameters for Modular Program 5 (MP5) version 2.1. Program specification requirements, formats, and default values of all parameters are defined. A [sample program specification](#) is provided along with [output from a sample scenario](#).

## 2 Terminology

For simplicity, the term “scenario” is used throughout this document to refer to a set of parameters and criteria used to execute the MP. The term “requester” refers to an individual (or group of individuals) who issues the MP request and defines the scenarios. The term “request programmer” refers to an individual who creates request Input Files and executes the MP.

The terms “exposure” and “exposure of interest” represent exposure to a medical product, procedure, or condition as defined by the MP requester. An exposure can therefore be defined using any set of pharmacy, procedures or diagnosis codes found in the MSCDM. For example, exposure to a drug product dispensed in the outpatient setting can be defined as the occurrence of one or more National Drug Codes (NDCs) in the pharmacy dispensing file, whereas exposure to a vaccine can be defined based on observation of specific procedure codes in the procedure file.

The term “claim” is used throughout the document to represent either an outpatient pharmacy dispensing or medical encounter/record with any of the codes for the exposure(s) of interest.

The term “member” is used to represent an individual with relevant enrollment and other criteria (as specified by the MP parameters). A member can be further defined as a “user” if evidence of use of exposure(s) of interest is observed. Whenever a user is identified, the service date on the claim of the first exposure of interest observed during the relevant period of interest is labeled the “index date”.

The execution of MP5 allows information for multiple scenarios and exposures to be generated at the same time. Results from all scenarios are included in the MP output tables and can be differentiated using “query group names” defined by the requester. This document describes the process for only one scenario to be tested.

## 3 Program Summary

MP5 is used to query the MSDD to characterize the use of outpatient pharmacy medication(s) and/or medical procedure(s) and diagnoses during a period defined by a start and end date (i.e., the query period). By default, MP5 generates various metrics for both prevalence (i.e., ever exposed during query period) and incidence (i.e., newly exposed during query period) based cohorts of users for each [group of] medications and/or procedures/diagnoses of interest.

<sup>1</sup> See [http://www.mini-sentinel.org/data\\_activities/](http://www.mini-sentinel.org/data_activities/) for more information about the MSCDM.

Four stratified output tables characterizing the exposures of interest are created for the prevalent and incident cohort: overall for the entire period, by year, by year/month, and age group/sex. The metrics reported in all four tables include the number of users, the number of claims with exposures of interest, and the total members and eligible member days for use to estimate background rates ([Section 9](#)). In addition, tables for [incident based cohorts](#) include the number of incident claims. A fifth output table contains results for all possible combinations of strata (*e.g.*, by year/age group/sex). Finally, a sixth table containing denominator information for all strata is generated. Combining the information from the fifth and sixth tables allows the requester to generate population-based background rates ([Section 6](#)).

MP5 requires the specification of several parameters to define a scenario. These include program parameters that specify a request identifier, query period, age range(s), and enrollment criteria. The names of input files (built as SAS datasets) containing several other parameters used to identify the exposure(s) of interest and incident users must also be specified.

One input file is the [Query File](#) that defines exposures of interest. The second file is the [Incident Query File](#); it is optional and is used to refine how incident users are defined.

All parameters and input file specifications are described in [Section 4](#).

## 4 Program Parameter and Input File Specifications

### 4.1 Program Parameter Specifications

There are seven main [program parameters](#) that may be specified. These include a request identifier, start and end dates for the query period, age stratifications, an allowed enrollment gap used to create continuous enrollment periods, and two input files (the [Query File](#) and the [Incident Query File](#)). Four of these parameters are specified by the requester (query period start and end dates, age stratifications, and enrollment gap); three parameters are specified by the request programmer (request identifier, names of the [Query File](#) and [Incident Query File](#)).

Table 1 contains detailed specifications for each of these required parameters.

**Table 1: Main Program Parameter Specification**

Parameter	Field Name	Description
Request Identifier	REQUESTID	<p><b>Details:</b> a prefix added to all output to track the various executions of the program.</p> <p><b>Note:</b> Cannot exceed 8 characters.</p> <p><b>Defined by:</b> Request programmer</p> <p><b>Input type:</b> Required</p> <p><b>Format:</b> Alphanumeric</p> <p><b>Example:</b> REQUESTID=mpr5r1</p>

Parameter	Field Name	Description
Enrollment Gap	ENROLGAP	<p><b>Details:</b> 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 coverage in periods 1/1/2007-3/27/2007 and 4/1/2007-12/21/2007 (<i>i.e.</i>, 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.</p> <p><b>Note:</b> A gap of 45-days is recommended for most uses.</p> <p><b>Defined by:</b> Requester  <b>Input type:</b> Required  <b>Format:</b> Numeric  <b>Example:</b> ENROLGAP=14 (gaps less than or equal to 14 days will be “bridged” to form one “continuously enrolled” sequence)</p>
Query Start Date	QUERYFROM	<p><b>Details:</b> date for the start of the query identification period. If QUERYFROM=03/01/2008, only claims with a service date on or after this date will be considered.</p> <p><b>Defined by:</b> Requester  <b>Input type:</b> Required  <b>Format:</b> mm/dd/yyyy  <b>Example:</b> QUERYFROM=03/01/2008</p>
Query End Date	QUERYTO	<p><b>Details:</b> date for the end of the query identification period. If QUERYTO=03/31/2009, claims after this date will not be considered.</p> <p><b>Defined by:</b> Requester  <b>Input type:</b> Required  <b>Format:</b> mm/dd/yyyy  <b>Example:</b> QUERYTO=03/31/2009</p>
<a href="#">Query File</a>	QUERYFILE	<p><b>Details:</b> name of the SAS dataset defining the query exposure(s) of interest. It lists the codes that define each exposure of interest. The file name, including its extension, must be entered. For specific details on the content of this file, see <a href="#">Section 4.2</a>.</p> <p><b>Named by:</b> Request programmer  <b>Input type:</b> Required  <b>Format:</b> .sas7bdat or .cport file format  <b>Example:</b> QUERYFILE=query.sas7bdat or query.cport</p>

Parameter	Field Name	Description
<a href="#">Incident Query File</a>	INCQUERYFILE	<p><b>Details:</b> name of the SAS dataset refining the incident exposure definition(s). It contains additional codes and various parameters to further refine how incident use of each exposure must be defined. The file name, including its extension, must be entered. For specific details on the content of this file and for an example of when this file is used and how it is different from the QUERYFILE, see <a href="#">Section 4.2</a>.</p> <p><b>Named by:</b> Request programmer  <b>Input type:</b> Optional  <b>Format:</b> .sas7bdat or .cport file format  <b>Example:</b> INCQUERYFILE=incquery.sas7bdat or incquery.cport</p>
Age Groups	AGESTRAT	<p><b>Details:</b> 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 the result tables. For example, to have results stratified by 20 year increments for members 40-99 years of age, AGESTRAT=40-59 60-79 80-99 would be entered.</p> <p><b>Note 1:</b> age is determined at the index date.</p> <p><b>Note 2:</b> various units of time can be used. Valid values are:</p> <ul style="list-style-type: none"> <li>• <b>D:</b> days;</li> <li>• <b>W:</b> weeks;</li> <li>• <b>Q:</b> quarters;</li> <li>• <b>M:</b> months;</li> <li>• <b>Y:</b> years (default value).</li> </ul> <p><b>Note 3:</b> the 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.</p> <p>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.</p> <p><b>Defined by:</b> Requester  <b>Input type:</b> Optional (default value is <b>00-01 02-04 05-09 10-14 15-18 19-21 22-44 45-64 65-74 75+ in years</b>)  <b>Format:</b> AA-AA BB-BB ZZ-ZZ  <b>Example:</b> AGESTRAT=<b>40-59 60-79 80-99</b></p>

## 4.2 Input File Specifications

In addition to the seven main [program parameters](#), several required and optional parameters need to be specified in the [Query File](#) and [Incident Query File](#).

### 4.2.1 Query File

The [Query File](#) is required. It contains the comprehensive set of codes used to define the exposure(s) of interest. National Drug Codes (NDCs), ICD-9-CM procedure codes, Healthcare Common Procedure Coding System (HCPCS) codes, and/or ICD-9/ICD-10/ICD-11 diagnosis codes can be used to define exposures of interest. Exposure(s) of interest can be defined using any mix of allowed code types.

The structure of the [Query File](#) must reflect how codes should be queried to define a unique exposure. The GROUP field must be used to group all codes pertaining to a given exposure of interest. For example, a group for “Exposure1” could be defined by all NDCs for any oral forms of anti-diabetic medications, a group for “Exposure2” by a mix of NDC and HCPCS codes for certain insulin products, and another group for “Exposure3” by only those NDCs for a recently approved oral form of anti-diabetic medication.

There are seven required parameters that must be specified in the [Query File](#); six must be specified by the requester: query codes of interest, query code type, washout period and type, care setting, and principal diagnosis.

Table 2 describes specifications for the Query File.

**Table 2: Query File Specification**

Parameter	Field Name	Description
Name of Query Group	GROUP	<p><b>Details:</b> standardized name used to refer to a query GROUP for exposure(s) of interest to be queried.</p> <p><b>Note:</b> multiple exposures/query groups can be defined within the same query file. In this case all exposures are queried independently and results are reported separately and labeled using each query GROUP name specified.</p> <p><b>Named by:</b> Request programmer  <b>Input type:</b> Required  <b>Format:</b> Alphanumeric; SAS character \$30; no special characters (<i>e.g.</i>, commas, periods, hyphens, etc) allowed, and underscores must be used to mark spaces.  <b>Example:</b> Insulin1</p>

Parameter	Field Name	Description
Query Code Type	CODETYPE	<p><b>Details:</b> type of each code value included in the CODE field (below) of this file.</p> <p>Valid values are all those allowed by the MSCDM:</p> <ul style="list-style-type: none"> <li>• <b>RX11:</b> 11-digit NDC</li> <li>• <b>RX09:</b> 9-digit NDC</li> <li>• <b>PX09:</b> ICD-9-CM procedure</li> <li>• <b>PX10:</b> ICD-10-CM procedure</li> <li>• <b>PX11:</b> ICD-11-CM procedure</li> <li>• <b>PXC4:</b> CPT-4 procedure (i.e., HCPCS Level I)</li> <li>• <b>PXHC:</b> HCPCS procedure (i.e., HCPCS Level II)</li> <li>• <b>PXH3:</b> HCPCS Level III procedure</li> <li>• <b>PXC2:</b> CPT Category II procedure</li> <li>• <b>PXC3:</b> CPT Category III procedure</li> <li>• <b>DX09:</b> ICD-9-CM diagnoses codes;</li> <li>• <b>DX10:</b> ICD-10-CM diagnoses codes;</li> <li>• <b>DX11:</b> ICD-11-CM diagnoses codes.</li> </ul> <p><b>Note:</b> type of codes used for various query groups within a given query file or incident query file (see <a href="#">Section 4.2.2</a>) determines the continuous enrollment requirements. If only drug code types are used (i.e., RX09 and/or RX11) in the <a href="#">Query File</a> and <a href="#">Incident Query File</a>, then only pharmacy benefit coverage is required. If only procedure/diagnosis codes are used in the <a href="#">Query File</a> and <a href="#">Incident Query File</a>, then only enrollment in medical benefits is required. If both drug and procedure and/or diagnosis code types are used in the <a href="#">Query File</a> and <a href="#">Incident Query File</a>, then enrollment in both pharmacy and medical benefits is required. See <a href="#">Section 5.1</a> for details.</p> <p><b>Defined by:</b> Requester, with support from the Mini-Sentinel Operations Center (MSOC) as needed</p> <p><b>Input type:</b> Required</p> <p><b>Format:</b> Alphanumeric; SAS character \$4.</p> <p><b>Example:</b> See valid values</p>

Parameter	Field Name	Description
Query Codes of Interest	CODE	<p><b>Details:</b> NDC, procedure and/or diagnosis code values for the exposure(s) of interest.</p> <p><b>Note 1:</b> for NDCs, either 9 or 11 digit codes can be entered.</p> <p><b>Note 2:</b> any decimal points in the code value should be removed.</p> <p><b>Note 3:</b> CODETYPE must be consistent with the expected format of the CODE value (<i>e.g.</i>, MP5 will not find any valid matches in the data for CODETYPE=RX11 and a 9-digit NDC value).</p> <p><b>Note 4:</b> duplicate CODETYPE-CODE combinations are removed by the MP2 algorithm.</p> <p><b>Defined by:</b> Requester, with support from the MSOC as needed  <b>Input type:</b> Required  <b>Format:</b> Alphanumeric; SAS character \$11.  <b>Example:</b> 12345678911</p>
Query Incidence Type	WASHTYP	<p><b>Details:</b> defines whether an individual can contribute at most 1 or more than 1 incident exposure. As detailed in <a href="#">Section 5</a>, the incidence type can take the following values:</p> <ul style="list-style-type: none"> <li>• <b>MIN:</b> <a href="#">minimum incidence</a> -- determines whether the first claim for the exposure of interest occurs during the query period and meets the incident definition. Any observed claim before the query period would disqualify the member from contributing an incident claim.</li> <li>• <b>MULT:</b> <a href="#">multiple incidence</a> -- identifies and reports metrics for all valid incident exposures during the query period; a person can contribute more than 1 incident exposure.</li> </ul> <p><b>Note 1:</b> WASHTYP parameter is used in conjunction with the Query Washout Period (below) to define a valid incident claim.</p> <p><b>Note 2:</b> MIN setting should be used with caution as interpretation is complex. A claim that meets the MIN criteria must occur during the query period and cannot be preceded by a disqualifying claim at any time during enrollment; that is, the look-back period to define incidence is not limited to the washout period. See <a href="#">Section 5.3.1</a> for details.</p> <p><b>Defined by:</b> Requester, with support from the MSOC as needed  <b>Input type:</b> Required  <b>Format:</b> Alphanumeric; SAS character \$4; case sensitive (upper case only).  <b>Example:</b> See valid values</p>

Parameter	Field Name	Description
Query Washout Period	WASHPER	<p><b>Details:</b> length of washout period in days. The washout period is a period of time before an incident exposure index date during which a member cannot have any evidence of exposure(s) of interest or any other exposure(s) specified in the <a href="#">Incident Query File</a>. The washout period is also used to ensure that appropriate pharmacy and/or medical benefit coverage requirements are met.</p> <p>Exposures of interest with fewer than WASHPER days of enrollment before the first exposure date are excluded from the incident evaluation.</p> <p><b>Note 1:</b> length of query washout period days must be the same within a given query group.</p> <p><b>Note 2:</b> required and used in conjunction with any Query Incidence Types.</p> <p><b>Defined by:</b> Requester  <b>Input type:</b> Required (0 must be entered if no washout is required)  <b>Format:</b> Numeric  <b>Example:</b> 183</p>
Query Care Setting	CARESETTING	<p><b>Details:</b> contains the care settings considered for the individual query group diagnosis or procedure codes. If all care settings are wanted, leave the field blank. The following are valid entries; all entries must be quoted and separated by a space:</p> <ul style="list-style-type: none"> <li>• <b>IP:</b> inpatient hospital stays</li> <li>• <b>IS:</b> non-acute institutional stays</li> <li>• <b>ED:</b> emergency department visits</li> <li>• <b>AV:</b> ambulatory visits</li> <li>• <b>OA:</b> other ambulatory visits.</li> </ul> <p><b>Defined by:</b> Requester  <b>Input type:</b> Optional  <b>Format:</b> Alphanumeric; SAS character \$20.  <b>Example:</b> 'IP' 'ED'</p>
Principal Diagnoses Only	PRINCIPAL	<p><b>Details:</b> for a diagnosis CODETYPE (<i>e.g.</i>, DX09, DX10, DX11), if PRINCIPAL is set equal to YES, only principal diagnoses in the IP and ED settings will be chosen. However, if PRINCIPAL is set equal to NO, all diagnoses for the specified care settings will be chosen. For codes other than diagnosis codes, PRINCIPAL must be set to NO.</p> <p><b>Defined by:</b> Requester  <b>Input type:</b> Required  <b>Format:</b> Alphanumeric; SAS character \$3.  <b>Example:</b> NO</p>

## 4.2.2 Incident Query File

The [Incident Query File](#) is optional. By default, for a given query group MP5 uses the list of codes included in the [Query File](#) to determine the incident status of identified exposures. That is, incidence is defined as no exposure to the exposure of interest in the query washout period (WASHPER). The [Incident Query File](#) is used to define incidence based on a different set of codes than those used to define the exposure of interest. For example, the requester may be interested in examining the use of Drug A, but wants all incident users of Drug A to be free of Drug A, B and C in the 183 days before the Drug A index date. In this case, the full list of codes for Drugs A, B and C would be included in the [Incident Query File](#) to define incident use of exposures included in the “Drug A” query group, and all of these codes would be used to assess the incident status of Drug A as specified in the required [Query File](#).

There are three required parameters that must be specified in the [Incident Query File](#); two must be specified by the requester: the list of incident query codes and the code types.

The MP5 [Incident Query File](#) allows inclusion of the following code types: National Drug Codes (NDCs), ICD-9-CM procedure codes, Healthcare Common Procedure Coding System (HCPCS) codes, and/or ICD-9/ICD-10/ICD-11 diagnosis codes to define exposures of interest.

Table 3 describes specifications for the [Incident Query File](#).

**Table 3: Incident Query File Specification**

Parameter	Field Name	Description
Name of Query Group	GROUP	<p><b>Details:</b> standardized name used to refer to a query GROUP for exposure(s) of interest to be queried.</p> <p><b>Note:</b> must match GROUP values from the <a href="#">Query File</a>.</p> <p><b>Named by:</b> Request programmer</p> <p><b>Input type:</b> Required</p> <p><b>Format:</b> Alphanumeric; SAS character \$30; no special characters (<i>e.g.</i>, commas, periods, hyphens, etc) allowed, and underscores must be used to mark spaces.</p> <p><b>Example:</b> Insulin1</p>

Parameter	Field Name	Description
Incident Only Code Type	CODETYPE	<p><b>Details:</b> type of each code value included in the CODE field (below) of this file.</p> <p>Valid values are all those allowed by the MSCDM:</p> <ul style="list-style-type: none"> <li>• <b>RX11:</b> 11-digit NDC</li> <li>• <b>RX09:</b> 9-digit NDC</li> <li>• <b>PX09:</b> ICD-9-CM procedure</li> <li>• <b>PX10:</b> ICD-10-CM procedure</li> <li>• <b>PX11:</b> ICD-11-CM procedure</li> <li>• <b>PXC4:</b> CPT-4 procedure (i.e., HCPCS Level I)</li> <li>• <b>PXHC:</b> HCPCS procedure (i.e., HCPCS Level II)</li> <li>• <b>PXH3:</b> HCPCS Level III procedure</li> <li>• <b>PXC2:</b> CPT Category II procedure</li> <li>• <b>PXC3:</b> CPT Category III procedure</li> <li>• <b>DX09:</b> ICD-9-CM diagnoses codes;</li> <li>• <b>DX10:</b> ICD-10-CM diagnoses codes;</li> <li>• <b>DX11:</b> ICD-11-CM diagnoses codes.</li> </ul> <p><b>Note:</b> type of codes used for various query groups within a given query file or incident query file (see <a href="#">Section 4.2.2</a>) determines the continuous enrollment requirements. If only drug code types are used (i.e., RX09 and/or RX11) in the <a href="#">Query File</a> and <a href="#">Incident Query File</a>, then only pharmacy benefit coverage is required. If only procedure/diagnosis codes are used in the <a href="#">Query File</a> and <a href="#">Incident Query File</a>, then only enrollment in medical benefits is required. If both drug and procedure and/or diagnosis code types are used in the <a href="#">Query File</a> and <a href="#">Incident Query File</a>, then enrollment in both pharmacy and medical benefits is required. See <a href="#">Section 5.1</a> for details.</p> <p><b>Defined by:</b> Requester, with support from MSOC as needed  <b>Input type:</b> Required  <b>Format:</b> Alphanumeric; SAS character \$4  <b>Example:</b> See valid values</p>

Parameter	Field Name	Description
Incident Only Query Codes	CODE	<p><b>Details:</b> NDC, procedure and/or diagnosis code values to be used to refine the incident definition for the exposure(s) of interest. For example, if a requester is examining exposure to DRUG A but wants incidence of DRUG A to be defined with respect to DRUG A <i>and</i> DRUG B, the Incident Only Query Code fields would include DRUG A and DRUG B codes.</p> <p><b>Note 1:</b> for NDCs, either 9 or 11 digit codes can be entered.</p> <p><b>Note 2:</b> any decimal points in the code value should be removed.</p> <p><b>Note 3:</b> CODETYPE must be consistent with the expected format of the CODE value (<i>e.g.</i>, MP5 will not find any valid matches in the data for CODETYPE=RX11 and a 9-digit NDC value).</p> <p><b>Note 4:</b> duplicate CODETYPE-CODE combinations are removed by the MP5 algorithm.</p> <p><b>Defined by:</b> Requester, with support from MSOC as needed  <b>Input type:</b> Required  <b>Format:</b> Alphanumeric; SAS character \$11  <b>Example:</b> 12345678922</p>

## 5 Key Definitions

### 5.1 Enrollment Type Requirements

The type of drug and medical benefit coverage required is determined by the code types used to define Query Groups and Incident Only Query Codes. If the [Query File](#) and [Incident Query File](#) include only NDCs, MP5 will only consider claims that occur during enrollment periods where a member has drug coverage. A specified WASHPER of 183 days in this instance would only look for 183 days of continuous enrollment in drug coverage prior to claim date to determine enrollment eligibility. Medical coverage is irrelevant in this instance (*i.e.*, the member may or may not have simultaneous medical coverage).

If the [Query File](#) and [Incident Query File](#) include only procedure and/or diagnosis codes, MP5 will only consider claims that occur during enrollment periods where a member has medical coverage. A specified WASHPER of 183 days in this instance would only require 183 days of continuous enrollment in medical coverage to determine enrollment eligibility. Drug coverage is irrelevant in this instance (*i.e.*, the member may or may not have simultaneous drug coverage).

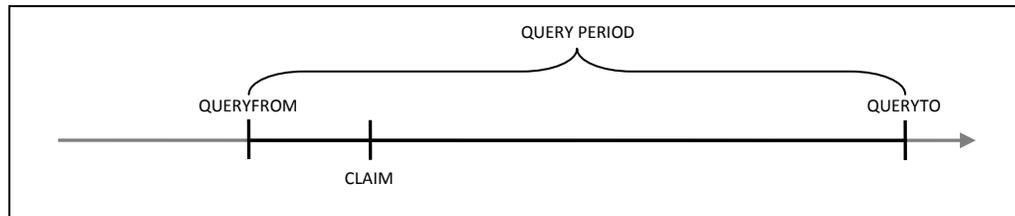
If the [Query File](#) and [Incident Query File](#) include drug and procedure and/or diagnosis codes, MP5 will only consider claims occurring during enrollment periods where a member has both drug *and* medical coverage. A specified WASHPER of 183 days in this instance would require 183 days of continuous enrollment in both drug and medical coverage to determine enrollment eligibility.

Please note that the enrollment type requirements described above differ from the requirements for Modular Programs 1 and 2. Those two modular programs always require drug coverage.

## 5.2 Prevalence-based Cohort

The [prevalence-based cohort](#) is composed of all members who have at least one claim that qualifies as an exposure of interest during the query period, delimited by the QUERYFROM to QUERYTO dates as illustrated in Figure 1. The claim date for each exposure of interest considered must fall during a continuous enrollment period of relevant type (See [Section 5.1](#)).

**Figure 1: Prevalence-based Cohort**



## 5.3 Incidence-based Cohort

The [incidence-based cohort](#) is a subset of the [prevalence-based cohort](#) that represents members who meet the incidence definition during the query period. When generating metrics for the [incidence-based cohort](#), the MP5 algorithm allows members to have one of the two following options: (1) [minimum incidence](#), where only one claim can be considered incident (specified using the WASHTYP=MIN option) or (2) [multiple incidence](#), where multiple claims can be considered incident (specified using the WASHTYP=MULT option).

### 5.3.1 Minimum Incidence (MIN)

Under the [minimum incidence](#) option, members can have only one incident claim during the query period and that claim must be the first one observed for that member in all enrollment periods of relevant types considered by the program. [Minimum incidence](#) is thus defined as evidence of one query group claim observed during the query period that satisfies the following three conditions. This option should be used with caution as interpretation can be complex.

1. Query Group claim date falls during a continuous enrollment period of relevant type.
2. Member is continuously enrolled WASHPER days before the Query Group claim date.
3. Member has no evidence of another claim with a Query Group code or [Incident Query File](#) code during any considered enrollment period before the Query Group claim date. That is, the MP5 algorithm queries the set of data available before the query claim date (regardless of the specified Query Period) to ensure that no claim with a Query Group code or [Incident Query File](#) code is observed during enrollment periods of relevant type under consideration.

### 5.3.2 Multiple Incidence (MULT)

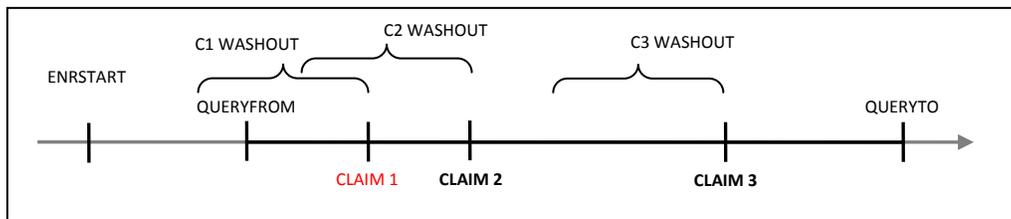
Under the [multiple incidence](#) option a member can have multiple incident claims as long as each claim meets the following incidence conditions. This definition is commonly used on epidemiologic assessments as it limits the period used to define incidence to a washout period equal for all members.

1. Query Group claim date falls during a continuous enrollment period of relevant type.
2. Member is continuously enrolled during the WASHPER days prior to the Query Group claim date.
3. Member has no evidence of a claim with a Query Group code or [Incident Query File](#) code during the WASHPER days prior to the Query Group claim date.

### 5.3.3 Examples

Figure 2A illustrates how the MP5 algorithm identifies a claim using the [minimum incidence](#) option.

**Figure 2A: Incident Claim Identification Using the Minimum Incidence Option\***

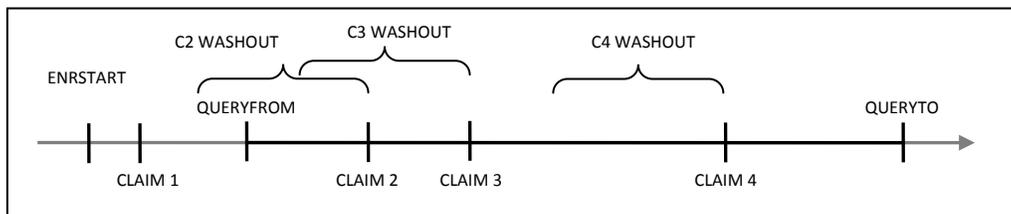


\*Incident claim in bold red

In the scenario depicted in Figure 2A, ENRSTART indicates the member's start of enrollment in drug and/or pharmacy benefits (depending on the CODETYPEs used in the [Query File](#) and [Incident Query File](#)). Claim 1 would be selected as the incident claim in this scenario, as the member satisfies the enrollment period condition (specified by the Washout Period parameter) and has no query group claim ever before Claim 1. Claims 2 and 3 also satisfy the enrollment period condition, but prior claims exclude them from being considered incident using the minimum incidence option.

Figure 2B depicts a scenario where a member does not have an incident claim under the [minimum incidence](#) option.

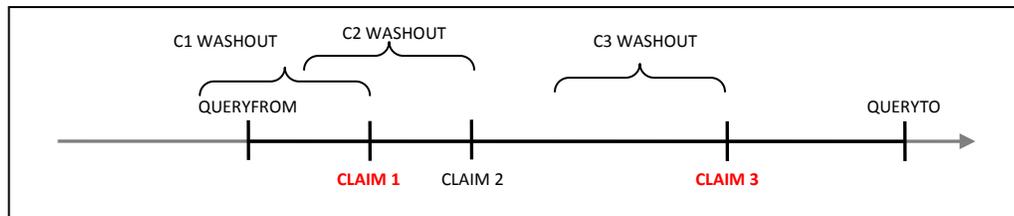
**Figure 2B: No Incident Claim Using the Minimum Incidence Option**



In this scenario Claim 1 is not considered incident since it did not occur within the query period. While claims 2, 3, and 4 satisfy the enrollment period condition, the presence of Claim 1 prior to the query period, but during the member’s history excludes these as incident claims under the [minimum incidence](#) option.

Figure 3 illustrates how the MP5 algorithm identifies a claim using the [multiple incidence](#) option.

**Figure 3: Incident Claim Identification Using the Multiple Incidence Option\***



\*Incident claims in bold red

In Figure 3, Claims 1 and 3 will be considered incident using the [multiple incidence](#) option since a washout period free of any claim (query or from the incident only list) is observed and both satisfy the enrollment condition.

## 6 Denominators and Background Rates

In addition to metrics on use of exposure(s) of interest, all output tables include metrics on denominators associated with each result stratum. Denominator metrics include (1) count of members and (2) eligible member days. These metrics are different for the [prevalence-based](#) vs. [incidence-based cohort](#). The members included in any denominator metrics are all the “eligible” members, *i.e.* members who can potentially be identified by the MP algorithm as a user (*i.e.*, either prevalent or incident) of the exposure(s) of interest defined in the query group.

### Prevalence-Based Denominators

The denominator member count for the [prevalence-based cohort](#) is composed of all members with at least one enrollment day (of relevant type) during the query period. See Table 2: Query File Specification for more details on the type of enrollment required based on the code types specified. Eligible member days for this cohort are simply the sum of all such enrollment days during the query period for the aforementioned members included in the denominator.

### Incidence-Based Denominators

The denominator member count for the [incidence-based cohort](#) is composed of all members enrolled for one period (or more) of continuous enrollment (of relevant type) of at least WASHPER+1 days (*i.e.*, WASHPER days for incidence requirement and one day to allow for exposure to be observed) overlapping but prior to the end of the query period. These members may be different depending on what Incidence Type is chosen. For the [minimum incidence](#) type the following metrics will be reported:

1. Members to be reported must be free of query claims or claims from exposure(s) found in the [Incident Query File](#) during all prior enrollment periods of relevant type in the member's history considered by the program.
2. Eligible member days reported are equal to the number of days of enrollment (of relevant type) falling during the query period that are (1) preceded by a period of at least WASHPER days long, free of any claims from exposure(s) found in the [Incident Query File](#), and b) prior to a query claim date (if exposed).

If, however, the [multiple incidence](#) option is chosen, the following will be reported:

1. Members to be reported must be free of query claims or claims from exposure(s) found in the [Incident Query File](#) during the washout period only.
2. Eligible member days are equal to the sum of all days of enrollment (of relevant type) falling during the query period that are preceded by a period of at least WASHPER days long, free of any query claims or claims from exposure(s) found in the [Incident Query File](#).

### **Background Rates**

Once denominator metrics are available, MP5 allows for the calculation of background rates for each Query Group specified. For example, the requester can calculate the proportion and rate of new users of exposures in Query Group Z in adults aged between 40 and 50 years old. The *proportion of new users* would be defined as the number of new users (*i.e.*, the numerator) per X enrolled members (*i.e.*, the denominator), whereas the *rate of new users* would be defined as the number of new users per Y person-time (*e.g.*, person-years, member-days).

## **7 Program Steps**

The general program steps are:

1. Extract claims of interest from the MSDD outpatient pharmacy/procedure/diagnosis tables to create exposure dates for each member.
2. For each exposure of interest/Query Group and for each member, assess prevalent and incident status, and extract all claims of interest during query period to calculate utilization by member and Query Group.
3. For each Query Group, calculate denominator member counts and days enrolled.
4. Calculate age at index date and extract sex for each cohort member using the MSDD demographic file.
5. Create output tables for both prevalent and incident users along with denominators and member days.

## **8 Program Execution**

When implementing modular programs within the MSDD, the Mini-Sentinel Operations Center (MSOC) uses a uniform folder structure across Data Partners to facilitate communications between MSOC and Data Partners and to streamline file management.

Each request distributed by MSOC is assigned a unique Request ID. Upon receipt of the request, Data Partners create a folder named after the Request ID and several subfolders to organize program inputs

and outputs. One of the folders contains output to be sent to MSOC and another contains intermediate files that remain with the Data Partner, but could be used to facilitate follow-up queries if necessary. Appropriate retention policies apply.

Table 4 defines the local environment variables that must be initialized by the user to execute the program (*i.e.*, defined by the Data Partner prior to execution of the program). Please note that these values cannot be left blank. Each Data Partner is required to enter user inputs at the beginning of the SAS Program sent with each request. These inputs are unique to each Data Partner.

**Table 4: Environment Variable Definitions**

Label	Field Name	Description
Data Partner ID	DPID	Two character Partner ID
Site ID of Data Partner	SITEID	Two character Site ID
Enrollment Table Name	ENRTABLE	Name of the Data Partner’s MSCDM Enrollment table
Demographics Table Name	DEMTABLE	Name of the Data Partner’s MSCDM Demographics table
Dispensing Table Name	DISTABLE	Name of the Data Partner’s MSCDM Dispensing table
Procedure Table Name	PROCTABLE	Name of the Data Partner’s MSCDM Procedure table
Diagnosis Table Name	DIATABLE	Name of the Data Partner’s MSCDM Diagnosis table
Input file folder	INFOLDER	Path where the input files will be saved
Output file folder	MSOC	Path where the shared output tables will be saved
Local data folder	DPLOCAL	Path where the local SAS datasets will be saved
Libname of the MSCDM	INDATA	Path where the local MSDD data is saved

## 9 Output Tables

Ten output tables are created by MP5 and each are output in both .sas7bdat and .csv formats. Four output tables each are created for the [incident-based cohort](#) and the [prevalence based cohort](#) and a numerator and denominator table are also output containing the full query summary statistics. The numerator and denominator tables contain all data necessary to calculate incident and prevalent rates stratified by month/year, year, and age/gender. Tables can be identified by the following suffixes:

- \_ITABLEX
- \_PTABLEX
- \_DENTABLE0
- \_NUMTAB0

The “I” and “P” correspond to incident and prevalent tables, respectively, and the “X” corresponds to the table number and takes the values 1 to 4. There is only one denominator table (DENTABLE0) and one numerator table (NUMTAB0). Below are examples of the incident and prevalent output tables, produced using the multiple incidence option.

**ITABLE1 - Counts of Patients and Claims over the Query Period**

Group	Users	Incident Claims	Claims	Eligible Members	Eligible Days
HOIGROUP1	20,509	24,144	104,952	435,280	232,940,477
HOIGROUP2	8,858	9,769	12,915	438,999	238,772,434

First row interpretation of ITABLE1: Within the QUERYFROM to QUERYTO period, 20,509 members (users) had at least one incident claim for any drug, procedure or diagnosis within the HOIGROUP1 category. For these members, 24,144 incident HOI claims and 104,952 individual claims were observed. Finally, 435,280 members could have been included as incident users (eligible members) for a total of 232,940,477 eligible days.

If the above table had been a prevalent table, it would appear as follows:

**PTABLE1 - Counts of Patients and Claims over the Query Period**

Group	Users	Incident Claims	Claims	Eligible Members	Eligible Days
HOIGROUP1	22,495	---	123,994	433,185	245,620,434
HOIGROUP2	11,143	---	14,984	451,401	265,716,473

First row interpretation of PTABLE1: Within the QUERYFROM to QUERYTO period, 22,495 members (users) had at least one prevalent claim for any drug, procedure or diagnosis in HOIGROUP1. For these members there were 123,994 individual claims. Finally, 433,185 members could potentially have been included as prevalent users (eligible members) for a total of 245,620,434 eligible days.

**ITABLE2 - Counts of Patients and Claims over the Query Period Stratified by Year**

Group	Year	Users	Incident Claims	Claims	Eligible Members	Eligible Days
HOIGROUP1	2007	5,604	5,757	17,038	232,751	59,839,910
HOIGROUP1	2008	8,067	8,275	24,276	309,745	81,966,398
HOIGROUP2	2007	2,170	2,200	2,635	235,402	61,317,667
HOIGROUP2	2008	3,386	3,439	4,166	313,240	83,911,011

First row interpretation of ITABLE2: In 2007, 5,604 members (users) had at least one incident claim for any drug, procedure or diagnosis in HOIGROUP1. In the same year for these members, 5,757 incident HOI claims and 17,038 individual claims were observed. Finally, 232,751 members could potentially have been included as incident users in 2007 (eligible members) for a total of 59,839,910 eligible days. The explanation above can also be applied for prevalent table interpretations. Replace “incident” with “prevalent” in the above explanation and set the value of “Incident Claims” to missing as this is irrelevant for prevalent rates. Note that counts should be greater for prevalent compared to incident tables.

**ITABLE3: Counts of Patients and Claims over the Query Period Stratified by Year/Month**

Group	Year	Month	Users	Incident Claims	Claims	Eligible Members	Eligible Days
HOIGROUP1	2007	6	395	395	521	128,884	3,852,881
HOIGROUP1	2007	7	555	555	722	197,126	6,087,607
HOIGROUP2	2007	6	133	133	142	131,963	3,952,396
HOIGROUP2	2007	7	211	211	228	201,432	6,231,763

First row interpretation of ITABLE3: In June 2007, 395 members (users) had at least one incident claim for any drug, procedure or diagnosis in HOIGROUP1. In the same month for these members, 395 incident HOI claims and 521 individual claims were observed. Finally, 128,884 members could potentially have been included as incident users (eligible members) for a total of 3,852,881 eligible days. The explanation above can also be applied for prevalent table interpretations. Replace “incident” with “prevalent” in the above explanation and set the value of “Incident Claims” to missing as this is irrelevant for prevalent rates. Note that counts should be higher for prevalent tables relative to incident tables.

**ITABLE4: Counts of Patients and Dispensings over the Query Period Stratified by Age Group and Sex**

Group	Age Groups	Sex	Users	Incident Claims	Claims	Eligible Members	Eligible Days
HOIGROUP1	20 to 44	F	6,292	7,203	32,475	113,297	49,589,636
HOIGROUP1	20 to 44	M	2,761	3,127	13,759	103,399	45,453,982
HOIGROUP1	45 to 64	F	6,476	7,725	32,684	97,217	54,305,185
HOIGROUP1	45 to 64	M	2,886	3,432	14,904	88,225	49,421,925
HOIGROUP1	65+	F	1,506	1,808	6,332	28,821	18,647,034
HOIGROUP1	65+	M	718	849	2,886	24,278	15,522,715
HOIGROUP2	20 to 44	F	1,868	2,034	2,656	114,645	51,382,675
HOIGROUP2	20 to 44	M	1,821	1,943	2,417	103,875	45,969,697
HOIGROUP2	45 to 64	F	2,266	2,566	3,461	98,573	56,460,035
HOIGROUP2	45 to 64	M	2,313	2,546	3,407	88,788	50,108,091
HOIGROUP2	65+	F	291	316	454	29,040	19,133,294
HOIGROUP2	65+	M	334	364	472	24,377	15,718,642

First row interpretation of ITABLE4: 6,292 female members (users) in the 20 to 44 years old age group had at least one incident claim for any drug, procedure or diagnosis in HOIGROUP1 within the QUERYFROM to QUERYTO period. For these members, 7,203 incident HOI claims and 32,475 individual claims were observed. Finally, 113,297 female members in the 20 to 44 year old age group could potentially have been included as incident users (eligible members) for a total of 49,589,636 eligible days. The explanation above can also be applied for prevalent table interpretations. Replace “incident” with “prevalent” in the above explanation and set the value of “Incident Claims” to missing as this is irrelevant for prevalent rates. Note that counts should be higher for prevalent tables relative to incident tables.

## 10 Example

Table 5 and Table 6 below show an example of a [Query File](#) and [Incident Query File](#) (e.g., queryfile.sas7bdat and incqueryfile.sas7bdat):

**Table 5: Example of Query File**

Code	Group	CodeType	WashPer	WashTyp	CareSetting	Principal
1111111100	HOIGROUP1	RX11	180	Mult	'IP'	NO
111111112	HOIGROUP1	RX09	180	Mult	'IP'	NO
99999	HOIGROUP2	PXC4	30	Min		NO
99998	HOIGROUP2	PXC4	30	Min		NO

**Table 6: Example of Incident Query File**

Code	Group	CodeType
1111111101	HOIGROUP1	RX11
77777	HOIGROUP2	PXC4

In this example, the requester specifies that, in addition to the codes listed in the [Query File](#), the codes from the [Incident Query File](#) for the HOIGROUP1 and HOIGROUP2 Query Groups need to be considered.

Furthermore, the requester specifies additional parameters to be used:

- Any enrollment gap of less than 45 days is considered administrative and must be ignored;
- The query period spans the years 2007 to 2010; and
- The results should be segmented according to 5 year age groups from 10 to 49 years old.

For this request, the program could be executed using the following SAS Macro call:

**%ModularProgram5(**

```

REQUESTID=mpr_ex5,
ENROLGAP=45,
QUERYFROM=01/01/2007,
QUERYTO=12/31/2010,
QUERYFILE=queryfile.sas7bdat,
INCQUERYFILE=incqueryfile.sas7bdat,
AGESTRAT=10-14 15-19 20-24 25-29 30-34 35-39 40-44 45-49);

```