Standard Operating Procedures

Sentinel Data Retention by Data Partners

Prepared by the Sentinel Operations Center

Version 2.2

December 6, 2021

Standard Operating Procedure

Sentinel Data Retention by Data Partners

Table of Contents

History of Modifications

Version	Date	Modification	Author
1.0	07/30/2018	First draft reviewed and approved	SOC
2.0	05/03/2019	Second draft incorporating additional policies for the Data Partners producing the Mother-Infant Linkage (MIL) Sentinel Common Data Model Table	SOC
2.1	10/10/2019	Replaced request type CCR with CCA for Phase A Data Quality Review and Characterization Data	SOC
2.2	12/06/2021	Added detail to Phases A and B Data Quality Review and Characterization retention policies; renamed TreeScan to Signal Identification	SOC

1   Terminology and Abbreviations

- DMQA: Data Management and Quality Assurance
- DP: Data Partner
- ETL: Extract, Transform, Load. Sentinel databases are given ETL numbers (e.g., ETL 1, ETL 2) for tracking purposes. The same ETL number is used to reference both the Sentinel Common Data Model tables included in Phase A (Sentinel Common Data Model tables except mother-infant linkage table) and in Phase B (all Sentinel Common Data Model tables) (e.g., ETL 1 Phase A, ETL 1 Phase B).
- MI: Mother-Infant
- MI ID: Mother-Infant Identification program package, which detects mothers and infants in Data Partners' Sentinel Common Data Models
- MIL: Mother-Infant Linkage
- MIL integration: Development of new quality assurance processes and programs, including common components, to routinely incorporate the MIL Table into the Sentinel Common Data Model
- Phase A: Review and approval of all Sentinel Common Data Model tables except the mother-infant linkage table
- Phase B: Review and approval of the mother-infant linkage table
- SCDM: Sentinel Common Data Model
- SOC: Sentinel Operations Center
- SOP: Standard Operating Procedure

2 Introduction

The Sentinel System is sponsored by the [U.S. Food and Drug Administration (FDA)](#) to proactively monitor the safety and effectiveness of FDA-regulated medical products and complements other existing FDA safety surveillance capabilities. The [Sentinel System](#) is one piece of FDA's [Sentinel Initiative](#), a long-term, multi-faceted effort to develop a national electronic system for active safety surveillance of FDA-regulated medical products. Sentinel Collaborators include Data and Academic Partners that provide access to healthcare data and ongoing scientific, technical, methodological, and organizational expertise. The Sentinel Operations Center (SOC) is funded by the FDA through the Department of Health and Human Services (HHS) Contract number HHSF223201400030I.

Given Sentinel's substantial number of participating organizations and scale of activity, Standard Operating Procedures (SOPs) are needed to support the project. SOPs are detailed instructions written and followed to achieve uniform performance of specific functions. They provide a basis for evaluation and improvement of processes.

3 Purpose and Scope

The purpose of this document is to define the processes for determining what Sentinel-related datasets and files must be retained by Data Partners, the duration of the required retention periods, and processes by which the Sentinel Operations Center (SOC) will advise Data Partners concerning datasets and files that can be deleted by Data Partners at their respective sites. This SOP adheres to the current [Sentinel Statement of Principles and Policies](#).

4 Roles and Responsibilities

Table 1 describes the Roles and Responsibilities associated with the Sentinel Data Retention by Data Partners process.

Table 1. Roles and Responsibilities

| Location	Role Name	Responsibilities
Sentinel Operations Center (SOC)	**Leadership**	• Approve policy and requirements for data retention by Data Partners
	SOP Owner	• Develop initial and subsequent versions of the SOP • Oversee SOP review process • Oversee policy implementation • Manage communications concerning this SOP
	Data Partner Relationship Manager	• Monitor the needs of SOC and Data Partners • Recommend amendments to the SOP, as needed
	List Generator	• Create list of all datasets/files that can be deleted by Data Partners based on the approved length of time since execution
	List Reviewer	• Review list of datasets/files that can be deleted by Data Partners, confirming accuracy, prior to distribution to Data Partners
	Communications Manager	• Serve as primary contact for Data Partners; refer questions to relevant teams as needed • Distribute communications to Data Partners concerning all datasets/files that can be deleted • Develop, maintain, and distribute Process Enablers
Data Partners	**Data Managers**	• Retrieve lists of datasets/files that can be deleted on a quarterly basis • Review lists of datasets/files that can be deleted • Take action consistent with SOC

5 Process

5.1 Process Enablers

Process Enablers are tools such as templates, forms, requirements, guidelines, and other documentation, separate from the SOP, which enable efficient management of the process. The following Enablers support this SOP:

- <u>Atlassian JIRA Software</u> – Tool used by SOC to track the distribution of data requests that enables SOC and Data Partners to identify query-related datasets and files that are due for deletion

- <u>Email</u> – Tool used as a mode of communication to Data Partners concerning datasets and files that can be deleted

- <u>Dataset/File Tracker</u> – List of information exported from the Atlassian JIRA Software distributed to Data Partners by SOC

5.2 Process Flow

Figure 1 depicts the process flow for determining what datasets and files can be deleted by Data Partners at their respective sites and communicating that information to Data Partners. For query-related datasets, communications about files eligible for deletion will be distributed on a quarterly basis (at the end of April, July, October, and January). For Extract, Transform, Load (ETL) infrastructure-related datasets, communications about files eligible for deletion will be distributed as ETLs are completed. Roles at the SOC are color-coded in green; the role at the Data Partners is color-coded in blue. The process is performed according to the numerical order of the process steps.

Figure 1. Data Retention by Data Partners Process Flow for Query-Related and ETL Infrastructure-Related Datasets and Files
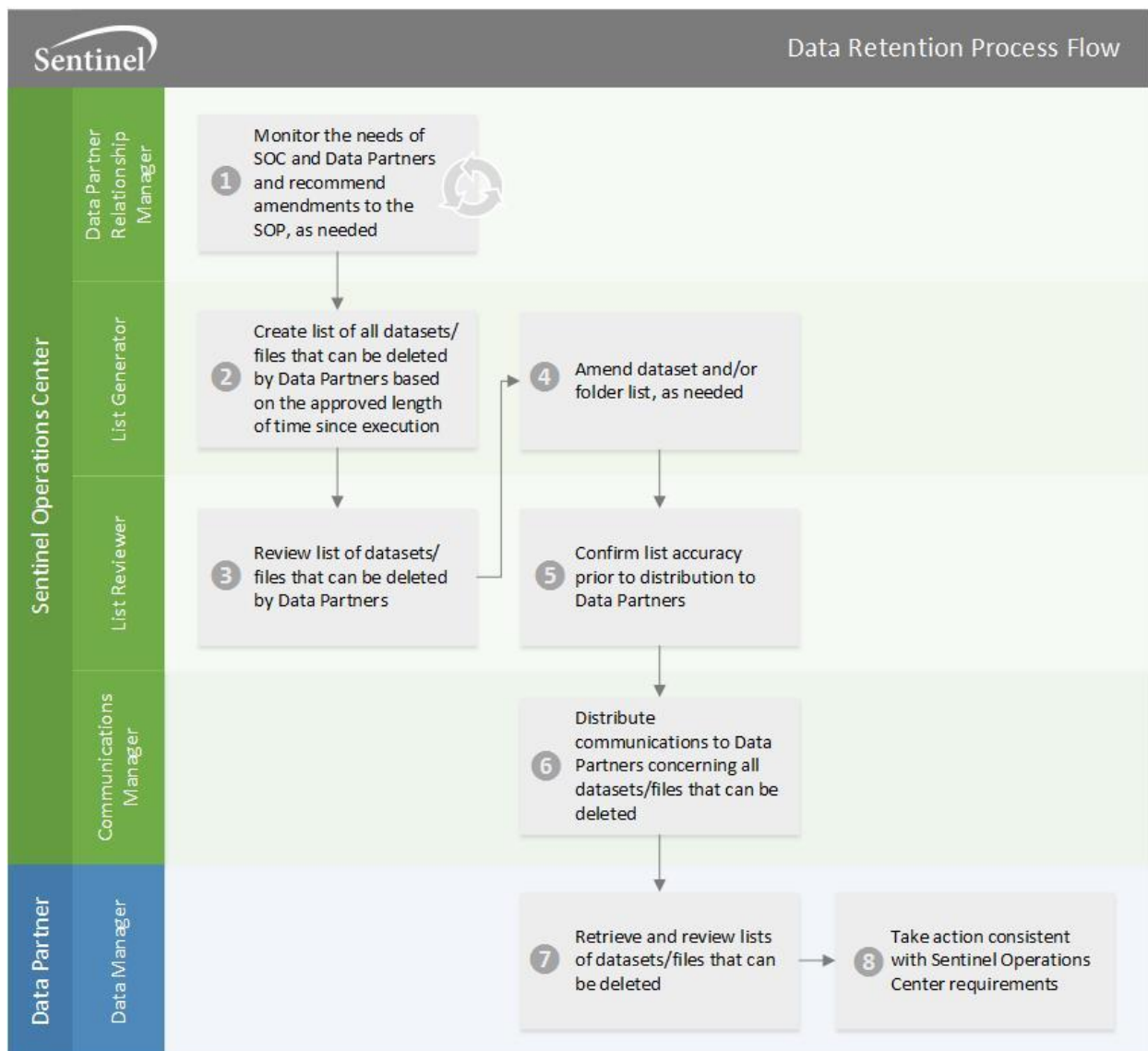
6 Data Retention Requirements by Specific Data Type

Sentinel data retention requirements depend on several factors, including FDA's regulatory needs and the types of datasets/files involved. Required data retention periods by type of data are shown in Table 2. **Note that request types represent unique tokens for use in defining a Sentinel analysis; tokens are not always an abbreviated term.**

Table 2: Data Retention Requirements by Specific Data Type

Data Type	Request Type	Description	Required Retention Policy
Query Related Datasets and Files			
1. Modular Program Queries	MPL1R	Modular Program Level 1 Routine Queries	2 years from request distribution
	MPL1P	Modular Program Level 1 Plus Queries	2 years from request distribution
	MPL2R	Modular Program Level 2 Routine Queries	3 years from request distribution
	MPL2P	Modular Program Level 2 Plus Queries	3 years from request distribution
	MPL3R	Modular Program Level 3 Routine Queries	3 years from request distribution
	MPL3P	Modular Program Level 3 Plus Queries	3 years from request distribution
2. Patient Episode Profile Retrieval Request Results	PEPR	Patient Episode Profile Retrieval Requests	3 years from request distribution
3. Signal Identification Request Results	SIR	Signal Identification Routine Requests	3 years from request distribution
	SIP	Signal Identification Plus Requests	3 years from request distribution
4. Ad Hoc Request Results	AHR	Ad Hoc Requests	3 years from request distribution
5. Survey Responses	SVR	Sentinel Distributed Database (SDD) Surveys	Retention is not required
6. Chart Reviews	N/A	External source data provided by clinicians/hospitals/other facilities in paper and/or electronic format	7 years from project completion, defined as acceptance of the final report by FDA
7. Sentinel Database Reporting Statistics	DGR	SDD Characterization Reports Results from "as needed" queries to characterize the SDD	3 months from request distribution
8. Continuous Data Improvement Statistics	CDR	Data Exploration to Support Infrastructure Development Results from internal Sentinel Operations Center (SOC) "deep dives" to provide information about	3 months from request distribution

Data Type	Request Type	Description	Required Retention Policy
		potential changes to the Sentinel Common Data Model (SCDM)	
9. Infrastructure Testing Results	QAD	Quality Assurance (QA) Program Package Development For internal SOC testing purposes	3 months from request distribution, or as instructed by the Sentinel Infrastructure Team
	CCD	Common Components Development For internal SOC testing purposes	3 months from request distribution, or as instructed by the Sentinel Infrastructure Team
	MPD	Modular Program Development For internal SOC testing purposes	3 months from request distribution, or as instructed by the Sentinel Infrastructure Team
ETL Infrastructure Related Datasets and Files			
10. Sentinel Common Data Model (SCDM) Tables	N/A	Sentinel Database in SCDM-format	For DPs not populating SCDM Mother-Infant Linkage (MIL) table 1. Current production ETL (ETL n Phase A[1], accessible) 2. Previous production ETL (ETL n-1 Phase A, available for retrieval within 1 week) For DPs populating SCDM MIL table 1. Current production ETL (ETL n Phase A**Error! Bookmark not defined.** or Phase B[2], accessible) 2. Previous production ETL (ETL n-1 Phase B, available for retrieval within 1 week) 3. Previous production ETL (ETL n-2 Phase B, if current prod is ETL Phase A,

---

[1] Review and approval of all SCDM tables except the mother-infant linkage table
[2] Review and approval of the MIL table

Data Type	Request Type	Description	Required Retention Policy
			available for retrieval within one week)[3]
11. Phase A Data Quality Review and Characterization Data	QAR MIR	Data Quality Review and Characterization Program Package • Results from QA program package (QAR) Results from the mother-infant identification package (for MIL DPs only) (QAR or MIR)	Retain results of QAR and MIR packages associated with the retained SCDM versions described in Data Type 10, Sentinel Common Data Model (SCDM) Tables
	CCA	QA Common Components Phase A Program Package Produces files specific to DP site, ETL, and Phase.	Retain results of CCR packages associated with the retained SCDM versions described in Data Type 10, Sentinel Common Data Model (SCDM) Tables
12. Phase B Data Quality Review and Characterization Data	QMR	MIL Data Quality Review and Characterization Program Package Results from MIL QA program package	Retain results of QMR packages associated with the retained SCDM versions described in Data Type 10, Sentinel Common Data Model (SCDM) Tables
	CCB	QA Common Components Phase B Program Package Produces files specific to DP site, ETL, and Phase.	Retain results of CCB packages associated with the retained SCDM versions described in Data Type 10, Sentinel Common Data Model (SCDM) Tables

---

[3] DPs creating the MIL table are required to have at least one ETL in production that contains a DMQA-approved MIL table. This may lead to DPs being required to support up to four ETLs at one time (up to three production databases, and ability to create a new ETL in a staging environment).