This document provides CDC guidance and best practices for SSuN recipients to follow for standardizing variables, response codes, data structures, formats, data quality procedures, and submitting SAS datasets via CDC’s Secure Access Management Service (SAMS).

Key definitions:

**SAMS**: Secure Access Management Service portal. The SAMS portal is a secure web-based application that provides secure transport of data between jurisdictions and CDC which encrypts data automatically. Only registered users have access to the SAMS portal.

**Data Quality Assurance**: Interconnected processes and procedures to assure validity of data, fidelity of data collection to pre-defined protocols and to assure that data outputs are properly formatted, contain the required records and can be seamlessly integrated into master datasets at the national level.

**Edit Checks**: SAS code that is used to ‘check’ the contents, formats and appropriateness of data reported. Edit check code will be supplied for recipients to run prior to sending data as well as run to check datasets as they arrive at CDC.

Considerations for implementation

**Data Submission Requirements for Strategy A**
Each strategy in SSuN consists of several SAS datasets transmitted to CDC via SAMS bi monthly. There are five (5) separate datasets required for Strategy A:

- **a.** STD Clinic visit records (containing a unique record for each visit to collaborating STD clinics)
- **b.** Diagnosis records (contains all diagnoses associated with each visit, linked through event id)
- **c.** Laboratory records (contains all laboratory records associated with each visit and from HIV match, linked through event and patient IDs)
- **d.** Treatment records (contains all STD-related treatments associated with each unique visit, linked through event ID)
- **e.** Facility reference file updated annually containing descriptive information about each collaborating STD clinic

**Data Submission Requirements for Strategy B**
There are three (3) separate datasets required for Strategy B:

- **a.** Case records with imbedded provider and patient investigation data
- **b.** Laboratory records (contains all laboratory observations and for cases and HIV match, linked through event and patient IDs)
- **c.** Provider reference records (contains information about diagnosing facility or provider, linked through case and provider IDs)

**Data Submission Requirements for Strategy C Activities**
Requirements may vary by specific activity and in some case may involve aggregate reporting only; data management protocols will be developed post-award for Strategy C Surveillance Focus Activities not previously covered by SSuN protocols.

**Data Quality Assurance Requirements**
CDC will provide detailed information on data structures with variable names, lengths, and types defined for all requested SAS datasets. SSuN recipients are required to ensure the correct data structures, validity of data elements and check for missing values in selected data elements before submitting data to CDC. Specifically, recipients must:

1. Verify all variables are submitted with the proper variable names and data types (e.g., character, numeric, date, etc.)
2. Verify all key identifiers in records (e.g., event, provider, laboratory and patient IDs) are correct and that these match across related datasets; ‘orphan’ records in related datasets must be resolved prior to sending to CDC.
3. Duplicate records must be merged or removed to assure uniqueness of visit and case ‘parent’ records.
4. Investigate and fix records that are missing any required data.

To aid jurisdictions with the data quality assurance requirements, CDC will provide SAS programs to check for correct file structure, data elements, and identify records with missing values. Jurisdictions must run these programs and correct any errors prior to submitting data to CDC. Additionally, each jurisdiction is responsible for conducting their own local data quality checks prior to the CDC data validation step for transmission.

Data Transmission Schedule
CDC will provide a data transmission schedule to SSuN recipients at the beginning of the project detailing which datasets (with a date range for included records) are required and due to CDC by specific dates. Data transmission will occur every month, alternating between Strategy A and Strategy B and including Strategy C activities as needed. Annual ‘cleaned’ and final datasets for each strategy are due in March for data from the prior calendar year; these data will serve as CDC’s archive for analytic and reporting purposes. Repeated failure to meet data transmission schedules may have fiscal impacts on recipients and could result in remedial action.

Evaluation and Resubmission Activities
CDC will perform edit checks on the SAS datasets before creating master files for analyses and reporting. If significant errors are identified in this process, or through subsequent analytic projects, recipients will be asked to fix and re-submit data for a specific time period. The success of SSuN critically depends on the collaborative efforts of all recipients and CDC project staff to assure the highest quality of data – used both nationally and locally to direct STD prevention and control activities.

SSuN recipients should consider data quality assurance as an integral surveillance activity requiring formal planning and adequate staff resources. When developing your project plans, budgets and staffing needs, be sure to include sufficient resources for data quality assurance activities. Creating a formal ‘Data Quality Plan’ early in your project cycle, and revisiting your plan regularly is a critical best practice for project management!

Other resources
- SSuN Cycle 4 Protocols: https://www.cdc.gov/std/funding/ssun/default.htm

CDC’s Division of STD Prevention, Surveillance and Special Studies Team created this series of documents for CDC staff working on PS19-1907 STD Surveillance Network (SSuN) and for applicants and recipients to help clarify strategies outlined in the NOFO and to support project implementation.