SAS® Clinical Standards Toolkit 1.2
Quick Start for SDTM Validation for SAS 9.2 and SAS 9.1.3

SAS® Clinical Standards Toolkit 1.2 Quick Start for SDTM Validation for SAS 9.2 and SAS 9.1.3

Contents

Getting Started

This version of the SAS® Clinical Standards Toolkit 1.2 – Quick Start Tutorial for SDTM Validation is for use with SAS 9.2 and SAS 9.1.3.

Overview

SAS® Clinical Standards Toolkit 1.2 (CST) is a powerful and flexible product. This tutorial provides a quick introduction to CST. Using this tutorial, you will walk through the main tasks required to use CST to validate CDISC SDTM 3.1.1 data. When you have completed this tutorial, you will understand how to configure CST to run CDISC SDTM 3.1.1 validations against your own data. While this tutorial specifically discusses CDISC SDTM 3.1.1 validation, many of the concepts apply to using CST to perform other tasks.

This tutorial is intended to show you the basic steps for using CST. It is not intended to provide comprehensive training in CST, the capabilities of CST, or the best practices. The goal is to present what the SAS code is doing. For example, the tutorial avoids using unnecessary SAS macros or SAS macro variables and all paths are hard-coded. The comments within the SAS programs provide additional information about the tasks being performed.

One of the biggest challenges you will face when using CST is the preparation of your own data and metadata. By definition, this data must be in CDISC SDTM 3.1.1 SAS data sets. The steps required to get your data into CDISC SDTM 3.1.1 SAS data sets is beyond the scope of this tutorial.

This tutorial uses a small number of the many options available within CST. To understand all the options provided by CST, see the *SAS Clinical Standards Toolkit 1.2: User's Guide*.

Tutorial Goal and Exercises

This tutorial walks you through the main steps that you need to use CST to validate your own data. Your own data will be represented as the fictitious study, MyStudy. The challenge is to test how well MyStudy complies with the CDISC SDTM 3.1.1 standard.

The tutorial consists of the following four steps:

1. Install this tutorial.
2. Exercise 1: Prepare the study data and metadata.
3. Exercise 2: Walk through the steps needed to setup CST to run a CDISC SDTM 3.1.1 Validation.
4. Exercise 3: Consolidate and simplify the steps from Exercise 2 into a single program.

Install this Tutorial

In this section, you will install the directories and files provided in this tutorial. To set up the tutorial, extract the following applicable file into the c:\ directory:

| For SAS 9.2 | cstQuickStartSDTMValidation.zip |
| For SAS 9.1.3 | cstQuickStartSDTMValidation913.zip |

The extraction creates the following directory which contains all of the subdirectories and files required by this tutorial:

| For SAS 9.2 | c:\cstQuickStart |
| For SAS 9.1.3 | c:\cstQuickStart913 |

The subdirectories and files are described in the following table.

Table 1. Quick Start Tutorial Folder Structure

🗀MyStudy	Contains the data and metadata that are specific to a study.
🗀MyData	o Study data in CDISC SDTM 3.1.1 SAS data set format.
🗀MyFormats	o Format catalogs including coding dictionaries.
🗀MyMetadata	o SAS data sets that describe the tables and columns that are used to store the data in MyData, for example, the study data's metadata.
🗀Exercise1_PrepareStudyData	Three SAS programs that create the data and metadata that comprise MyStudy. Technically, these programs are not part of CST. The programs are included in the tutorial because you might need to perform similar tasks.
🗀Exercise2_StepByStep	Three SAS programs that walk through the set up and execution of the CST CDISC SDTM 3.1.1 validation.
🗀Exercise3_AllInOne	One SAS program that combines the steps outlined in Exercise 2.

Note: When created, some directories are empty. The tutorial exercises will add content to these directories.

The tutorial zip file is located on the SAS download page. If you need help locating this file, search the SAS support Web site (http://support.sas.com) or contact your SAS Administrator for assistance. A copy of this document is provided within the tutorial zip file.

Default Directories and Installation

This tutorial and the supporting SAS programs and SAS data sets all assume that you have extracted them into their default directories.

Table 2. Installed Software and the Associated Default Directory

Installed Software	SAS Software Version	Default Directory
Tutorial	For SAS 9.2	c:\cstQuickStart
	For SAS 9.1.3	c:\cstQuickStart913

SAS	For SAS 9.2	c:\Program Files\SAS\SASFoundation\9.2
	For SAS 9.1.3	c:\Program Files\SAS\SAS 9.1
	Note: This directory also is known as the SAS Root folder.	
SAS Clinical Standards Toolkit	For SAS 9.2	c:\Program Files\SAS
	For SAS 9.1.3	c:\Program Files\SAS
	Note: This directory also is known as the SAS Home folder.	
SAS Clinical Standards Toolkit's Global Library	For SAS 9.2	c:\cstGlobalLibrary
	For SAS 9.1.3	c:\cstGlobalLibrary
	Note: Because this directory sometimes is shared by multiple users, it is the most likely installed in a non-default directory.	

It is not required that you install SAS, CST, or this tutorial in their respective default directories. However, if you choose to install any of this software into a different directory, then you must modify the paths in:

- all of the SAS programs
- in the SAS References data set

The SAS References data set is created in Exercise 2A and updated in Exercise 3.

Exercise 1 – Prepare Study Data and Metadata

In this section you will generate the data and metadata for the fictitious study called MyStudy. Your data must be in CDISC SDTM 3.1.1-formatted SAS data sets. Also, you must have the supporting metadata files that describe the tables and columns for the domains. Therefore, one of the most critical parts of using CST is the preparation of your own data and metadata.

The preparation can vary depending on your organization's requirements, data, metadata, tools, capabilities, policies, and procedures. You might have everything you need already in place and the preparations are complete or it might take you and your co-workers months to prepare. To assist you with preparing your data and metadata, SAS Institute provides SAS® Clinical Data Integration. Clinical Data Integration integrates seamlessly with CST.

The exercises in this section are intended to give you a general idea about the *kinds of tasks* you must perform to generate your data and metadata for use with CST.

Exercise 1A – Prepare Data

To simulate the process of creating your data, copy the data from the sample study.

☐ Edit the following applicable SAS program and update any paths if needed.

For SAS 9.2	c:\cstQuickStart\Exercise1_PrepareStudyData\Exercise1A_PrepareMyStudyData.sas
For SAS 9.1.3	c:\cstQuickStart913\ Exercise1_PrepareStudyData\Exercise1A_PrepareMyStudyData.sas

☐ Submit the edited SAS program.

This program copies several study data sets into the MyData subdirectory. Each data set represents one SDTM domain.

Exercise 1B – Prepare Format Catalogs

To simulate the process of creating SAS formats, which can include controlled terminology, copy the SAS catalogs from the sample study.

☐ Edit the following applicable SAS program and update any paths if needed.

For SAS 9.2	c:\cstQuickStart\Exercise1_PrepareStudyData\Exercise1B_PrepareMyStudyFormats.sas
For SAS 9.1.3	c:\cstQuickStart913\ Exercise1_PrepareStudyData\Exercise1B_PrepareMyStudyFormats.sas

☐ Submit the edited SAS program.

This program copies several SAS catalogs to the MyFormats subdirectory.

Exercise 1C – Prepare Metadata Files

To simulate the process of creating the metadata files that are specific to MyStudy (that is, the source metadata), copy the SAS standardized set of metadata (that is, the reference metadata).

The purpose of study-level metadata (source metadata) is to enable you to use a modified version of the SDTM formats (for example, SDTM+) or your own internal standards.

☐ Edit the following applicable SAS program and update any paths if needed.

For SAS 9.2	c:\cstQuickStart\Exercise1_PrepareStudyData\Exercise1C_PrepareMyStudyMetadata.sas
For SAS 9.1.3	c:\cstQuickStart913\ Exercise1_PrepareStudyData\ Exercise1C_PrepareMyStudyMetadata.sas

☐ Submit the edited SAS program.

This program copies two SAS data sets into the MyMetadata subdirectory. The MyTableMetadata.sas7bdat file contains the metadata required to define the tables (domains) for the given study. The MyColumnMetadata.sas7bdat file contains the metadata required to define the columns within the tables (domains.)

Unlike the earlier steps, you made some changes to the contents of these two data sets rather than copying them unchanged. These changes represent the minimum required to use a copy of the CST reference metadata as your source metadata.

If you are validating against the standard SAS CDISC SDTM 3.1.1 format, then the process shown in the Exercise1C_PrapreMyStudyMetadata.sas program might be all that you require for validation.

Exercise 2 – Step by Step

In this exercise, you will explore the tasks required to run a CST validation program. You will complete the steps necessary to execute a CST SDTM validation against the data that was configured in Exercise 1.

See the *SAS Clinical Standards Toolkit 1.2: User's Guide* for more information about each of these steps.

Exercise 2A – Create a SAS References data set

The SAS References data set is used to pass parameters to CST. It defines the SAS libraries, data sets, and other files that CST requires. The SAS References data set **_must_** contain a row that references itself.

The full details of the SAS References data set are beyond the scope of this document. See the *SAS Clinical Standards Toolkit 1.2: User's Guide* for more information.

☐ Edit the following applicable SAS program and update any paths if needed.

For SAS 9.2	c:\cstQuickStart\Exercise2_StepByStep\Exercise2A_CreateSASReferences.sas
For SAS 9.1.3	c:\cstQuickStart913\Exercise2_StepByStep\Exercise2A_CreateSASReferences.sas

☐ Submit the edited SAS program.

The SAS References data set is written to the study-level directory in the following file:

For SAS 9.2	c:\cstQuickStart\Exercise2_StepByStep\ SASReferences.sas7bdat
For SAS 9.1.3	c:\cstQuickStart913\Exercise2_StepByStep\ SASReferences.sas7bdat

You will reuse this SAS References data set in Exercise 3.

Exercise 2B – Select the Validation Checks

In this exercise, you will select the validation checks that you wish to run against the MyStudy data. The validation_master data set contains the complete list of checks that the CST predefines. One row is equal to one check. When running a validation job, you create a subset of the validation_master data set. This resulting data set is called the validation_control data set.

The validation_control data set contains the list of checks that are executed by the validation job. The validation_master is not used directly by the validation process itself.

Some rows in the validation_master data set appear to conduct the same validation check, that is, some of the tests performed by the checks are repeated by multiple checks. However, upon close review, you can see that these rows do differ in other details. The differences often occur in the source of the check (for example, WebSDM orJANUS) or in the level of severity that is reported when the data fails the check

(for example, WARNING or ERROR.) This enables greater flexibility in choosing which checks you want to run. For example, you can choose to run all of the JANUS checks and only the JANUS checks.

Note: You also can add your own checks to CST, but this capability is beyond the scope of this tutorial.

☐ Edit the following applicable SAS program and update any paths if needed.

| For SAS 9.2 | c:\cstQuickStart\Exercise2_StepByStep\Exercise2B_ChooseChecks.sas |
| For SAS 9.1.3 | c:\cstQuickStart913\Exercise2_StepByStep\Exercise2B_ChooseChecks.sas |

☐ Submit the edited SAS program.

The SAS References data set is written to the study-level directory in the following file:

| For SAS 9.2 | c:\cstQuickStart\Exercise2_StepByStep\checkstorun.sas7bdat |
| For SAS 9.1.3 | c:\cstQuickStart913\Exercise2_StepByStep\checkstorun.sas7bdat |

This data set contains two rows that are similar. Both rows test for the same condition. However, one check is defined by JANUS , which reports the condition as a note. The other check is defined by WebSDM, which reports the condition as warning. Typically, in a production system, you would choose to run only one of these checks.

Exercise 2C – Run Validation

In this exercise, you will validate MyStudy data against CDISC SDTM 3.1.1.

When reviewing the SAS code, notice several lines that are referencing properties. In addition to the SAS References and Control data sets, you can pass parameters into CST using properties. Properties are options that impact how CST performs various tasks. For the purposes of this tutorial, you will use only default values. However, the tutorial does include an example properties file, MyValidation.properties, in the MyStudy folder. For more information about properties, see the *SAS Clinical Standards Toolkit 1.2: User's Guide*.

To run the validation, complete the following steps:

☐ Edit the following applicable SAS program and update any paths if needed.

| For SAS 9.2 | c:\cstQuickStart\Exercise2_StepByStep\Exercise2C_RunValidation.sas |
| For SAS 9.1.3 | c:\cstQuickStart913\Exercise2_StepByStep\Exercise2C_RunValidation.sas |

☐ Submit the edited SAS program.

The SAS References data set is written to the study-level directory in the following file:

| For SAS 9.2 | c:\cstQuickStart\Exercise2_StepByStep\validationresult.sas7bdat |
| For SAS 9.1.3 | c:\cstQuickStart913\Exercise2_StepByStep\validationresult.sas7bdat |

The results data set contains the findings reported by the validation process.

The following data set file contains some information about the validation process itself, for example, the number of checks the process ran:

For SAS 9.2	c:\cstQuickStart\Exercise2_StepByStep\validationmetrics.sas7bdat
For SAS 9.1.3	c:\cstQuickStart913\Exercise2_StepByStep\ validationmetrics.sas7bdat

Exercise 3 – All in One

In this exercise, you will combine all of the steps from Exercise 2 into a single program.

The SAS References data set created in Exercise 2A is very close to what you need in this exercise. You will update the data set by changing a few values.

This program also demonstrates how to run all of the JANUS checks, but for the purposes of this exercise, you will run only the first few JANUS checks.

☐ Edit the following applicable SAS program and update any paths if needed.

For SAS 9.2	c:\cstQuickStart\Exercise3_AllInOne\Exercise3_RunCSTValidation.sas
For SAS 9.1.3	c:\cstQuickStart913\Exercise3_AllInOne\Exercise3_RunCSTValidation.sas

☐ Submit the edited SAS program.

The following job-specific files are written to the Exercise3_AllInOne subdirectory:

- CST CDISC SDTM 3.1.1 validation job-specific input:
  - The SAS references data set file sasreferences.sas7bdat
  - The control checks data set file checks.sas7bdat
- CST CDISC SDTM 3.1.1 validation job-specific output:
  - The results data set file results.sas7bdat
  - The metrics data set file metrics.sas7bdat

Summary

This Quick Start tutorial provided an overview so you could learn the basic process for using Clinical Standards Toolkit. While this tutorial focused on CDISC SDTM 3.1.1 validation, other CST tasks, such as CDISC CRT-DDS validation, use a similar process. For detailed information, see the *SAS Clinical Standards Toolkit 1.2: User's Guide*.