Implementing, Managing, and Validating a Clinical Standard Using SAS Clinical Standards Toolkit 1.3®

Copyright © 2010, SAS Institute Inc. All rights reserved.

SAS Clinical StandardsToolkit

- What is SAS Clinical Standards Toolkit?
- SAS CST Framework
- Metadata Files
- Supported Standards
- Validating the SDTM Standard
- Future Functionality
- More Information

Copyright © 2010, SAS Institute Inc. All rights reserved.

What is SAS Clinical Toolkit?

- Focuses primarily on support of clinical research activities.
- Initially focusing on standards as defined by CDISC, but not limited to CDISC.
- A collection of “tools”, providing an initial set of standards and functionality that will evolve and grow with future updates and releases.
- Designed as an integral part of Clinical Data Integration (CDI), but is available to all SAS users as open source SAS Macros.

Copyright © 2010, SAS Institute Inc. All rights reserved.

What is SAS Clinical Toolkit?

- Can modify, but be careful! Macros!
- CDISC models available 3.1.1 and 3.1.2 (November 2010)
- Controlled Terminology 2008-08, 2010-03
- CRT-DDS and define.xml capability
- Completely customizable to fit various standards
- Free with 9.1.3 and 9.2

Copyright © 2010, SAS Institute Inc. All rights reserved.

SAS Clinical Standards Toolkit Framework

Consists of two distinct pieces (default installation):

- The components that are installed as part of SAS Foundation and shared files. (SAS Macros, Java JAR files, etc.)
  - *SASROOT:\ClinicalStandardsToolkitFramework\...*
  - *SASROOT:\SASClinicalStandardsToolkitSDTM312\...*
  - *SASROOT:\SASFoundation\9.2\cstframework\sasmacro*

- The global standards library that the framework macros operate on top of.
  - *C:\cstGlobalLibrary\...*

Copyright © 2010, SAS Institute Inc. All rights reserved.

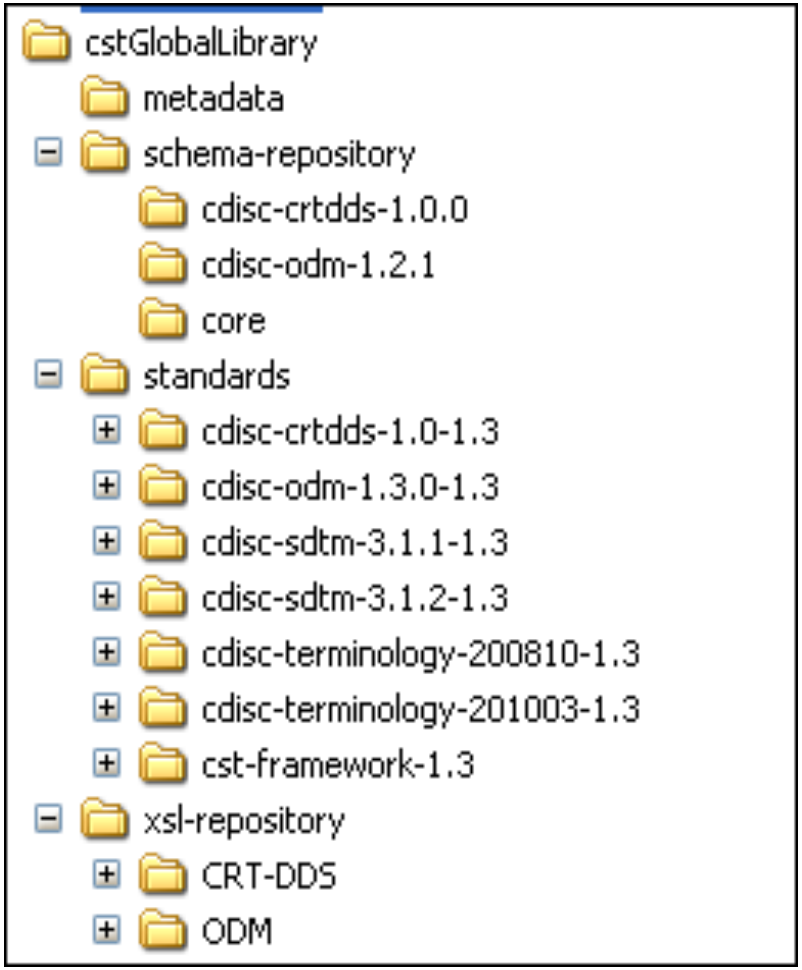
CST Framework – Global Standards Library

The Global Standards Library is created during installation. A series of directories are created here:

- **/metadata :** contains data sets that have information about the registered standard versions.
- **/schema-repository :** contains schemas for XML-based standards that are supported
- **/standards :** contains directories for each of the supported standards.
- **/xsl-repository :** contains directories and XSL files used in reading and writing XML files.

Copyright © 2010, SAS Institute Inc. All rights reserved.

CST Framework – Global Standards Library



Example of default installation of CST Global Library installed on C: drive. This is Toolkit 1.3

Copyright © 2010, SAS Institute Inc. All rights reserved.

SAS Clinical Standards Toolkit Framework

Some usage scenarios for the framework

- Initializing the global macro variables
- Referencing the default version of a standard
- Getting a list of installed standards
- Getting a list of files and data sets associated with a registered standard
- Creating data sets used by the framework
- Creating table shells based on a standard
- Getting a copy of the reference metadata for a data standard
- Inserting information from registered standards into a SASReferences file
- ….

Copyright © 2010, SAS Institute Inc. All rights reserved.

SAS Clinical Standards Toolkit Framework

Getting a list of files and data sets associated with a registered standard

When standards are registered so is the information about the files and data that make up the standard. Use the following macro call.

```
%cst_getStandardSASReferences( _cstStandard=CST-FRAMEWORK,
                               _cstStandardVersion=1.2,
                               _cstOutputDS=sasrefs);
```

	Name of Standard	Standard version	CST input/output data or metadata	Data or metadata subtype within type	SAS libref or fileref	Reference type (libref or fileref)	Relative path
1	CST-FRAMEWORK	1.2	control	reference	csttmp	libref	&_cstGRoot./standards/cst-framework-1.
2	CST-FRAMEWORK	1.2	control	validation	csttmp	libref	&_cstGRoot./standards/cst-framework-1.
3	CST-FRAMEWORK	1.2	lookup		control	libref	&_cstGRoot./standards/cst-framework-1.
4	CST-FRAMEWORK	1.2	messages		cstmsg	libref	&_cstGRoot./standards/cst-framework-1.
5	CST-FRAMEWORK	1.2	properties	initialize	cstprop	fileref	&_cstGRoot./standards/cst-framework-1.
6	CST-FRAMEWORK	1.2	results	metrics	csttmp	libref	&_cstGRoot./standards/cst-framework-1.
7	CST-FRAMEWORK	1.2	results	results	csttmp	libref	&_cstGRoot./standards/cst-framework-1.
8	CST-FRAMEWORK	1.2	standards	registeredsasreferences	csttmp	libref	&_cstGRoot./standards/cst-framework-1.
9	CST-FRAMEWORK	1.2	standards	registeredstandards	csttmp	libref	&_cstGRoot./standards/cst-framework-1.

Copyright © 2010, SAS Institute Inc. All rights reserved.

SAS Clinical Standards Toolkit Framework

Creating data sets used by the framework

Allows the user to create table shells used by the Toolkit, useful if creating a new standard. Use the following macro call.

```
%cst_CreateDS( _cstStandard=CST-FRAMEWORK,
               _cstStandardVersion=1.3,
               _cstType=control,
               _cstSubType=reference,
               _cstOutputDS=work.sasrefs);
```

The type and subtype identify it as a SASReferences table

	CST input/output data or metadata	Data or metadata subtype within type	SAS libref or fileref	Reference type (libref or fileref)	Relative path	Order within type (autocall,fmtseach)	Filename (null for libraries)
1	control	reference	csttmp	libref	&_cstGRoot./standards/cst-framework-1.	.	sasreferences
2	control	validation	csttmp	libref	&_cstGRoot./standards/cst-framework-1.	.	validation_master

Copyright © 2010, SAS Institute Inc. All rights reserved.

SAS Clinical Standards Toolkit Framework

Creating table shells based on a standard

Data standards, like SDTM, have reference metadata that describe the tables and columns making up the standard. To create table shells using this metadata, use the following macro call.

```
%cst_createTablesForDataStandard( _cstStandard=CDISC-SDTM,
                                  _cstOutputLibrary=work);
```

This creates the SDTM tables/domain shells depending on the default version of SDTM (3.1.1 or 3.1.2).

Copyright © 2010, SAS Institute Inc. All rights reserved.

SAS Clinical Standards Toolkit Framework

Framework Maintenance Macros

%cst_registerStandard()

%cst_setStandardVersionDefault()

%cst_unregisterStandard()

There are 16 SAS Framework macros (%cst_)

Copyright © 2010, SAS Institute Inc. All rights reserved.

SAS Clinical Standards Toolkit Metadata

Toolkit provides and uses a series of metadata files to support basic core functions of the toolkit.

Standards

StandardSASReferences

Standardlookup

SASReferences

Properties

Messages

Validation_Master(Validation_Control)

Reference_Tables (Source_Tables)

Reference_Columns (Source_Columns)

Validation_Metrics

Validation_Results

Copyright © 2010, SAS Institute Inc. All rights reserved.

SAS Clinical Standards Toolkit Metadata

Standards data set is used to store information about a standard version. One stored in the cstGlobalLibrary standards area for each standard and one stored containing all standards in the cstGlobalLibrary metadata area.

	standard	mnemonic	standardversion	comment	rootpath	isstandarddefault	iscstframework
1	CDISC-CRTDDS	CRT	1.0	CDISC CRT-DDS V1.0	&_cstGRoot./standards/cdisc-crtdds-1.0	Y	N
2	CDISC-SDTM	SDTM	3.1.1	CDISC SDTM V3.1.1	&_cstGRoot./standards/cdisc-sdtm-3.1.1	Y	N
3	CDISC-TERMINOLOGY	CT	200810	CDISC Terminology, Packages 1, 2A, 2B, LABTEST	&_cstGRoot/standards/cdisc-terminology	Y	N
4	CST-FRAMEWORK	CST	1.2	Clinical Standards Toolkit Framework	&_cstGRoot./standards/cst-framework	Y	Y

	standard	isdatastandard	supportsvalidation	isxmlstandard	importxsl	exportxsl	schema
1	CDISC-CRTDDS	Y	Y	Y	CRT-DDS/1.0/import/Root.xsl	CRT-DDS/1.0/export/Root.xsl	cdisc-crtdds-1.0.0/define1-0-0.xsd
2	CDISC-SDTM	Y	Y	N			
3	CDISC-TERMINOLOGY	N	N	N			
4	CST-FRAMEWORK	N	N	N			

Copyright © 2010, SAS Institute Inc. All rights reserved.

SAS Clinical Standards Toolkit Metadata

StandardSASReferences is identical in structure to ***sasreferences***, however has complete data. StandardSASReferences is used to "fill in the spots" when sasreferences has missing information. One stored at the standards level and one at the metadata level that contains all standards.

	standard	standardversion	type	subtype	SASref	reftype	path
5	CDISC-CRTDDS	1.0	referencemetadata	column	crtref	libref	&_cstGRoot/standards/cdisc-crtdds-1.0/metadata
6	CDISC-CRTDDS	1.0	referencemetadata	table	crtref	libref	&_cstGRoot/standards/cdisc-crtdds-1.0/metadata
7	CDISC-CRTDDS	1.0	referencexml	stylesheet	xslt01	fileref	&_cstGRoot/standards/cdisc-crtdds-1.0/stylesheet
8	CDISC-SDTM	3.1.1	autocall		sdtmauto	fileref	&_cstGRoot/standards/cdisc-sdtm-3.1.1/macros

	standard	standardversion	type	subtype	order	memname	comment
5	CDISC-CRTDDS	1.0	referencemetadata	column	.	reference_columns.sas7bdat	
6	CDISC-CRTDDS	1.0	referencemetadata	table	.	reference_tables.sas7bdat	
7	CDISC-CRTDDS	1.0	referencexml	stylesheet	1	define1-0-0.xsl	Default stylesheet to display define.xml
8	CDISC-SDTM	3.1.1	autocall		1		

Copyright © 2010, SAS Institute Inc. All rights reserved.

SAS Clinical Standards Toolkit Metadata

The SASReferences data set is the "brain center" of the Toolkit. Used to tell Toolkit where things are located. Can be created by the user or generated through the Toolkit. Many users will create their own SASReferences data set. Further discussion later.

standard	standardversion	type	subtype	SASref	reftype	path	order	memname	comment
CDISC-SDTM	3.1.1	sourcedata		srcdata	libref	&studyRootPath\data	.		Path to study-specific SDTM domain data sets
CDISC-SDTM	3.1.1	sourcemetadata	table	srcmeta	libref	&studyRootPath\metadata	.	source_tables.sas7bdat	Source of study-specific SDTM table metadata
CDISC-SDTM	3.1.1	sourcemetadata	column	srcmeta	libref	&studyRootPath\metadata	.	source_columns.sas7bdat	Source of study-specific SDTM column metadata
CDISC-SDTM	3.1.1	referencecontrol	validation	refcntl	libref		.		Derived from standardsasreferences
CDISC-SDTM	3.1.1	referencemetadata	table	refmeta	libref		.		Derived from standardsasreferences
CDISC-SDTM	3.1.1	referencemetadata	column	refmeta	libref		.		Derived from standardsasreferences
CDISC-SDTM	3.1.1	autocall		sdtmcode	fileref	&_cstGRoot\standards\cdisc-sdt	1		Standard-specific macro library
CDISC-SDTM	3.1.1	fmtsearch		srcfmt	libref	&studyRootPath\terminology\for	1	formats.sas7bcat	Standard- and study-specific formats
CDISC-TERM	200810	fmtsearch		cstfmt	libref	&_cstGRoot\standards\cdisc-ter	2	cterms.sas7bcat	Global Library formats
CUSTOM		referencecterm		ctref	libref	&studyRootPath\terminology\cod	1	meddra.sas7bdat	Customers coding dictionary location (may be a global standard)

Copyright © 2010, SAS Institute Inc. All rights reserved.

SAS Clinical Standards Toolkit Metadata

Property files (i.e initialize.properties) set default preferences for each process. Properties are a series of name-value pairs that are translated into global macro variables available for the duration of a Toolkit process. Invoked by the ***cst_setProperties*** macro.

```
_cstDebug=0
_cstDebugOptions=mprint mlogic symbolgen mautolocdisplay
_cst_rc=0
_cst_MsgID=
_cst_MsgParm1=
_cst_MsgParm2=
_cstResultSeq=0
_cstSeqCnt=0
_cstSrcData=
_cstResultFlag=0
_cstResultsDS=work._cstresults
_cstMessages=work._cstmessages
_cstReallocateSASRefs=0
_cstFMTLibraries=work
_cstMessageOrder=APPEND
_cstSASRefsLoc=
_cstSASRefsName=
_cstSASRefs=work._cstsasrefs
```

Copyright © 2010, SAS Institute Inc. All rights reserved.

SAS Clinical Standards Toolkit Metadata

Messages data sets are used to store information about the framework and standards validation checks. There are framework messages (CST prefix) and standards validation messages (e.g SDTM prefix)

	Result identifier	Standard version	Source of check	Record identifier used by checksource	Severity of check	Rule description from checksource	Message text
1	CST0001	***	CST	CST0001	Error		Fatal error encountered, process cannot continue
2	CST0002	***	CST	CST0002	Warning: Check not run		No tables evaluated-check validation control data set
3	CST0003	***	CST	CST0003	Warning: Check not run		&_cstparm1 could not be found
4	CST0004	***	CST	CST0004	Warning: Check not run		No columns evaluated - check validation_control specification

Result identifier	Standard version	Source of check	Record identifier used by checksource	Severity of check	Rule description from checksource	Message text
SDTM0001	***	Janus	IR4000	Note	Identifies domain table that has zero rows and therefore contains no data	Domain &_cstparm1 contains 0 observations or is missing
SDTM0001	***	WebSDM	IR4000	Warning	Identifies domain table that has zero rows and therefore contains no data	Domain &_cstparm1 contains 0 observations or is missing
SDTM0002	***	JanusFR	SAS0017	Error	A load of data into JANUS requires that the DM, DS and EX domains be submitted for each study to be loaded.	Missing (or empty) DM, DS or EX domain
SDTM0003	***	WebSDM	SAS0018	Error	WebSDM and the SDTM model require only the DM domain be present.	Missing (or empty) DM domain
SDTM0004	***	SAS	SAS0033	Note	Source metadata includes domain data set not found in reference metadata	Study data set not found in reference standard

Copyright © 2010, SAS Institute Inc. All rights reserved.

SAS Clinical Standards Toolkit Metadata

Results data set. Each Toolkit process generates a results data set that can optionally be persisted beyond the SAS session based on SASReferences data set settings. Each results data set captures the outcome of specific process actions, using the ***messages*** data sets to standardize output.

	Result identifier	Validation check identifier	Unique invocation of resultid	Sequence number within resultseq	Source data	Resolved message text from message file	Result severity (e.g., warning, error)	Problem detected? (0=no, otherwise yes)	Process status (Non-zero, aborted)
24	CST0200		1	8	SDTM_VALIDATE	PROCESS GLOBALLIBRARY: c:/cstGlobalLibrary	Info	0	0
25	CST0200		1	9	SDTM_VALIDATE	PROCESS CSTVERSION: 1.3	Info	0	0
26	CST0100	SDTM0001	1	1	SRCDATA.AE	No errors detected in SRCDATA.AE	Info	0	0
27	CST0100	SDTM0001	1	1	SRCDATA.AE	No errors detected in SRCDATA.AE	Info	0	0
28	SDTM0001	SDTM0001	1	2	SRCDATA.CE	Domain SRCDATA.CE contains 0 observations or is missing	Warning	1	0

Copyright © 2010, SAS Institute Inc. All rights reserved.

SAS Clinical Standards Toolkit Metadata

Validation_Master & Validation_Control

Validation_master contains ALL **standard specific** validation information. Validation_control contains **study specific** validation information and is created from validation_master. The "brain center" of the Toolkit validation process. Will discuss later.

	Validation check identifier	Standard model	Standard version	Source of check	Record identifier used by checksource	Severity of check	Category of check	SAS macro module name	Check should use source metadata
1	SDTM0001	CDISC-SDTM	***	Janus	IR4000	Note	Metadata	cstcheck_zeroobs	Y
2	SDTM0001	CDISC-SDTM	***	WebSDM	IR4000	Warning	Metadata	cstcheck_zeroobs	Y
3	SDTM0002	CDISC-SDTM	***	JanusFR	SAS0017	Error	Metadata	cstcheck_zeroobs	N
4	SDTM0003	CDISC-SDTM	***	WebSDM	SAS0018	Error	Metadata	cstcheck_zeroobs	N
5	SDTM0004	CDISC-SDTM	***	SAS	SAS0033	Note	Metadata	cstcheck_dsmismatch	Y

	Validation check identifier	Standard model	Standard version	Source of check	Record identifier used by checksource	Severity of check	Category of check	SAS macro module name	Check should use source metadata
1	SDTM0001	CDISC-SDTM	***	WebSDM	IR5000	Warning	Metadata	cstcheck_zeroobs	Y
2	SDTM0201	CDISC-SDTM	***	WebSDM	IR5001	Error	Column	cstcheck_violatesstd	Y
3	SDTM0101	CDISC-SDTM	***	WebSDM	IR5002	Warning	Date	cstcheck_column	Y

Copyright © 2010, SAS Institute Inc. All rights reserved.

SAS Clinical Standards Toolkit Metadata

Reference_Tables & Source_Tables

Reference_tables contains ALL **standard specific** table/ domain information. Source_tables contains **study specific** table/domain information and is created from reference_tables. Determines tables to be validated. For SDTM, information is gathered from the CDISC SDTM Implementation Guide.

SASreferences sourcedata libref	Table Name	Table Label	Observation Class within Standard	(Relative) path to xpt file	Title for xpt file	Table Structure	Purpose
REFDATA	AE	Adverse Events	Events	..\ae.xpt	Adverse Events SAS transport file	One record per event per subject	Tabulation
REFDATA	CM	Concomitant Medications	Interventions	..\cm.xpt	Concomitant Medications SAS transport file	One record per medication intervention episode per subject	Tabulation
REFDATA	CO	Comments	Special Purpose	..\co.xpt	Comments SAS transport file	One record per comment per subject	Tabulation
REFDATA	DM	Demographics	Special Purpose	..\dm.xpt	Demographics SAS transport file	One record per subject	Tabulation
REFDATA	DS	Disposition	Events	..\ds.xpt	Disposition SAS transport file	One record per disposition status or protocol milestone per subject	Tabulation
REFDATA	DV	Protocol Deviations	Events	..\dv.xpt	Protocol Deviations SAS transport file	One record per protocol deviation per subject	Tabulation

Copyright © 2010, SAS Institute Inc. All rights reserved.

SAS Clinical Standards Toolkit Metadata

Reference_Columns & Source_Columns

Reference_columns contains all **standard specific** table/domain column information. Source_columns contains **study specific** table/domain column information and is created from reference_columns. Determines columns to be validated. For SDTM, information is gathered from the CDISC SDTM Implementation Guide.

SASreferences sourcedata libref	Table Name	Column Name	Column Description	Column Order	Column Type	Column Length	Display Format	XML Data Type	SAS Format/XML Codelist	Column Required or Optional
REFDATA	AE	STUDYID	Study Identifier	1	C	40		text		Req
REFDATA	AE	DOMAIN	Domain Abbreviation	2	C	8		text		Req
REFDATA	AE	USUBJID	Unique Subject Identifier	3	C	40		text		Req
REFDATA	AE	AESEQ	Sequence Number	4	N	8		integer		Req
REFDATA	AE	AEGRPID	Group ID	5	C	40		text		Perm
REFDATA	AE	AEREFID	Reference ID	6	C	40		text		Perm
REFDATA	AE	AESPID	Sponsor-Defined Identifier	7	C	40		text		Perm

Copyright © 2010, SAS Institute Inc. All rights reserved.

SAS Clinical Standards Toolkit Supported Standards

Toolkit 1.3 shipped with:

- CDISC-SDTM 3.1.1
- CDISC-SDTM 3.1.2
- CDISC-CRTDDS 1.0
- CDISC-Terminology 200810
- CDISC-Terminology 201003

Focus Area (For 1.2 users)

- CDISC-Terminology 201003
- CDISC-SDTM 3.1.2 Domains

Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

- Open a SAS session with toolkit installed
- Open validate_data from SDTM sample study
- Submit

Log - (Untitled)

```
NOTE: Copyright (c) 2002-2008 by SAS Institute Inc., Cary, NC, USA.
NOTE: SAS (r) Proprietary Software 9.2 (TS2M2)
      Licensed to Lightfoot Aubrey E-5774, Site 70068112.
NOTE: This session is executing on the XP_PRO  platform.


NOTE: SAS initialization used:
      real time           1.34 seconds
      cpu time            0.87 seconds
```

validate_data.sas

```
*******************************************************************************;
* validate_data.sas                                                           *;
*                                                                             *;
* Sample driver program to perform a primary Toolkit action, in this case,    *;
* validation, to assess compliance of some source data and metadata with a    *;
* registered standard.  A call to a standard-specific validation macro is     *;
* required later in this code.                                                *;
*                                                                             *;
* Assumptions:                                                                *;
*   The SASReferences file must exist, and must be identified in the call to  *;
*    cstutil_processsetup if it is not work.sasreferences.                    *;
*                                                                             *;
* CSTversion  1.3                                                             *;
*******************************************************************************;

*******************************************************************************;
* The following %let statement sets the studyrootpath.Studyrootpath is used to make *;
```

Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

- Runs successfully
- No errors or warning detected in the log
- Open validation_results data set in the results library
- Review the checks

VIEWTABLE: Results.Validation_results

	Result identifier	Validation check identifier	Unique invocation of resultid	Sequence number within resultseq	Source data	Resolved message text from message file	Result severity (e.g., warning, error)	Problem detected? (0=no, otherwise yes)	Process status (Non-zero, aborted)	Actual
28	SDTM0001	SDTM0001	1	14	SRCDATA.MB	Domain SRCDATA.MB contains 0 observations or is missing	Medium	1	0	
29	CST0100	SDTM0001	1	15	SRCDATA.MH	No errors detected in SRCDATA.MH	Info	0	0	
30	SDTM0001	SDTM0001	1	16	SRCDATA.MS	Domain SRCDATA.MS contains 0 observations or is missing	Medium	1	0	
31	SDTM0001	SDTM0001	1	17	SRCDATA.PC	Domain SRCDATA.PC contains 0 observations or is missing	Medium	1	0	
32	CST0100	SDTM0001	1	18	SRCDATA.PE	No errors detected in SRCDATA.PE	Info	0	0	
33	SDTM0001	SDTM0001	1	19	SRCDATA.PP	Domain SRCDATA.PP contains 0 observations or is missing	Medium	1	0	
34	SDTM0001	SDTM0001	1	20	SRCDATA.QS	Domain SRCDATA.QS contains 0 observations or is missing	Medium	1	0	
35	CST0100	SDTM0001	1	21	SRCDATA.RELREC	No errors detected in SRCDATA.RELREC	Info	0	0	
36	CST0100	SDTM0001	1	22	SRCDATA.SC	No errors detected in SRCDATA.SC	Info	0	0	
37	CST0100	SDTM0001	1	23	SRCDATA.SE	No errors detected in SRCDATA.SE	Info	0	0	
38	CST0100	SDTM0001	1	24	SRCDATA.SU	No errors detected in SRCDATA.SU	Info	0	0	
39	CST0100	SDTM0001	1	25	SRCDATA.SUPPAE	No errors detected in SRCDATA.SUPPAE	Info	0	0	
40	SDTM0001	SDTM0001	1	26	SRCDATA.SUPPQUAL	Domain SRCDATA.SUPPQUAL contains 0 observations or is missing	Medium	1	0	
41	CST0100	SDTM0001	1	27	SRCDATA.SV	No errors detected in SRCDATA.SV	Info	0	0	
42	CST0100	SDTM0001	1	28	SRCDATA.TA	No errors detected in SRCDATA.TA	Info	0	0	
43	CST0100	SDTM0001	1	29	SRCDATA.TE	No errors detected in SRCDATA.TE	Info	0	0	
44	CST0100	SDTM0001	1	30	SRCDATA.TI	No errors detected in SRCDATA.TI	Info	0	0	
45	CST0100	SDTM0001	1	31	SRCDATA.TS	No errors detected in SRCDATA.TS	Info	0	0	
46	SDTM0001	SDTM0001	1	32	SRCDATA.TV	Domain SRCDATA.TV contains 0 observations or is missing	Medium	1	0	

Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

So, what drives this??

- "Driver" program (validate_data.sas)
- SASReferences data set (the "brain center" of the Toolkit
- Validation_control data set (the "brain center" of the validation process)
- Source Data – the actual study data
- Source Metadata
  - Source_tables, Source_columns
- Reference Metadata
  - Reference_tables, Reference_columns
- Messages Data

Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

Driver programs

- Toolkit is shipped with driver programs to facilitate Toolkit processes outside of CDI
- Located in the sample study as **validate_data.sas**
- Initializes information vital to the macro calls
- Runs in base SAS
- Modifiable by the user for other standards or studies
- 17 Driver programs shipped with Toolkit 1.3

Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

Validate_data.sas

```
%let studyRootPath=!sasroot/../../
    SASClinicalStandardsToolkitSDTM312/1.3 /sample/cdisc-sdtm-3.1.2/
    sascstdemodata;

%let _cstSetupSrc=SASREFERENCES;

%let workPath=%sysfunc(pathname(work));

%let _cstStandard=CDISC-SDTM;

%let _cstStandardVersion=3.1.2;
```

%SDTM_VALIDATE; (same as used by CDI)

Examples above are from a snippet of code within validate_data.sas

Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

SASReferences data set

- The roadmap and “Brain Center” for any Toolkit process
- At this point all required fields must be populated

Name of Standard	Standard version	CST input/output data or metadata	Data or metadata subtype within type	SAS libref or fileref	Reference type (libref or fileref)	Relative path	Order within type (autocall,fmtseach)	Filename (null for libraries)
CDISC-SDTM	3.1.2	autocall		sdtmauto	fileref	&_cstGRoot./standards/cdisc-sdtm-3.1.2-1.3/macros	1	
CDISC-SDTM	3.1.2	control	reference	work	libref	C:\DOCUME~1\GENELI~1.IBI\LOCALS~1\Temp\SAS Temporary Files_TD3696	1	sasreferences.sas7bdat
CDISC-SDTM	3.1.2	control	validation	control	libref	!sasroot/../../SASClinicalStandardsToolkitSDTM312/1.3/sample/cdisc-sdtm-3.1.2/sascstdemodata/control	2	validation_control.sas7bdat
CDISC-SDTM	3.1.2	fmtsearch		srcfmt	libref	!sasroot/../../SASClinicalStandardsToolkitSDTM312/1.3/sample/cdisc-sdtm-3.1.2/sascstdemodata/termino	1	formats.sas7bcat
CDISC-TERMINOLOGY	201003	fmtsearch		cstfmt	libref	c:/cstGlobalLibrary/standards/cdisc-terminology-201003-1.3/formats	2	cterms.sas7bcat
CST-FRAMEWORK	1.2	messages		messages	libref	&_cstGRoot./standards/cst-framework-1.3/messages	1	messages
CDISC-SDTM	3.1.2	messages		sdtmmsg	libref	&_cstGRoot./standards/cdisc-sdtm-3.1.2-1.3/messages	2	messages.sas7bdat
CDISC-SDTM	3.1.2	properties	initialize	inprop	fileref	&_cstGRoot./standards/cdisc-sdtm-3.1.2-1.3/programs	1	initialize.properties
CDISC-SDTM	3.1.2	properties	validation	valprop	fileref	!sasroot/../../SASClinicalStandardsToolkitSDTM312/1.3/sample/cdisc-sdtm-3.1.2/sascstdemodata/progra	2	validation.properties
CDISC-SDTM	3.1.2	referencecontrol	standardref	refcntl	libref	&_cstGRoot./standards/cdisc-sdtm-3.1.2-1.3/validation/control	.	validation_stdref.sas7bdat
CDISC-SDTM	3.1.2	referencecontrol	validation	refcntl	libref	&_cstGRoot./standards/cdisc-sdtm-3.1.2-1.3/validation/control	.	validation_master.sas7bdat
CDISC-SDTM	3.1.2	referencemetadata	column	refmeta	libref	&_cstGRoot./standards/cdisc-sdtm-3.1.2-1.3/metadata	.	reference_columns.sas7bdat
CDISC-SDTM	3.1.2	referencemetadata	table	refmeta	libref	&_cstGRoot./standards/cdisc-sdtm-3.1.2-1.3/metadata	.	reference_tables.sas7bdat
CDISC-SDTM	3.1.2	results	validationmetrics	results	libref	!sasroot/../../SASClinicalStandardsToolkitSDTM312/1.3/sample/cdisc-sdtm-3.1.2/sascstdemodata/results	.	validation_metrics.sas7bdat
CDISC-SDTM	3.1.2	results	validationresults	results	libref	!sasroot/../../SASClinicalStandardsToolkitSDTM312/1.3/sample/cdisc-sdtm-3.1.2/sascstdemodata/results	.	validation_results.sas7bdat
CDISC-SDTM	3.1.2	sourcedata		srcdata	libref	!sasroot/../../SASClinicalStandardsToolkitSDTM312/1.3/sample/cdisc-sdtm-3.1.2/sascstdemodata/data	.	
CDISC-SDTM	3.1.2	sourcemetadata	column	srcmeta	libref	!sasroot/../../SASClinicalStandardsToolkitSDTM312/1.3/sample/cdisc-sdtm-3.1.2/sascstdemodata/metad	.	source_columns.sas7bdat
CDISC-SDTM	3.1.2	sourcemetadata	table	srcmeta	libref	!sasroot/../../SASClinicalStandardsToolkitSDTM312/1.3/sample/cdisc-sdtm-3.1.2/sascstdemodata/metad	.	source_tables.sas7bdat

Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

Validation_control data set

- Study level
- Controls the validations being submitted to the Toolkit
  - The "Brain Center" for the validation checks
- Derived from validation_master data set
- Allows the user to pick and choose the validations needed to implement.
- Modifiable by the user.

Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

Validation_control data set

- The control center for the validation process

Validation check identifier	Standard model	Standard version	Source of check	Record identifier used by checksource	Severity of check	Category of check	SAS macro module name	Check should use source metadata
SDTM0034	CDISC-SDTM	3.1.2	WebSDM	IR5251	Warning	Metadata	cstcheck_notimplemented	Y
SDTM0209	CDISC-SDTM	***	WebSDM	IR5100	Error	Date	cstcheck_columncompare	Y
SDTM0221	CDISC-SDTM	***	WebSDM	IR5136	Note	CntlTerm	cstcheck_notincodelist	Y

Validation check identifier	Domains/data sets to which check applies	Columns to which check applies	Code logic used within code	Code logic type	Lookup standard type	SAS format name	Reference in standard supporting check
SDTM0034	_ALL_	_ALL_	**not implemented**;	1			
SDTM0209	_ALL_-DS	[**STDY][**ENDY]	data work._cstProblems; set &_cstDSName.; if (.<&_cstColumn2<&_cstColumn1); run;	2			
SDTM0221	_ALL_	_ALL_		0	METADATA		

	Validation check identifier	Column values to be reported	Current check status	Report all possible records in error	Unique check identifier
1	SDTM0034		-2	Y	SDTM003401CST130SDTM3122010-05-07T14:45:00CST
2	SDTM0209		1	Y	SDTM020901CST130SDTM3122010-05-07T14:45:00CST
3	SDTM0221		1	Y	SDTM022101CST130SDTM3122010-05-07T14:45:00CST

Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

Validation check identifier	SAS macro module name	Check should use source metadata	Domains/data sets to which check applies	Columns to which check applies	Code logic used within code
SDTM0034	cstcheck_notimplemented	Y	_ALL_	_ALL_	**not implemented**;
SDTM0209	cstcheck_columncompare	Y	_ALL_-DS	[**STDY][**ENDY]	data work._cstProblems; set &_cstDSName.; if (.<&_cstColumn2<&_cstColumn1); run;
SDTM0221	cstcheck_notincodelist	Y	_ALL_	_ALL_	

Domains/data sets to which check applies

- Determines tables to process (&_cstDSName)
- Conventions – what can be used is macro specific
  - _ALL_ — All domains/tables
  - SUPP** — All SUPPAE, SUPPLB, etc.
  - CLASS:FINDINGS — All Findings domain classifications
  - TS — Single domain
  - _ALL_-DS-LB-PC — All domains EXCEPT DS, LB, and PC
  - CE+CM+SU — Three domains CE, CM, & SU
  - [DM][DS] — Paired - check values from DM with DS
  - [_ALL_-DM][DM] — Paired – check all except DM against DM

Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

Validation check identifier	SAS macro module name	Check should use source metadata	Domains/data sets to which check applies	Columns to which check applies	Code logic used within code
SDTM0034	cstcheck_notimplemented	Y	_ALL_	_ALL_	**not implemented**;
SDTM0209	cstcheck_columncompare	Y	_ALL_-DS	[**STDY][**ENDY]	data work._cstProblems; set &_cstDSName.; if (.<&_cstColumn2<&_cstColumn1); run;
SDTM0221	cstcheck_notincodelist	Y	_ALL_	_ALL_	

Columns to which check applies

- Determines the column(s) to process (&_cstColumn)
- Conventions – what can be used is macro specific
  - _ALL_ — All columns
  - **DUR — Prefaced with domain name (AEDUR, etc)
  - TSPARM — Single column
  - **DY+**STDY+VISITDY — Three columns DMDY, SVSTDY, & VISITDY
  - [**STDY][**ENDY] — Paired - check values from AESTDY with AEENDY
- Use source metadata - if Y uses Source_Tables, Source_Columns (Study Level)

Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

SAS macro module name
cstcheck_notimplemented
cstcheck_columncompare
cstcheck_notincodelist

Code logic used within code
not implemented;
data work._cstProblems; set &_cstDSName.; if (.<&_cstColumn2<&_cstColumn1); run;

SAS Macro Module Name

- 14 cstcheck_* validation SAS macros
- This column directs which macro will be used per edit check
- May require Code Logic:
  - At the statement level
  - At the data step (or proc sql) level
  - May not

Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

Source_tables data set

- Study level
- Determines which domains/tables are available for validation
- Derived from reference_tables data set
- Allows the user to pick and choose the tables to run against
- Modifiable by the user.

Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

Source_tables data set

- Contains Study tables
- Validates the standard – if not here will not be submitted

	SASreferences sourcedata libref	Table Name	Table Label	Observation Class within Standard	(Relative) path to xpt file	Title for xpt file	Table Structure	Purpose	Table Keys	Date Set State (Final, Draft)	Release Date
1	SRCDATA	AE	Adverse Events	Events	..\ae.xpt	Adverse Events SAS transport file	One record per adverse event per subject	Tabulation	STUDYID USUBJID AEDECOD AESTDTC	Final	November 12, 2008
2	SRCDATA	CE	Clinical Events	Events	..\ce.xpt	Clinical Events SAS transport file	One record per event per subject	Tabulation	STUDYID USUBJID CETERM CESTDTC	Final	November 12, 2008
3	SRCDATA	CM	Concomitant Medications	Interventions	..\cm.xpt	Concomitant Medications SAS transport file	One record per recorded medication occurrence or constant-dosing interval per subject	Tabulation	STUDYID USUBJID CMTRT CMSTDTC	Final	November 12, 2008

	SASreferences sourcedata libref	Table Name	Name of Standard	Version of Standard	Associated reference(s) in Standard	Comment
1	SRCDATA	AE	CDISC-SDTM	3.1.2		
2	SRCDATA	CE	CDISC-SDTM	3.1.2		
3	SRCDATA	CM	CDISC-SDTM	3.1.2		

Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

Source_columns data set

- Study level
- Determines which columns are available for validation
- Derived from reference_columns data set
- Allows the user to pick and choose the columns to run against
- Modifiable by the user.

Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

Source_columns data set

- Contains Study columns
- Validates the standard – if not here will not be submitted

	SASreferences sourcedata libref	Table Name	Column Name	Column Description	Column Order	Column Type	Column Length	Display Format	XML Data Type	SAS Format/XML Codelist	Column Required or Optional
1	SRCDATA	AE	STUDYID	Study Identifier	1	C	40		text		Req
2	SRCDATA	AE	DOMAIN	Domain Abbreviation	2	C	8		text		Req
3	SRCDATA	AE	USUBJID	Unique Subject Identifier	3	C	40		text		Req
4	SRCDATA	AE	AESEQ	Sequence Number	4	N	8		integer		Req

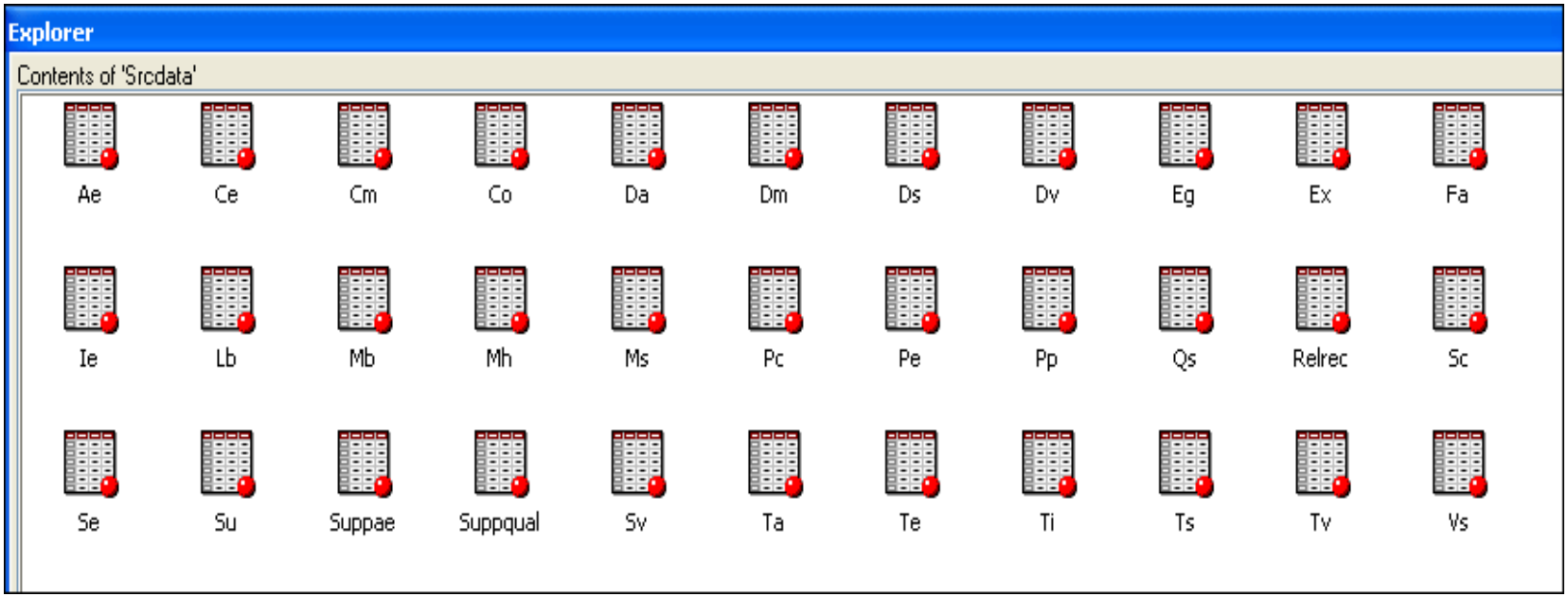
	SASreferences sourcedata libref	Table Name	Column Role	Controlled Term or Format in Standard	Computational Algorithm or Method	Column qualifiers (space delimited)	Name of Standard	Version of Standard	Associated reference(s) in Standard	Comment
1	SRCDATA	AE	Identifier			UPPERCASE	CDISC-SDTM	3.1.2	SDTM2.2.4	Unique identifier for a study.
2	SRCDATA	AE	Identifier	AE		UPPERCASE	CDISC-SDTM	3.1.2	SDTM2.2.4,SDTMIG4.1.2.2,AppendixC2	Two-character abbreviation for the domain.
3	SRCDATA	AE	Identifier			UPPERCASE	CDISC-SDTM	3.1.2	SDTM2.2.4,SDTMIG4.1.2.3	Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product.

Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

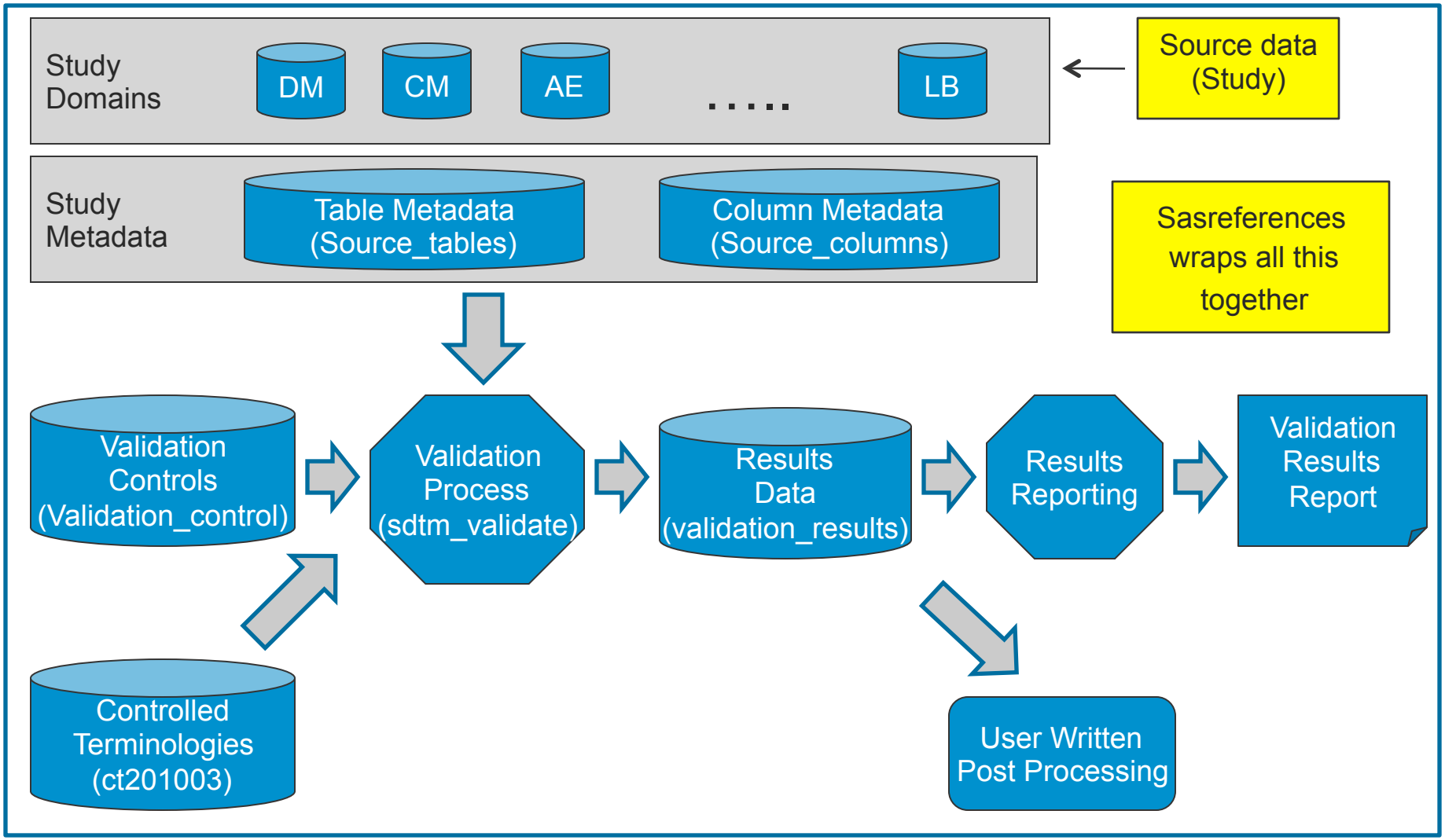
Source Data – The actual study data

- SRCDATA libname
- Sample data for this example
- C:\Program Files\SAS\SASClinicalStandardsToolkitSDTM312\1.3\sample\cdisc-sdtm-3.1.2\sascstdemodata\data



Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM



Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

Customization – Modifying a check

- Validation_contol, Source_tables, Source_columns, Messages.
- Modify codelogic field
- Modify columns and/or tables field if needed
- Modify Source metadata if needed
- Modify message in Messages data set if needed.
- Test
- Move to Validation_master if needed for other studies

Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

Customization – Adding a check using existing macros

- Validation_contol, Source_tables, Source_columns, Messages.
- Locate similar edit and copy it within validation_control.
- Assign unique check number.
- Insert new information in codelogic field, columns, tables
- Modify Source metadata if needed
- Add new message to Messages data set.
- Test
- Move to Validation_master if needed for other studies

Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

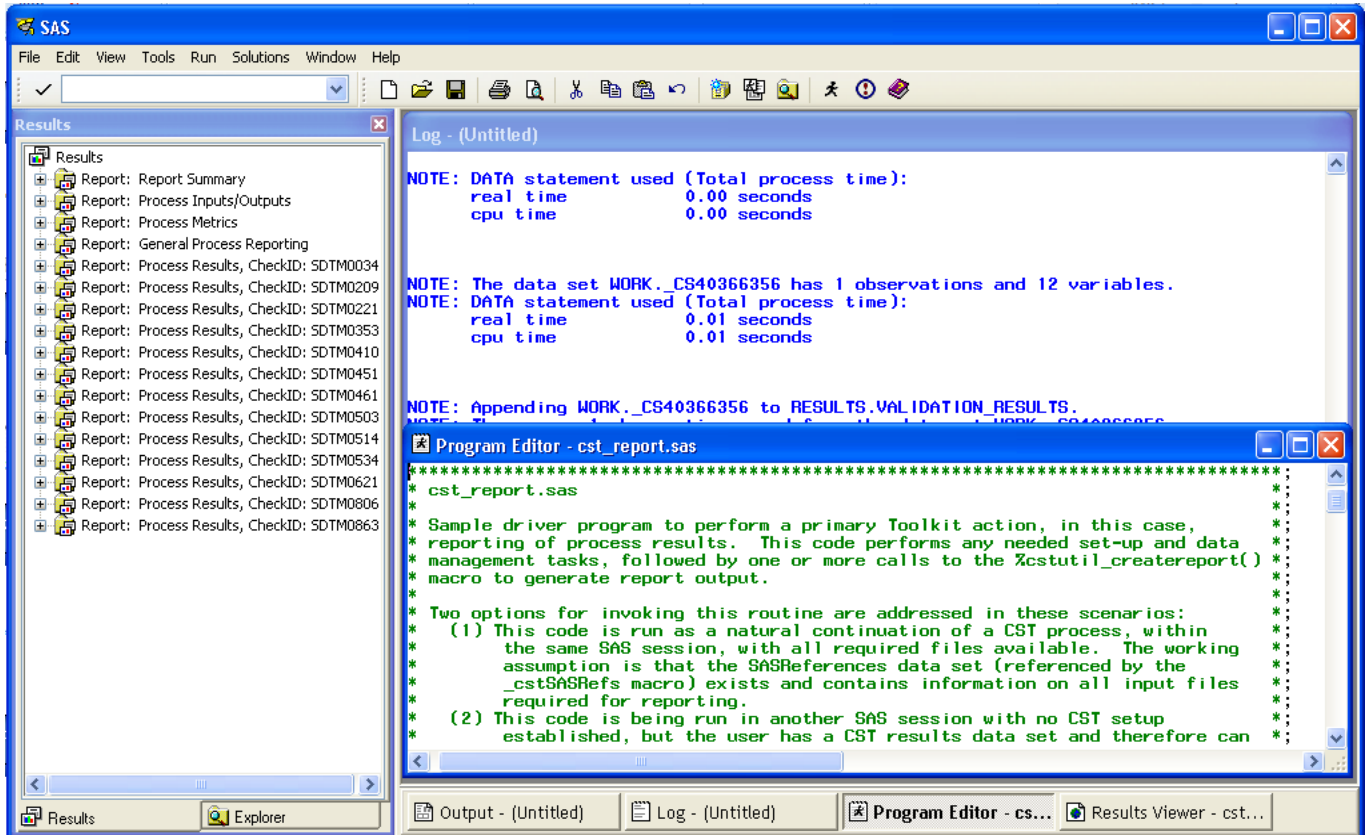
Two Report Macros now available

- Validation Report – reports on validation checks
  - Validation_results
  - Validation_metrics
  - Options available to the user
- Metadata Report – information about the checks
  - Validation_control
  - Messages
  - Validation_stdrefs
  - Options available to the user

Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM – Validation Report

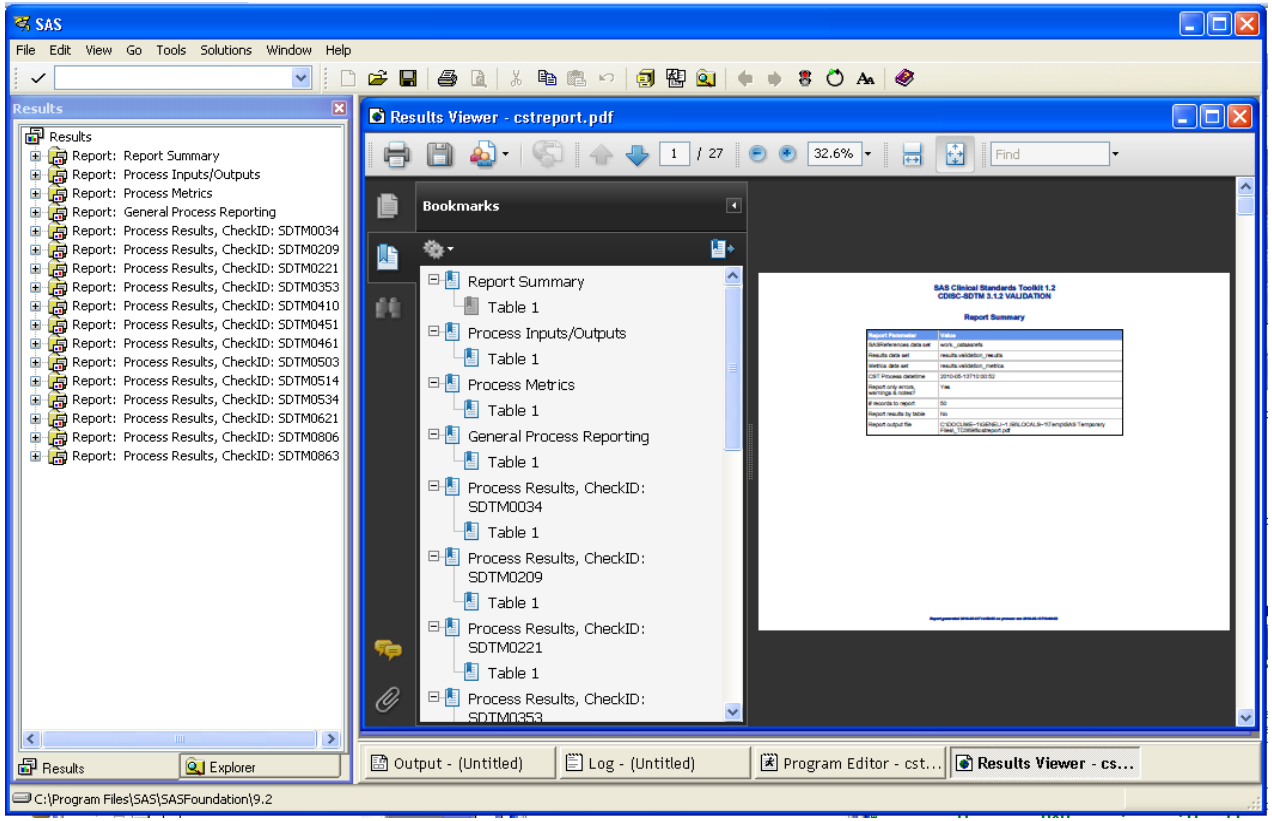
- Open a SAS session with toolkit installed
- Open cst_report.sas from sample study
- Submit



Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM – Validation Report

- Runs successfully
- No errors or warning detected in the log
- Reports display automatically in Results Viewer



Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

Validation Report

SAS Clinical Standards Toolkit 1.2
CDISC-SDTM 3.1.2 VALIDATION

Process Metrics

Summary Metrics		Check Metrics				
Metric	#	Check ID	# Check Invocations	# Recs (if available)	# Errors	# Check Invocations Not Run
# of distinct check invocations	13	SDTM0034	1	1	0	0
# check invocations not run	3	SDTM0209	1	6	2	0
Errors (severity=High) reported	1	SDTM0221	1	8848	8770	1
Warnings (severity=Medium) reported	0	SDTM0353	1	4	2	0
Notes (severity=Low) reported	0	SDTM0410	1	1	0	0
Structural errors, warnings and notes	1	SDTM0451	1	2	0	2
Content errors, warnings and notes	8837	SDTM0461	1	4	2	0
		SDTM0503	1	4	4	0
		SDTM0514	1	4	2	0
		SDTM0534	1	2	2	0
		SDTM0621	1	4	0	0
		SDTM0806	1	46	46	0
		SDTM0863	1	4	4	0

Note: "# Check Invocations Not Run" includes both checks that did not run and checks that failed to complete successfully.

Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

Validation Report

SAS Clinical Standards Toolkit 1.2
CDISC-SDTM 3.1.2 VALIDATION

Process Results, CheckID: SDTM0209

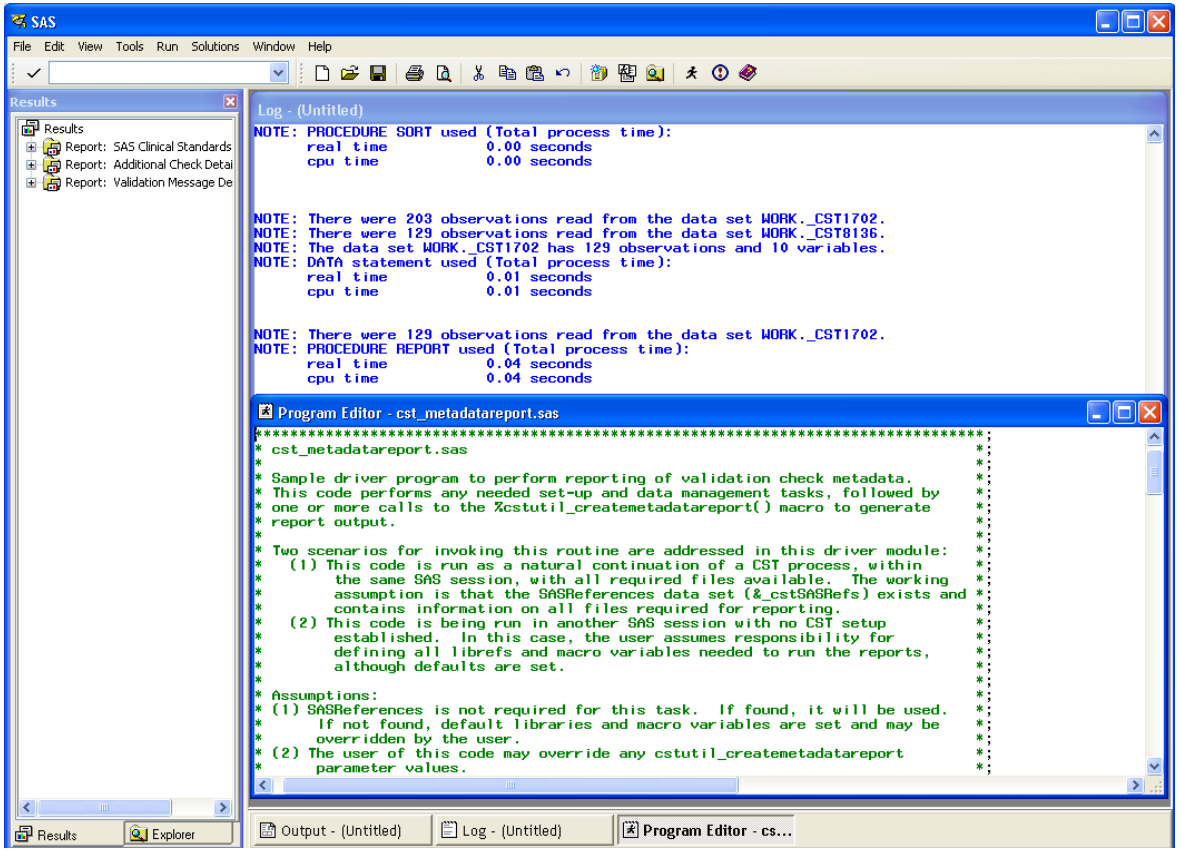
Description: Identifies records that violate the condition [Study Day of Start of Observation (STDY) less than or equal to Study Day of End of Observation (**ENDY)], limited to records where **STDY is not null and **ENDY is not null**
Check scope: (Tables) _ALL_-DS, (Columns) [STDY][**ENDY]**
Source: WebSDM (IR5100)
Validation check macro: cstcheck_columncompare, using source metadata

Check Invocation	Seq #	Source Data	Result Identifier	Message	Severity	Problem Detected?	Actual Value	Keys
1	1	SRCDATA.AE	SDTM0209	Begin day must be less than or equal to end day	High	Yes	AESTDY=12,AEENDY=11	STUDYID=SASCSTDEMODATA, USUBJID=S001P002, AEDECOD=ZQRA Syndrome, AESTDTC=2008-02-13T12:00
1	2	SRCDATA.AE	SDTM0209	Begin day must be less than or equal to end day	High	Yes	AESTDY=12,AEENDY=11	STUDYID=SASCSTDEMODATA, USUBJID=S001P002, AEDECOD=ZQRA Syndrome, AESTDTC=2008-02-13T12:00

Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM – Metadata Report

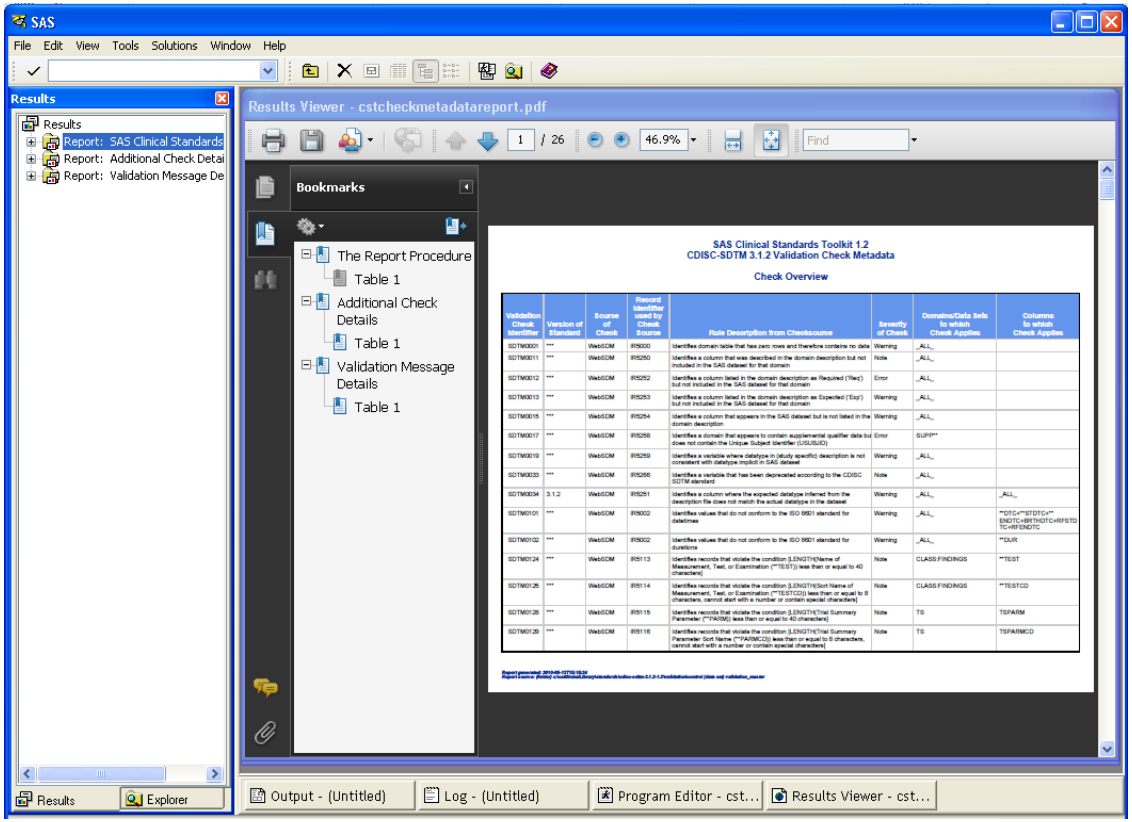
- Open a SAS session with toolkit installed
- Open cst_metadatareport.sas from sample
- Submit



Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM – Metadata Report

- Runs successfully
- No errors or warning detected in the log
- Reports display automatically in Results Viewer



Copyright © 2010, SAS Institute Inc. All rights reserved.

Validation of SDTM

Metadata Report

SAS Clinical Standards Toolkit 1.2
CDISC-SDTM 3.1.2 Validation Check Metadata

Additional Check Details

Validation Check Identifier	Source of Check	Type of Check	Code Source	Use Source Metadata	Code Logic	Lookup Standard Type	SAS Format Name	Check Status	Report All?
SDTM0001	WebSDM	Metadata	cstcheck_zeroobs	Yes	No codelogic for this check			Active	Yes
SDTM0011	WebSDM	Metadata	cstcheck_ metamismatch	Yes	proc sql noprint;create table _csttempds1 as select ref.table, ref.column from work._cstrefcolumnmetadata ref left join (select distinct table from work._cstsrccolumnmetadata) src on upcase(ref.table)=upcase(src.table) where upcase(ref.table)=upcase(src.table);create table _csttempds as select upcase(base.table) as table,upcase(base.column) as column from _csttempds1 base except select upcase(comp.table) as table,upcase(comp.column) as column from work._cstcolumnmetadata comp;select count(*) into: _cstMetricsCntNumRecs from _csttempds1;quit;			Active	Yes
SDTM0012	WebSDM	Metadata	cstcheck_ metamismatch	Yes	proc sql noprint;create table _csttempds1 as select ref.table, ref.column from work._cstrefcolumnmetadata (where=(upcase(core)="REQ")) ref left join (select distinct table from work._cstsrccolumnmetadata) src on upcase(ref.table)=upcase(src.table) where upcase(ref.table)=upcase(src.table);create table _csttempds as select upcase(base.table) as table,upcase(base.column) as column from _csttempds1 base except select upcase(comp.table) as table,upcase(comp.column) as column from work._cstcolumnmetadata comp;select count(*) into: _cstMetricsCntNumRecs from _csttempds1;quit;			Active	Yes
SDTM0013	WebSDM	Metadata	cstcheck_ metamismatch	Yes	proc sql noprint;create table _csttempds1 as select ref.table, ref.column from work._cstrefcolumnmetadata (where=(upcase(core)="EXP")) ref left join (select distinct table from work._cstsrccolumnmetadata) src on upcase(ref.table)=upcase(src.table) where upcase(ref.table)=upcase(src.table);create table _csttempds as select upcase(base.table) as table,upcase(base.column) as column from _csttempds1 base except select upcase(comp.table) as table,upcase(comp.column) as column from work._cstcolumnmetadata comp;select count(*) into: _cstMetricsCntNumRecs from _csttempds1;quit;			Active	Yes
SDTM0015	WebSDM	Metadata	cstcheck_	Yes	proc sql noprint;select count(*) into:_cstMetricsCntNumRecs from			Active	Yes

Copyright © 2010, SAS Institute Inc. All rights reserved.

Future

- 1.3
  - SDTM 3.1.2
  - Controlled Terminology 2010-03
  - Reporting Tools
- SDTM Validation Team
  - Hoping to consolidate SDTM 3.1.2 validation rules by November 2010
- Post 1.3 – Looking at the following
  - ODM 1.3.1
  - ADaM
  - Beyond CDISC – HL7
  - Reactive to users thru Forum, Product Management, Tech Support, Me
  - Our customers can influence any of this – be vocal

Copyright © 2010, SAS Institute Inc. All rights reserved.

CST information available from support.sas.com

http://support.sas.com/rnd/base/cdisc/cst/index.html

Availability of updates will generally be posted to the SAS Clinical Trials Forum, available at:

http://support.sas.com/forums/forum.jspa?forumID=9&start=0

SAS 9.1.3 Download Site for Clinical Standards Toolkit

http://ftp.sas.com/techsup/download/hotfix/12clintlkt.html

Copyright © 2010, SAS Institute Inc. All rights reserved.

Periodic product updates and preproduction downloads will be available at:

http://support.sas.com/rnd/base/cdisc/cst/preproduction.html

Currently available downloads include the following zip files:

CSTReporting92.zip (SAS 9.2 Reporting package)

CSTReporting913.zip (SAS 9.1.3 Reporting package)

CSTSDTM312_92.zip (SAS 9.2 CDISC-SDTM 3.1.2 metadata and domain templates)

CSTSDTM312_913.zip (SAS 9.1.3 CDISC-SDTM 3.1.2 metadata and domain templates)

CSTCT201003.zip (CDISC-Terminology-201003 format catalog and data set of CDISC controlled terminology)

Copyright © 2010, SAS Institute Inc. All rights reserved.

Questions?

Contact information:

Gene Lightfoot
00-1-919-531-3103

Gene.lightfoot@sas.com

(Development)

Bill J. Gibson
00-1-919-531-0996

billj.gibson@sas.com

(Product Management)

Copyright © 2007, SAS Institute Inc. All rights reserved.