

In-Depth Review of Validation Tools to Check Compliance of CDISC SDTM- Ready Clinical Datasets

Bhavin Busa
Kim Lindfield

May 25th, 2010
PharmaSUG Annual Meeting

Presentation Outline

- FDA and CDISC
- SDTM data flow from Sponsor to FDA
- Cubist needs and requirements for a SDTM validation tool
- Types of SDTM validation tools evaluated
 - Previously evaluated (2007/2008)
 - Recently evaluated (2009/2010)
- Side-by-side comparison of SDTM validation tools
- Conclusions

FDA and CDISC - Streamlining Clinical Trial Data

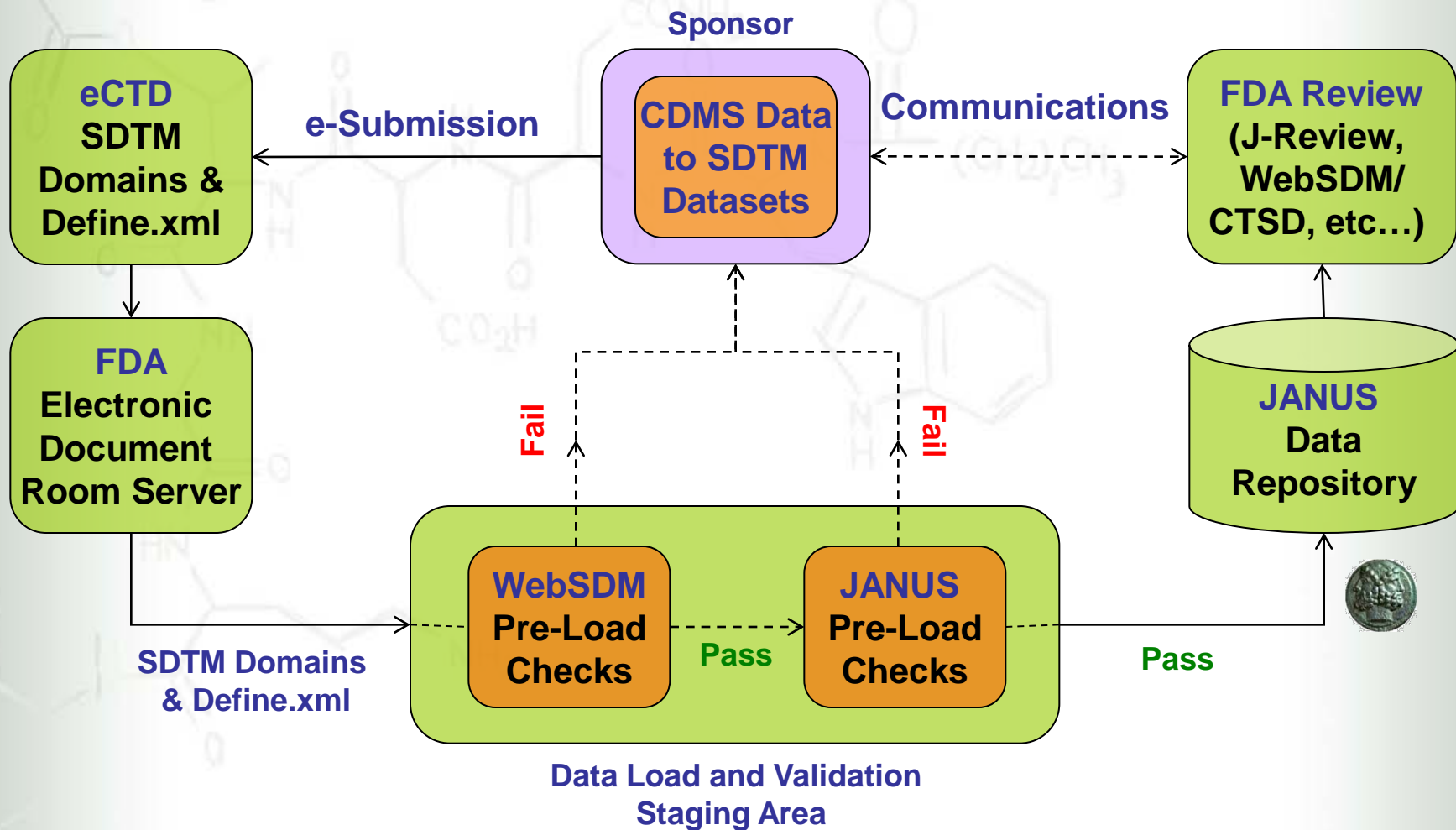
- FDA developed standards-based clinical data repository, Janus
- Critical Path Initiative: standardize format for clinical data submission
- eCTD Guidance*: FDA endorsing CDISC SDTM as the preferred model
- CDISC Team: active collaboration from FDA
- The Regulatory Plan: FDA proposed rule would require use of CDISC SDTM standards

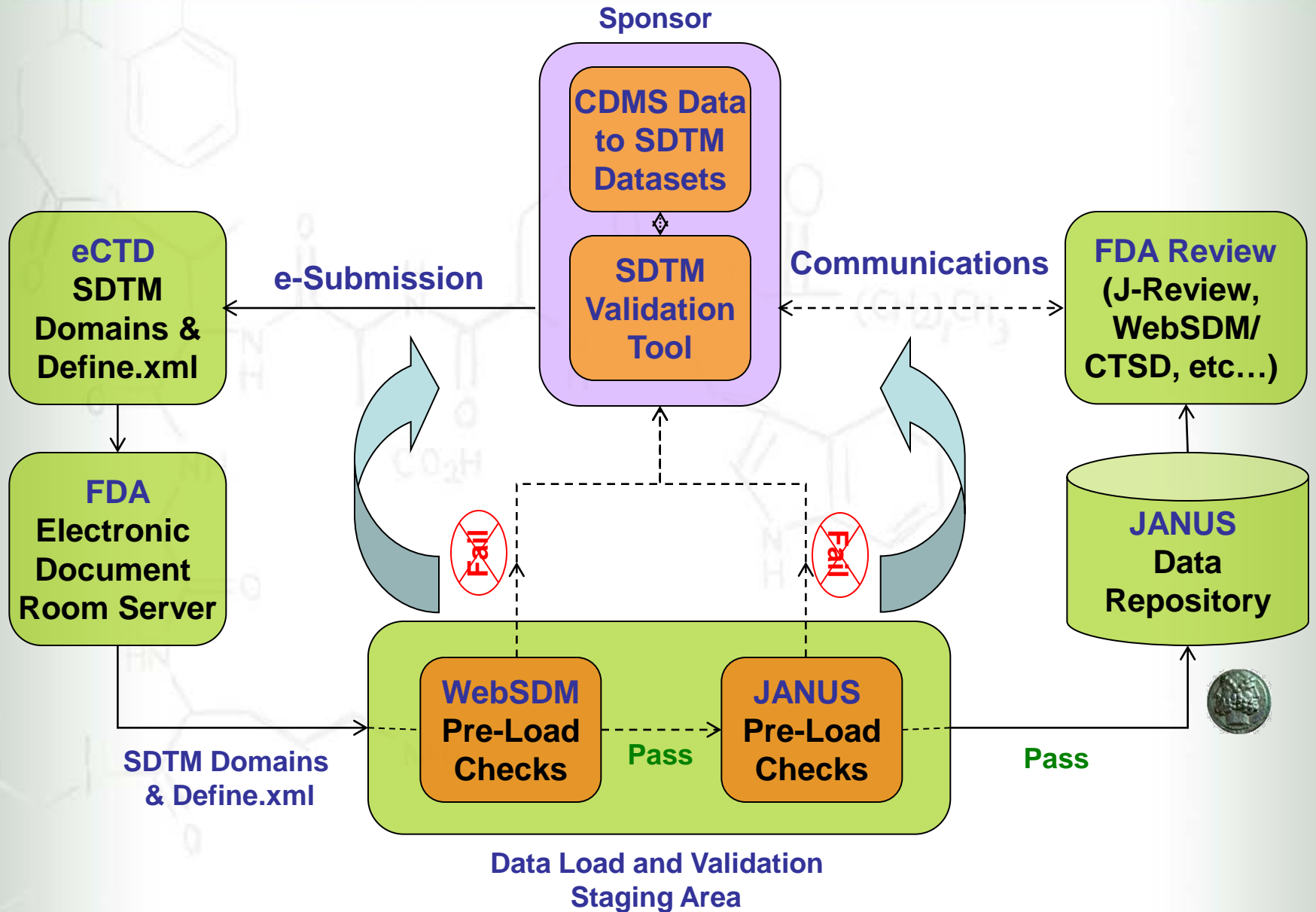
!! CDISC SDTM !!
!! CDISC SDTM !!

FDA and CDISC - Submission Data Warehouse

- Janus: Data model to collect and analyze clinical trial data
- Central access to standardize data
- Provides ability to perform cross-trial analysis for data mining
 - detects clinical trends
 - performs more advanced, robust analysis
- Creates an integrated data platform for most commercial tools for review, analysis and reporting

SDTM Data Flow From Sponsor to FDA - Pilot





Sanofi-Aventis Experience Submitting SDTM & Janus Compliant Datasets*

- “The most valuable information we learned from our experience submitting SDTM data to the FDA was that compliance with the SDTM IG is not sufficient to ensure the data will be successfully loaded by or useful to the FDA tools within Janus. For all planned submissions, the industry needs to address both SDTM IG rules and Janus rules.”

SDTM Validation Tools - Needs and Requirements

- Needs of a validation tool are:
 - Check the compliance of the SDTM domains for successful load into Janus
 - Reduce risk of delays in the submission review process
- Sponsor requirements of a validation tool are:
 - Check the compliance of the SDTM domains as per the most recent SDTM Implementation Guide (SDTM IG)
 - Apply SDTM IG and Janus pre-load checks
 - Identify structural and consistency errors, controlled terminology
 - Validate define.xml (or may be generate as well)
 - Versatile for use throughout the Clinical Data Life Cycle
 - Flexible enough to handle variations across release and as new standards are developed and become available

SDTM Validation Tools - Previously Evaluated Options

- Cubist evaluated SDTM validation tools (2007/2008):
 - **Web Submission Data Manager (WebSDM™)**
Lincoln Technologies – checks the compliance of submission-ready files according to the SDTM IG (V 3.1.1 and V 3.1.2)
 - **PROC CDISC**
SAS - procedure that provides checks of data content against the domain definitions outlined in the SDTM IG (V 3.1)
 - **SAS Based Macro Solution**
In-house - includes library of SAS macros that checks each SDTM domain for compliance with the SDTM IG (V 3.1.1 and V 3.1.2)

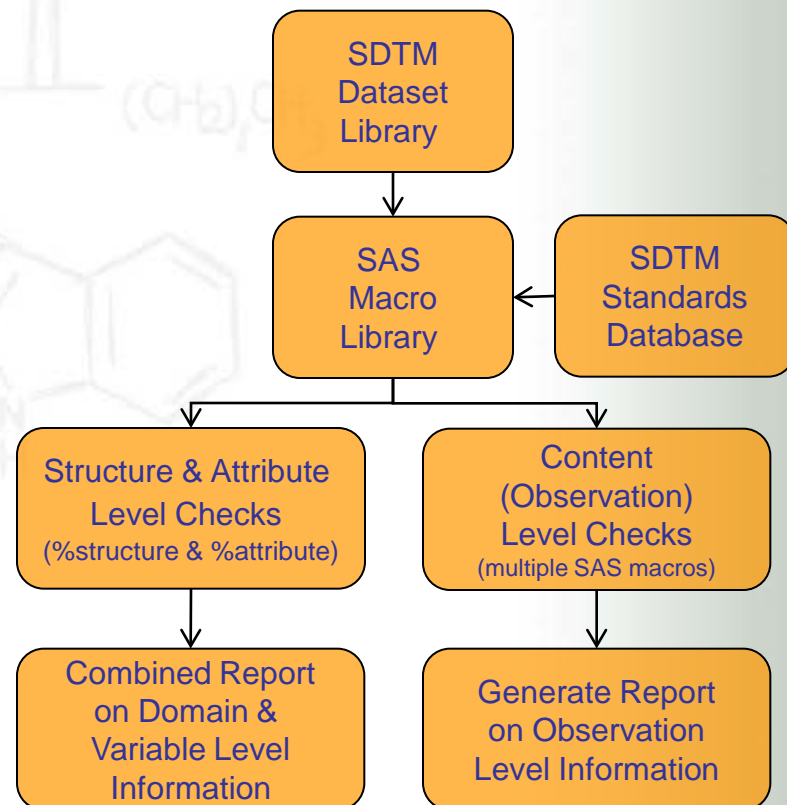
SDTM Validation Tools - Newly Evaluated Options

- Cubist evaluated SDTM validation tools (2009/2010):
 - SAS Clinical Standards Toolkit (CST)
SAS – SAS macro based solution to check the compliance of the clinical data and metadata against a reference standard (e.g. SDTM)
 - OpenCDISC Validator
OpenCDISC – open source Java based project that provides validation of clinical data against CDISC models (e.g. SDTM)

Previously Evaluated SDTM Validation Tools

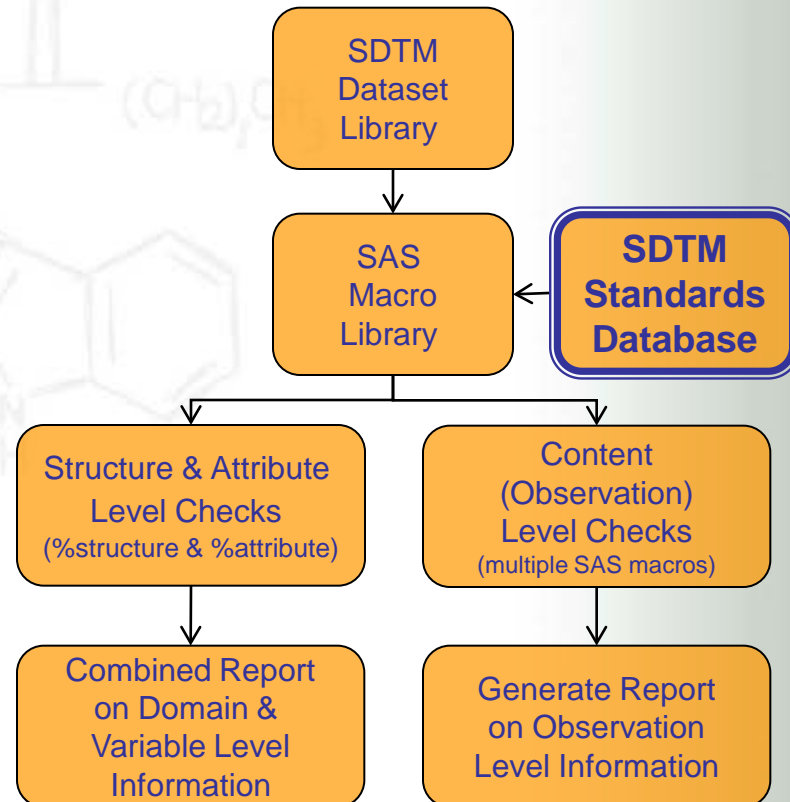
SDTM Validation Tools - SAS Macro Based Solution

- **SDTM Standards Database:** holds the metadata of all SDTM domains
- **SAS Macro Library:** collection of SAS macros to check the submission compliance of the **SDTM Dataset Library**
- Includes **structural, attribute and content** level checks
- Generate customizable **reports** for domain, variable and observation level information



SDTM Validation Tools - SAS Macro Based Solution

- Repository to maintain & access data standards
- Database: Filemaker® Pro, Microsoft® Excel/Access, etc..
- Holds the metadata (domain, variable and codelist/controlled terminology) level information
- Based on SDTM IG 3.1.1 but can be upgraded to SDTM IG 3.1.2 and also be customized for 'near-SDTM' or sponsor defined datasets



SDTM Validation Tools - SAS Macro Based Solution

- SDTM Standards Database (FileMaker Pro) tables

Table 1: The domain level database layout

Column Name	Description
Domain	The name of the domain (SDTM/Custom-Defined)
Description	A description of the domain
Structure	The data structure of the domain observations
KeyFields	The fields that uniquely identify the domain

Table 2: The variable level database layout

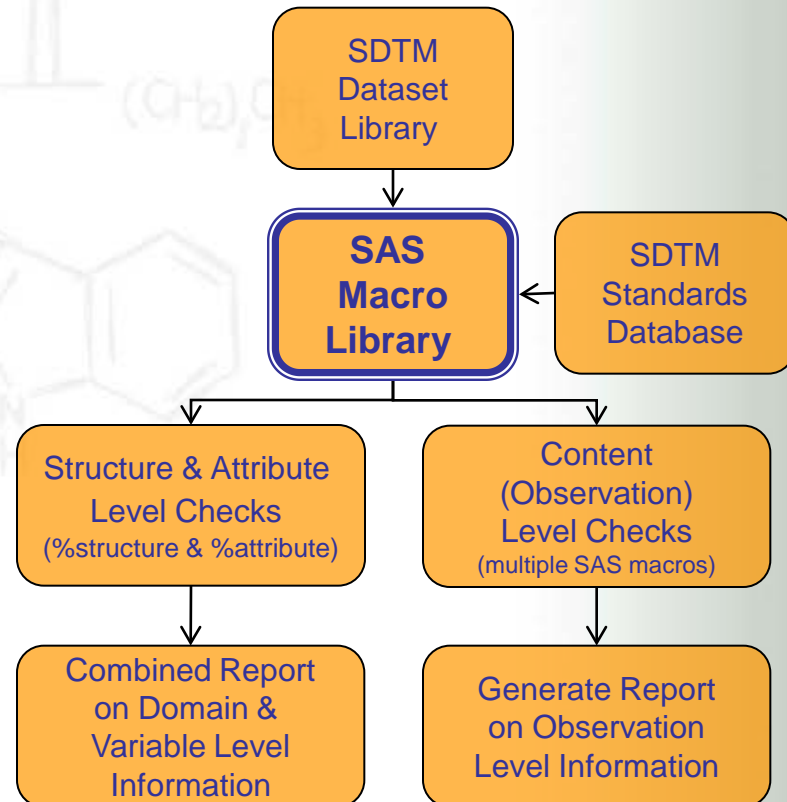
Column Name	Description
Domain	The name of the domain (SDTM/Custom-Defined)
VarPrefix	The variable name minus domain prefix
VarName	Variable name
VarType	Variable type
VarLabel	Variable label
VarLength	Variable length
VarFormat	Variable format or controlled terms, study-specific
CodeValue	Variable is coded, Expected or possible
DecodedValue	Decoded value of the variable in the domain (e.g.: 'Identifier', 'Timing', etc)
Role	'Timing', etc)
SortOrder	The order of the variables in the domain

Table 3: The format/controlled terminology level information

Column Name	Description
VarFormat	The variable format or controlled terms, study-specific
CodeValue	The coded value of the format/controlled terms
DecodedValue	The decoded value of the format/controlled terms

SDTM Validation Tools - SAS Macro Based Solution

- Consists of multiple validated SAS macros
- The macros are categorized into structure, attribute and content level checks
- Uses SDTM dataset library consisting of SAS or XPT files
- Access SDTM standards database to get the metadata information
- Requires Base SAS license only



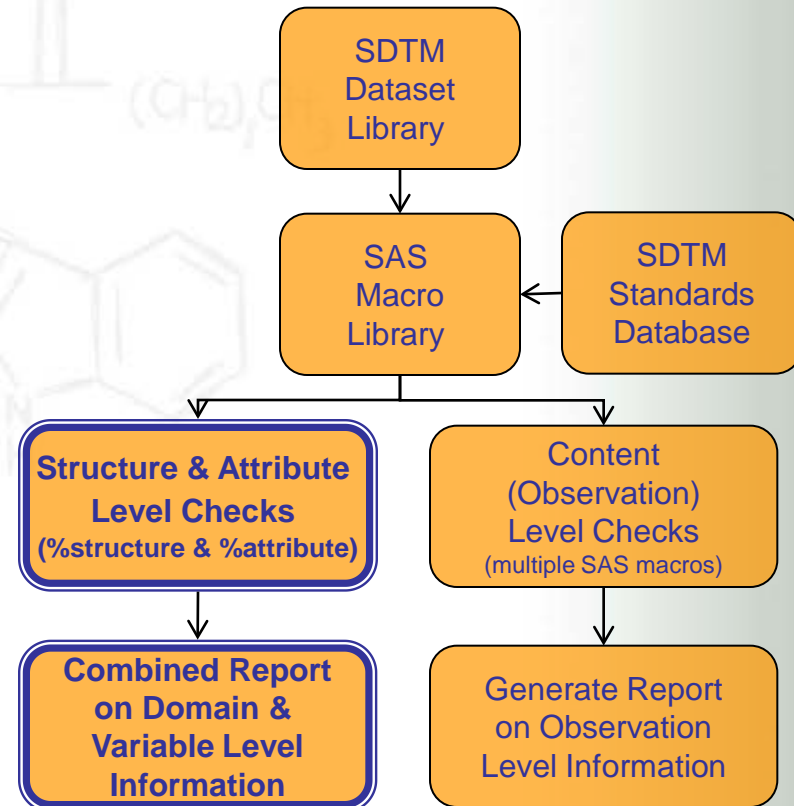
WebSDM, JANUS and Sponsor Defined Checks

Rule ID	Domain	Description of Rule	Severity-WebSDM	Severity-Janus	Severity-Cubist	Message
IR4000	All	Identifies domain table that has zero rows and therefore contains no data.	Medium	Low	Medium	No rows in domain table
IR4001	All	Identifies a null (empty) value found in a column where (Standard) Core attribute is 'Req'	High	Medium	High	Null value in column
IR4002	All	Identifies records where the value for a date doesn't conform to the ISO 8601 standard.	Medium	High	High	Invalid ISO 8601 value
SD1001	All	Identifies domain table that has domain description (label) that does not match standards.	N/A	N/A	SD-Low	Domain description (label) does not match standards

- SD: Sponsor Defined, unless otherwise stated, the severity level is derived based on the WebSDM and Janus validation checks

SDTM Validation Tools - SAS Macro Based Solution

- Structural level macro (%structure & %attribute):
 - Uses PROC CONTENTS to get domain and variable level information from the SDTM dataset library
 - Access domain and variable level metadata table from the SDTM standards database
 - The checks * include:
 - domain name
 - domain description
 - variable name
 - variable order
 - variable label
 - variable type
 - variable length



Example of the SDTM Dataset Error Log Reports - Structure & Attribute Level Information

Microsoft Excel - Study ABC - SDTM Datasets Structural-Attributes Level Error Log

	A	B	C	D	E	F
1	Structural-Domain-Level Error Log for Study ABC					
2	SDTM Domain Name	SDTM Domain Description	Standards Domain Description	Error Message	Severity Level	
3	DM	Demo Dataset	Demographics	Domain description (label) does not match standards	SD-Low	
4	DM	Demo Dataset	Demographics	Variable order in the domain does not match standards	SD-Low	
5	MG	Micro Genetics		Domain is not expected in the study	SD-High	
6	VS		Vital Signs	Domain description (label) does not match standards	SD-Low	
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						

Structural-Domain-Level | Structural-Variable-Level | Attributes-Variable-Level

* SD: Sponsor Defined, unless otherwise stated, the severity level is derived based on the WebSDM and Janus validation checks

Example of the SDTM Dataset Error Log Reports - Structure & Attribute Level Information

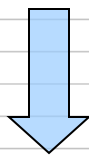
Microsoft Excel - Study ABC - SDTM Datasets Structural-Attributes Level Error Log

File Edit View Insert Format Tools Data Window Help Adobe PDF

Type a question for help

1	Structural-Variable-Level Error Log for Study ABC								
2	SDTM Domain Name	Variable Name	Variable Label	Variable Type	Variable Length	Core	Error Message	Severity Level	
3	DM	ARM	Description of Planned Arm	Char	36	Required	Required variable not present in the domain	High	
4	DM	SITEID	Study Site Identifier	Char	4	Required	Required variable not present in the domain	High	
5	DM	RDTXGP	Randomized Treatment Group Name	Char	20	N/A	Additional variable present in the domain	Medium	
6	DM	RACE	Race	Char	40	Expected	Expected variable not present in the domain	Low	
7	DM	ETHNIC	Ethnicity	Char	22	Permissible	Permissible variable not present in the domain	SD-Low	
8									
9									
10									
11									
12									
13									
14									
15									
16									
17									
18									
19									
20									
21									

Structural-Domain-Level | Structural-Variable-Level | Attributes-Variable-Level



* SD: Sponsor Defined, unless otherwise stated, the severity level is derived based on the WebSDM and Janus validation checks

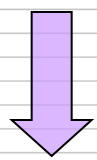
Example of the SDTM Dataset Error Log Reports - Structure & Attribute Level Information

Microsoft Excel - Study ABC - SDTM Datasets Structural-Attributes Level Error Log

File Edit View Insert Format Tools Data Window Help Adobe PDF Type a question for help

1	Attributes-Variable-Level Error Log for Study ABC										
2	Domain Name	Variable Name	SDTM Variable Label	Standards Variable Label	SDTM Variable Type	Standards Variable Type	SDTM Variable Length	Standards Variable Length	Core	Error Message	Severity Level
3	DM	AGE	Age	Age	Char	Num	36	36	Expected	Expected variable attribute information does not match	Medium
4	DM	SUBJID	Subject Identifier	Subject Identifier for the Study	Char	Char	4	4	Required	Required variable attribute information does not match	SD-Low
5	DM	SEX	Sex	Sex	Char	Char	1	6	Required	Required variable attribute information does not match	SD-Low
6	DM	INVNAM	Investigator ID	Investigator Name	Char	Char	20	20	Permissible	Permissible variable attribute information does not match	SD-Low
7											
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											
20											
21											
22											
23											
24											
25											
26											
27											
28											
29											
30											
31											
32											
33											

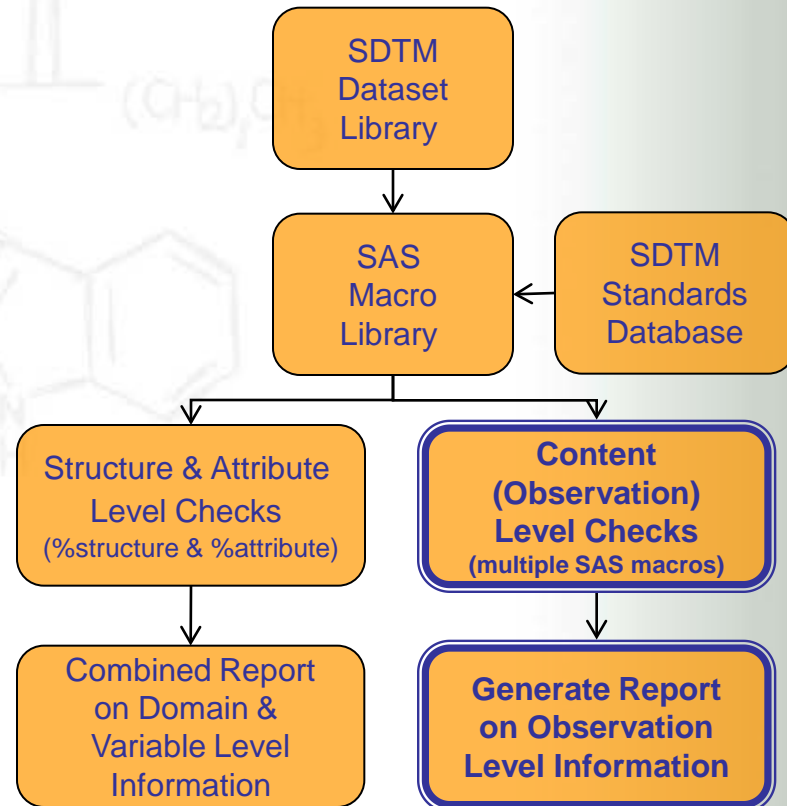
Structural-Domain-Level | Structural-Variable-Level | **Attributes-Variable-Level**



* SD: Sponsor Defined, unless otherwise stated, the severity level is derived based on the WebSDM and Janus validation checks

SDTM Validation Tools - SAS Macro Based Solution

- Content (Observation) level checks:
 - Consist of multiple SAS macros that are designed to check SDTM datasets against the pre-defined set of rules
 - Based on SDTM IG/Janus Operational Pilot SDTM validation specification
 - Access the format/controlled terminology level information from the SDTM standards database
 - Generate single report which is a combination of output from several content level macros in a library



SDTM Validation Tools - SAS Macro Based Solution

- List of some content level macros programmed to WebSDM/Janus pre-load checks:

Macro Name	Purpose (Check Rules)	Applicable Domain(s)	Applicable Variable(s)	Severity
<i>%agechk</i>	reports records with negative Age (AGE) value. Missing AGE is ignored.	DM	AGE	High
<i>%zerorows</i>	reports domain table that has zero observations and contains no data	ALL	N/A	Low
<i>%domain</i>	identifies records where the value in the Domain Abbreviation column (DOMAIN) does not match with the domain name	ALL	DOMAIN	Low
<i>%iso8601</i>	identifies records where the value for a date does not conform to the ISO8601 standard	ALL	ALL Date variables	High

Example of the SDTM Dataset Error Log Reports - Content Level Information

Microsoft Excel - ContentLevelCheck-SampleOutput

File Edit View Insert Format Tools Data Window Help Adobe PDF Type a question for help

	A	B	C	D	E
1	Content-Level Error Log for Study ABC				
2	SDTM Domain Name	Affected Variable(s)	Affected Unique Subject Identifier	Error Message	Severity
3	DM	AGE	0601-009-0082	Negative Age Value	High
4	DM	RFENDTC	0601-003-0069	RFENDTC cannot be null when ARMCD NE 'SCRNFL'	High
5	DM	RFSTDTC	0601-001-0074	RFSTDTC cannot be null when ARMCD NE 'SCRNFL'	High
6	DM	DMDTC	0601-009-0085	Invalid ISO 8601 value	High
7	DM	ARM/ARMCD	0601-004-0101	If ARMCD equals SCRNFL then ARM must equal 'Screen Failure'	Medium
8	DM	DOMAIN	0601-005-0077	Inconsistent value for DOMAIN	Low
9	VS	N/A	N/A	No rows in domain table	Low
10					
11					
12					
13					
14					
15					
16					
17					
18					

Content-Level

* SD: Sponsor Defined, unless otherwise stated, the severity level is derived based on the WebSDM and Janus validation checks

Example of the SDTM Dataset Error Log Reports - Content Level Information

Microsoft Excel - ContentLevelCheck-SampleOutput

File Edit View Insert Format Tools Data Window Help Adobe PDF

Type a question for help

	A	B	C	D	E
1	Content-Level Error Log for Study ABC				
2	SDTM Domain Name	Affected Variable(s)	Affected Unique Subject Identifier	Error Message	Severity
3	DM	AGE	0601-009-0082	Negative Age Value	High
4	DM	RFENDTC	0601-003-0069	RFENDTC cannot be null when ARMCD NE 'SCRNFL'	High
5	DM	RFSTDTC	0601-001-0074	RFSTDTC cannot be null when ARMCD NE 'SCRNFL'	High
6	DM	DMDTC	0601-009-0085	Invalid ISO 8601 value	High
7	DM	ARM/ARMCD	0601-004-0101	If ARMCD equals SCRNFL then ARM must equal 'Screen Failure'	Medium
8	DM	DOMAIN	0601-005-0077	Inconsistent value for DOMAIN	Low
9	VS	N/A	N/A	No rows in domain table	Low
10					
11					
12					
13					
14					
15					
16					
17					
18					

Content-Level

* SD: Sponsor Defined, unless otherwise stated, the severity level is derived based on the WebSDM and Janus validation checks

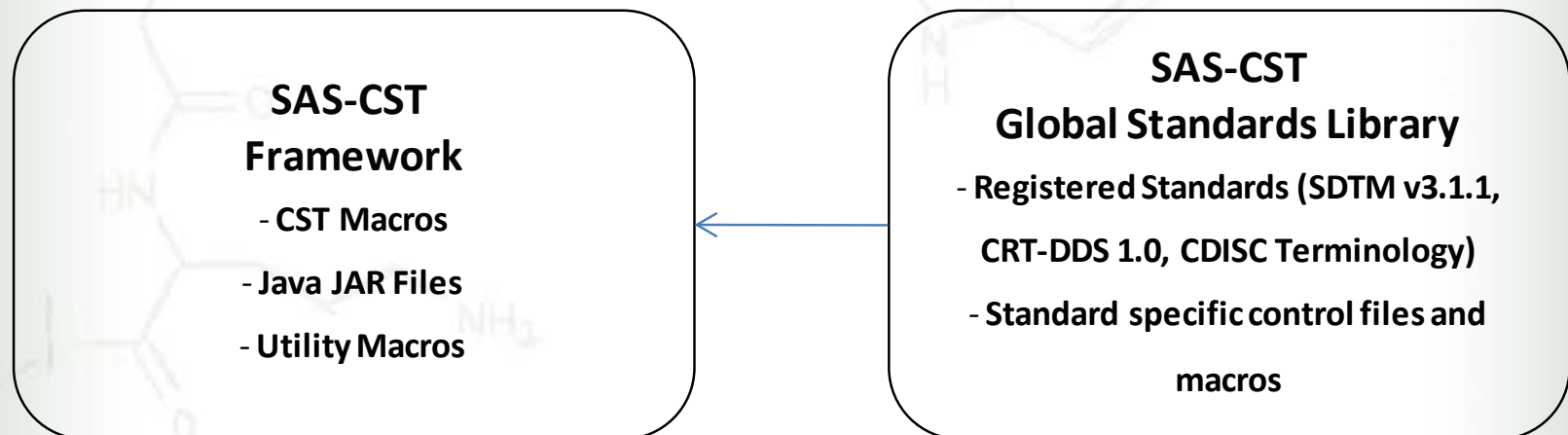
Newly Evaluated SDTM Validation Tools

SDTM Validation Tools - SAS Clinical Standards Toolkit (CST)

- **SAS CST:** SAS macro based solution to check the compliance of the clinical data and metadata against a reference standard
- Toolkit version 1.2 is licensed as a part of Base SAS (SAS 9.1.3 and SAS 9.2) and is available as a download at no additional cost
- Supports SDTM 3.1.1, CRT-DDS 1.0 and Controlled Terminology
 - Provides 143 unique validation checks which includes both WebSDM and Janus rules as well as SAS defined checks
 - Capable of generating and validating define.xml
 - Checks data against controlled terminology list (September 2008 version)

SDTM Validation Tools - SAS Clinical Standards Toolkit (CST)

- **Installation:** the toolkit is available as a hot fix for SAS 9.1.3 and 9.2, pre-requisite to have Apache Ant v1.7.1 and Java environment on the system
- **Framework:** SAS macro based framework consisting of 2 distinct pieces: **SAS-CST Framework** and **Global Standards Library**



SDTM Validation Tools - SAS Clinical Standards Toolkit (CST)

- Types of Checks Included:

Validation Type	# of checks
WebSDM	105
Janus	51
Janus FR	58
SAS	32
<i>Total</i>	246
<i>Total Unique Checks</i>	143

Validation Sub-Type	# of checks
Controlled Terminology	47
Column	64
Column Attribute	10
Column Value	25
Date	12
Metadata	26
Multi-record	24
Multi-table	38

SDTM Validation Tools - SAS Clinical Standards Toolkit (CST)

Control & Reference Files - CST

Reference
Metadata
(incl lookups)

Validation
Control

Execution
Properties

Controlled
Terminology

Input

Source Data
DM, AE,
LB,.....DS

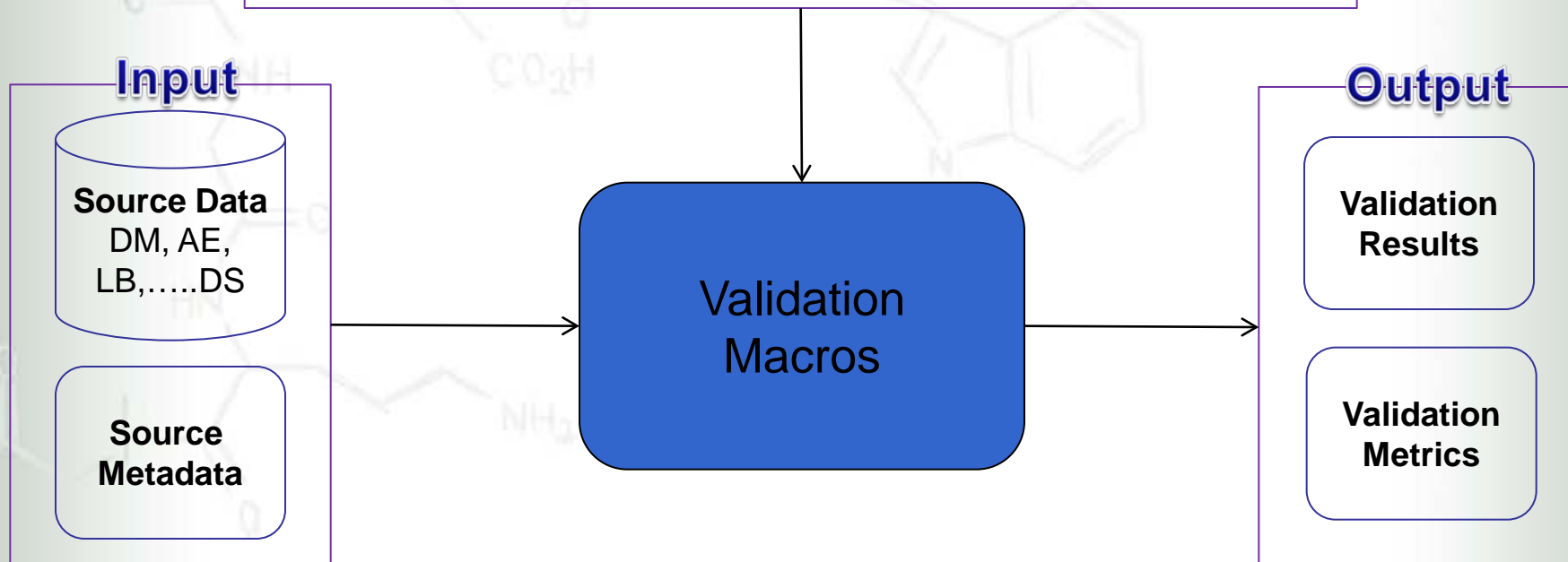
Source
Metadata

Validation
Macros

Output

Validation
Results

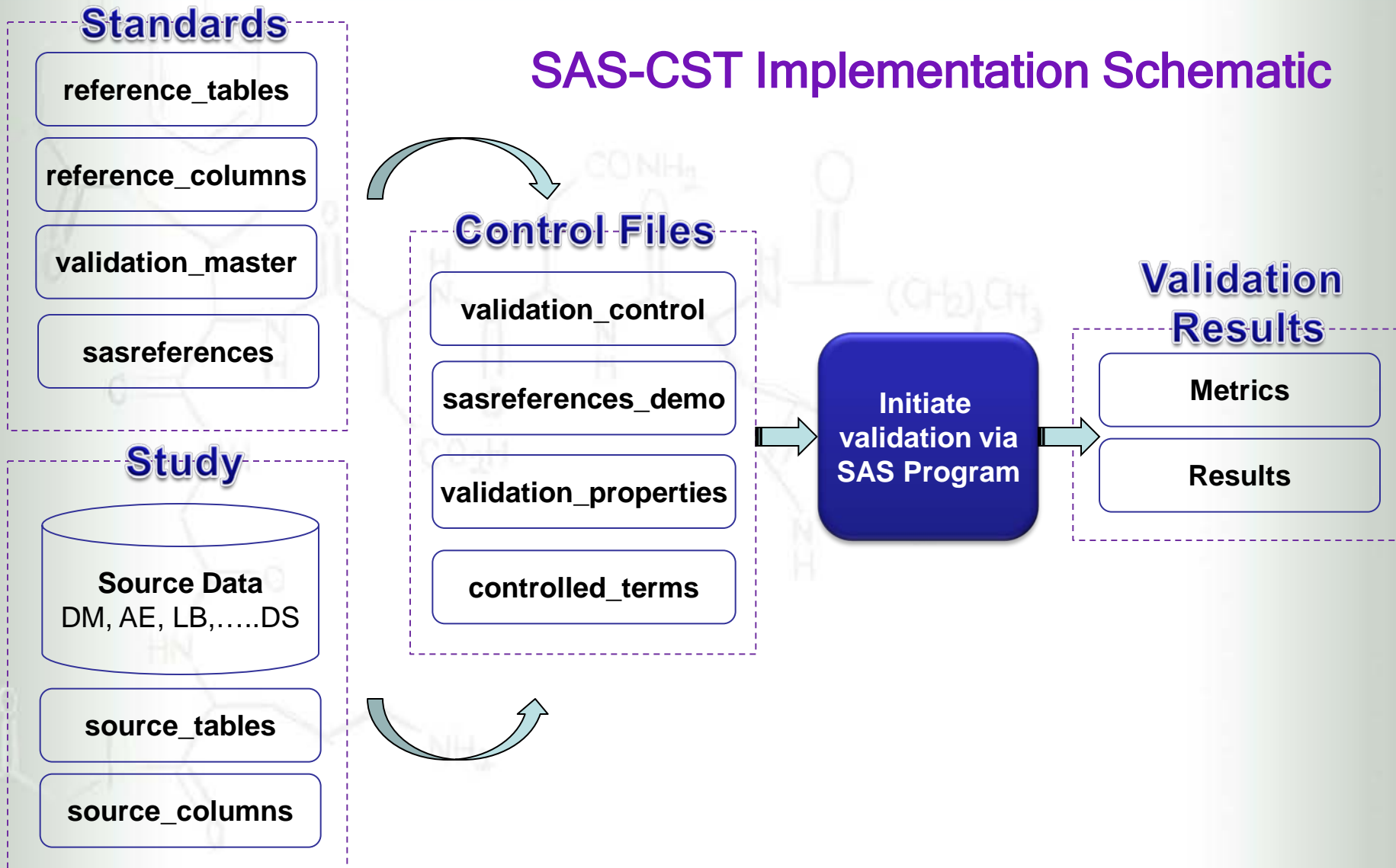
Validation
Metrics



SDTM Validation Tools - SAS Clinical Standards Toolkit (CST)

- Input from the user:
 - Create source metadata dataset (source_tables and source_columns datasets, PROC CONTENTS)
 - Create a list of validation checks to run by populating validation_control dataset (using validation master dataset provided as part of a global standard)
 - Create and populate a reference file called SASREFERENCES
 - Contents of the SASREFERENCES dataset
 - Location of reference metadata
 - Location of study data
 - Location of study metadata (source_tables, source_columns)
 - Locations of formats, codelists, dictionaries
 - Location of validation_control dataset
 - Output destination datasets
 - Run initializing macros and sas codes to generate validation check results/metrics

SAS-CST Implementation Schematic



```
1 %let studyRootPath = C:\CSTDemo;

2 /* Set up the CST-FRAMEWORK standard. */
   %cst_setStandardProperties (_cstStandard=CST-FRAMEWORK,
                             _cstStandardVersion=1.2,
                             _cstSubType=initialize);

3 /* Set up the CDISC-SDTM standard. */
   %cst_setStandardProperties (_cstStandard=CDISC-SDTM,
                             _cstStandardVersion=3.1.1,
                             _cstSubType=initialize);

4 /* >>> Define where the sasreferencestable lives */
   %let _cstSASRefsLoc = &studyRootPath/control;

5 /* >>> Define the name of the sasreferencestable */
   %let _cstSASRefsName = sasreferences_demo;

6 /* Allocate filerefs and librefs */
   %cstutil_allocatesasreferences;

7 /* Run the validation checks. */
   %sdm_validate;
```

SDTM Validation Tools - SAS Clinical Standards Toolkit (CST)

- Review of SAS-CST (Pros):
 - Includes checks that cover all WebSDM and Janus checks.
 - User-defined checks are possible.
 - Available with Base SAS (v9.1.3 and v9.2) at no additional cost.
 - SAS has used modular approach for this tool, no revalidation when the new standards become available.
 - User can add their own standard module.
 - Supports controlled terminology and CRT-DDS standards.
 - Future expansion to CDISC ADaM and HL7.
 - Sits on top of Base SAS – which already provides a compliant environment from a quality assurance perspective.
 - SAS technical support available if needed.

SDTM Validation Tools - SAS Clinical Standards Toolkit (CST)

- Review of SAS-CST (Cons):
 - Installation of this tool for SAS 9.1.3 was a daunting task and required IT help in order to figure out all the steps towards successful installation.
 - Built for SAS programmers (strong programming background). The tool cannot be used by non-SAS programmers (Standards Developer, Data Manager, etc.).
 - Not intuitive – steep learning curve to figure out implementation on real submission study.
 - Staff will require training – which can contribute to additional cost.
 - As the newer versions of the CDISC SDTM standards become available, SAS Institute will need to release new modules to register to the framework. Users will have to work around SAS release timeline.
 - Although a quick start document for SDTM validation is available, clear examples are not provided.

SDTM Validation Tools - OpenCDISC Validator

- **OpenCDISC Validator:** an open source Java based project that provides validation of datasets against CDISC models (e.g. SDTM)
- The tool is available to download for free. Requires no installation of the program – can run off a USB flash drive!
- Supports SDTM 3.1.1 & 3.1.2, CRT DDS 1.0, Controlled Terminology
 - Provides both SDTM IG and Janus rules as well as OpenCDISC defined checks
 - Capable of generating and validating define.xml
 - Checks data against controlled terminology list

SDTM Validation Tools - OpenCDISC Validator

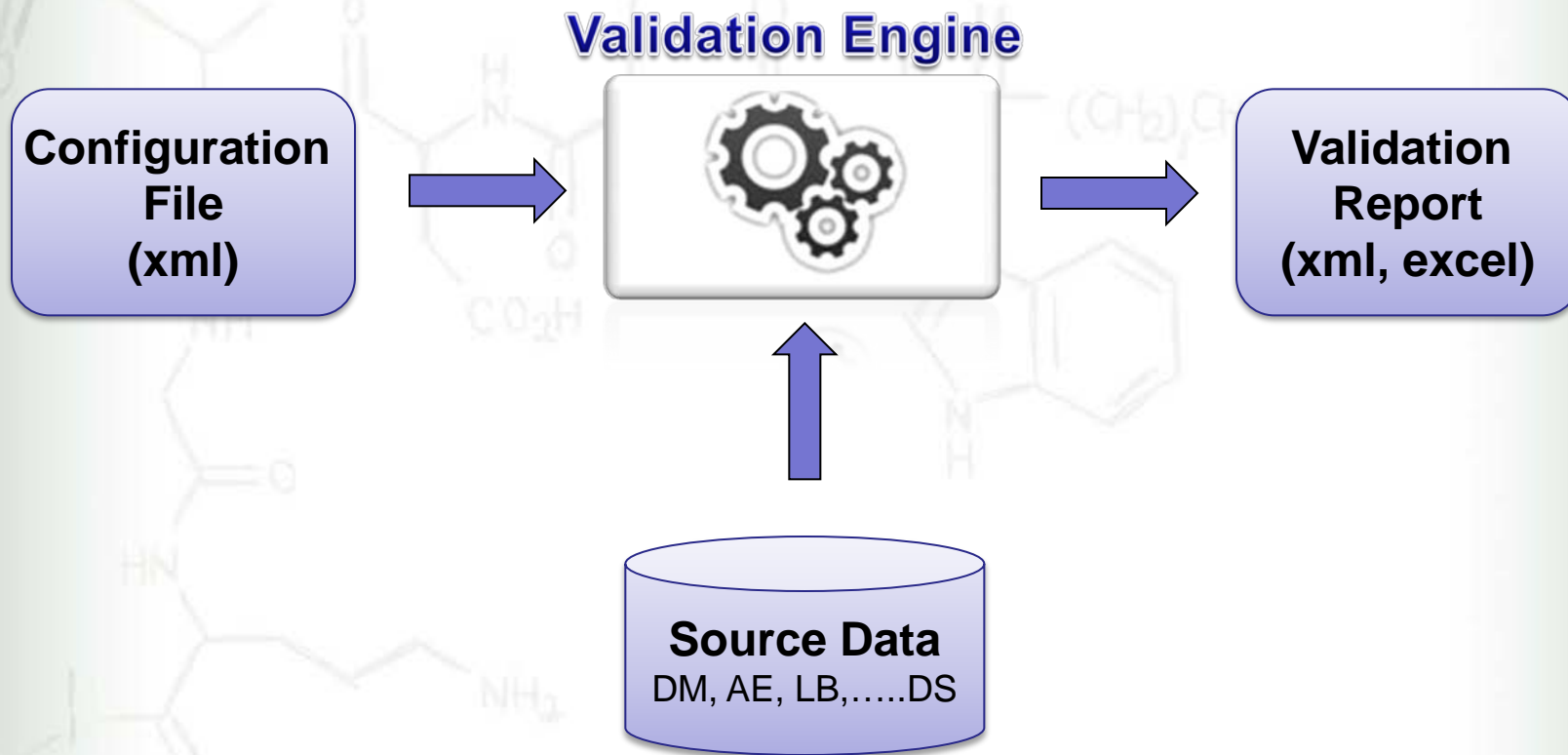
- **Installation:**

- OpenCDISC validator 1.0.1 production version is available for download as a zip file on OpenCDISC website.
- Requires Java runtime environment (JRE) version 1.5 or higher and a system with 2 GB RAM.
- To install the validator simply unzip the files in the directory where you want the validator to reside.

- **Framework:**

- The tool is an open source java based project which works on a XML-based validation framework.
- The tool architecture separates validation rules from application logic and comprises of two major components – the validation engine and configuration files (xml format)

SDTM Validation Tools - OpenCDISC Validator



SDTM Validation Tools - OpenCDISC Validator

The screenshot shows the OpenCDISC Validator application window. The title bar reads "OpenCDISC Validator" and includes standard window controls. The menu bar contains "File" and "Help".

The main interface is titled "What would you like to do?" and features two radio buttons: "Validate Data" (selected) and "Generate Define.xml".

Below this, there are two dropdown menus: "Standard:" set to "SDTM" and "Source Format:" set to "SAS® Transport (XPORT)".

The "Source Data:" section contains a large empty text area for listing sources. To its right are three buttons: "Browse", "Remove", and "Clear". Below this area is the text: "You can select multiple files or folders as sources".

The "Configuration:" section has a dropdown menu that is currently blank.

The "Define.xml:" section has a text input field and a "Browse" button. Below the input field is the text "Optional".

The "Report Format:" section has a dropdown menu set to "Excel" and a link labeled "Report Settings".

The "Progress:" section has a progress bar and the text "Waiting to begin.".

At the bottom right, there are two buttons: "Start" and "Stop".

SDTM Validation Tools - OpenCDISC Validator

- Review of OpenCDISC Validator (Pros):
 - Includes checks that cover all WebSDM and Janus checks.
 - Additional user-defined checks are possible.
 - Supports controlled terminology and CRT-DDS standards.
 - Available for windows platform at no additional cost and can be deployed on server or hosted environment.
 - Installation in minutes and can be run off a USB flash drive.
 - Flexible in upgrading to newly available standards.
 - Ready to use tool (intuitive) – no training required. The tool can be used by anyone (SAS programmers, Standards Developer, Data Manager, etc.).
 - Future expansion to CDISC ADaM and SEND.
 - Technical support available through OpenCDISC forum on their website.

SDTM Validation Tools - OpenCDISC Validator

- Review of OpenCDISC Validator (Cons):
 - No validation documentation provided with the installation. This will be available via OpenCDISC commercial entities.
 - Requires user to create their own configuration files if additional sponsor defined checks are to be included.
 - Since configuration and validation rules are both in xml format – it requires prior knowledge of the language to be able to edit or generate these files.
 - As the newer versions of the CDISC SDTM standards become available, OpenCDISC will need to make the configuration files available to the user. This means users will have to work around OpenCDISC timeline.

Side-By-Side Comparison of Validation Tools

No.	Items ⁺	In-house SAS Macro Soln	SAS-CST	OpenCDISC
1.	Validates domains as per SDTM IG			
	<i>Version 3.1.1</i>	✓	✓	✓
	<i>Version 3.1.2</i>	✓		✓
2.	Includes both WebSDM and JANUS checks	✓	✓	✓
3.	Allows for custom defined checks	✓	✓	✓
4.	Generate and validates Define.xml	✓*	✓	✓

⁺ Based on Cubist Defined Needs and Requirements. Will defer from Sponsor to Sponsor

Side-By-Side Comparison of Validation Tools

No.	Items ⁺	In-house SAS Macro Soln	SAS-CST	OpenCDISC
5.	Independence from external parties (with regards to updates)	✓		
6.	Provides timely updates	✓*		✓
7.	SAS Based Environment	✓	✓	
8.	21 CFR Compliant (validation documents IQ/OQ provided)	✓*	✓	
9.	Requires very minimal installation/implementation steps			✓

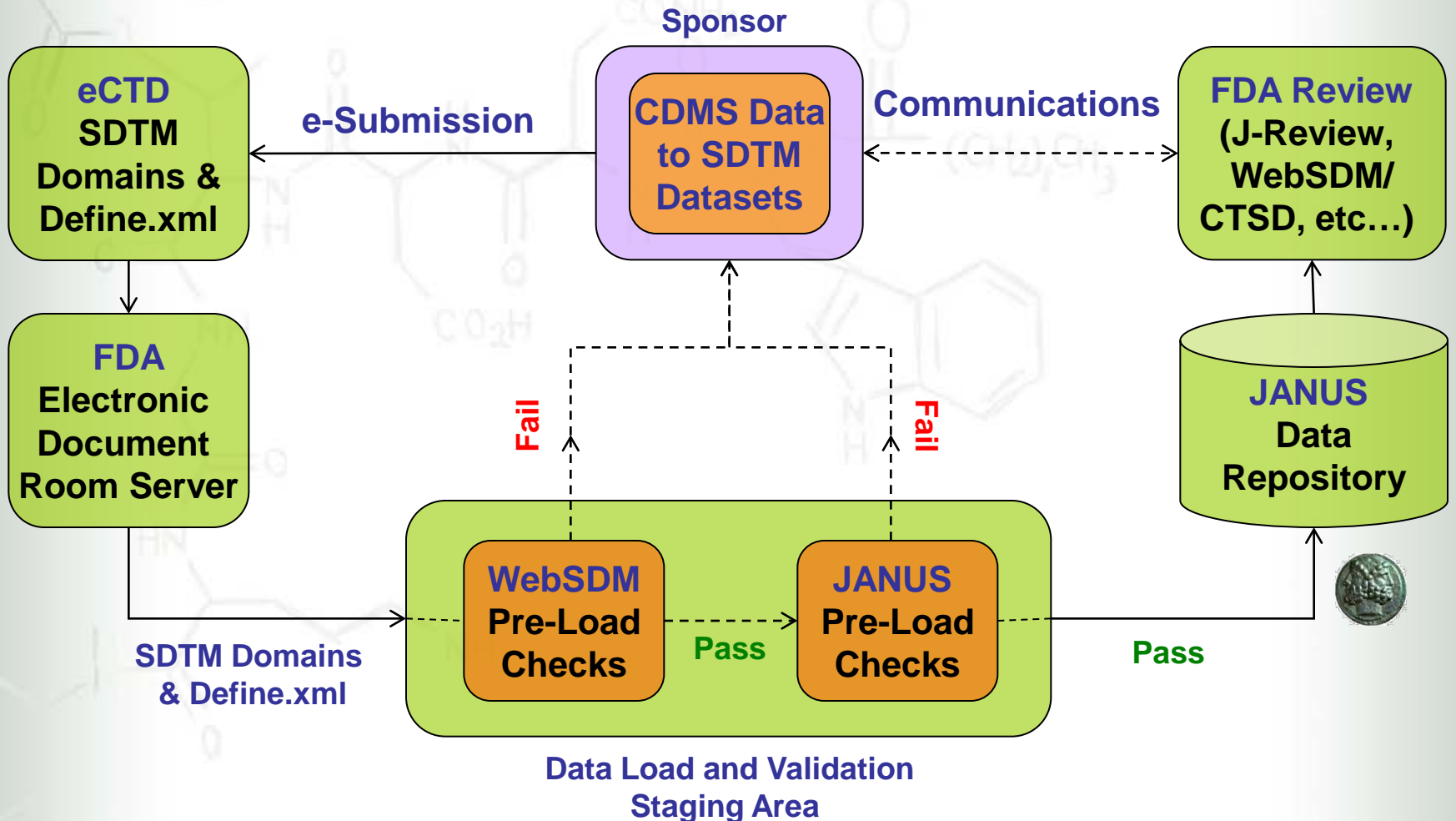
⁺ Based on Cubist Defined Needs and Requirements. Will defer from Sponsor to Sponsor

Side-By-Side Comparison of Validation Tools

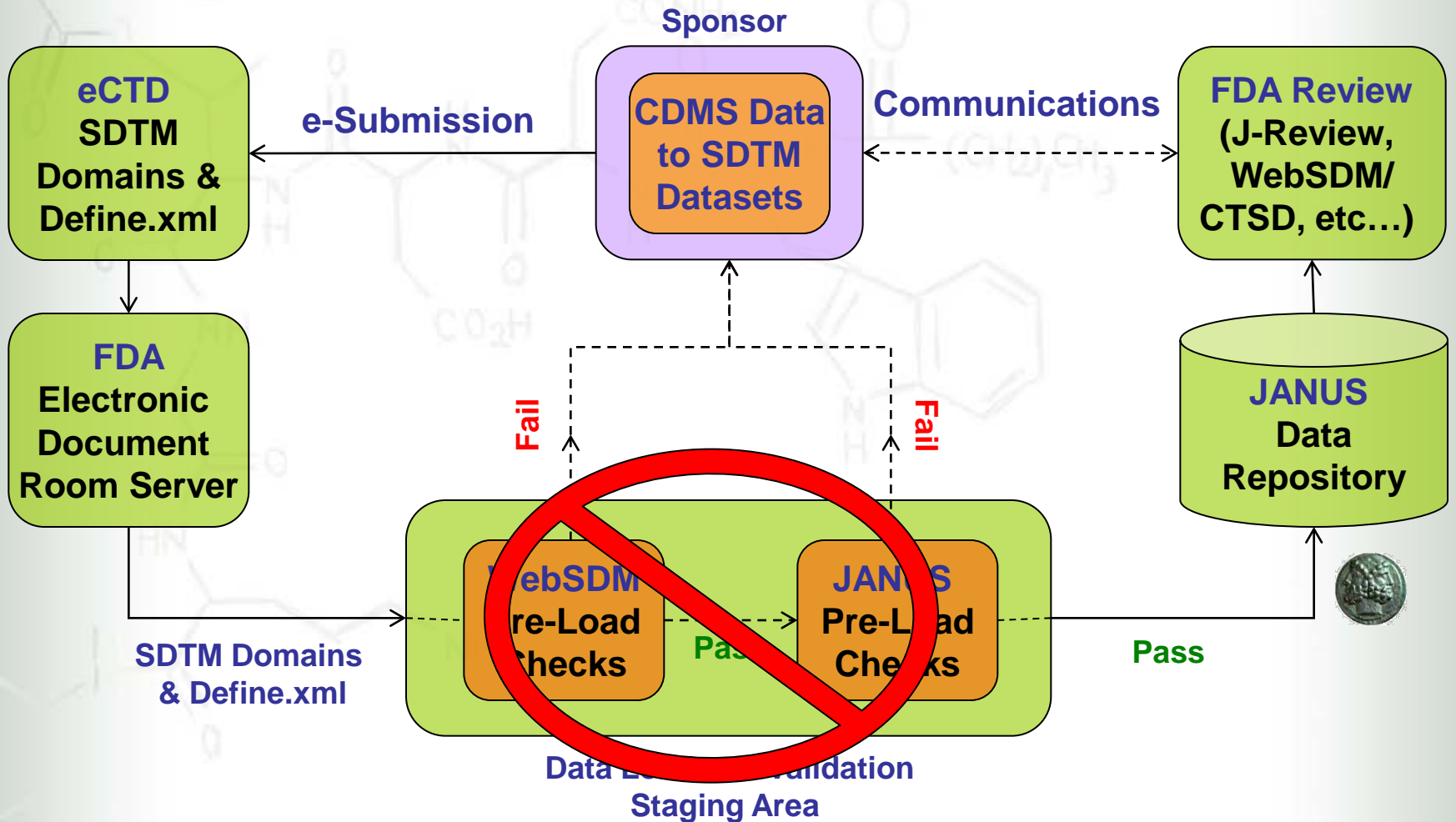
No.	Items ⁺	In-house SAS Macro Soln	SAS-CST	OpenCDISC
10.	Easy to use/requires minimal training (can be used by Data Manager, Standards Coordinator)			✓
11.	Supported by existing clinical programming (SAS) skill sets (No additional programming language to learn)	✓	✓	
12.	Generate reports directly into a user-friendly format	✓		✓

⁺ Based on Cubist Defined Needs and Requirements. Will defer from Sponsor to Sponsor

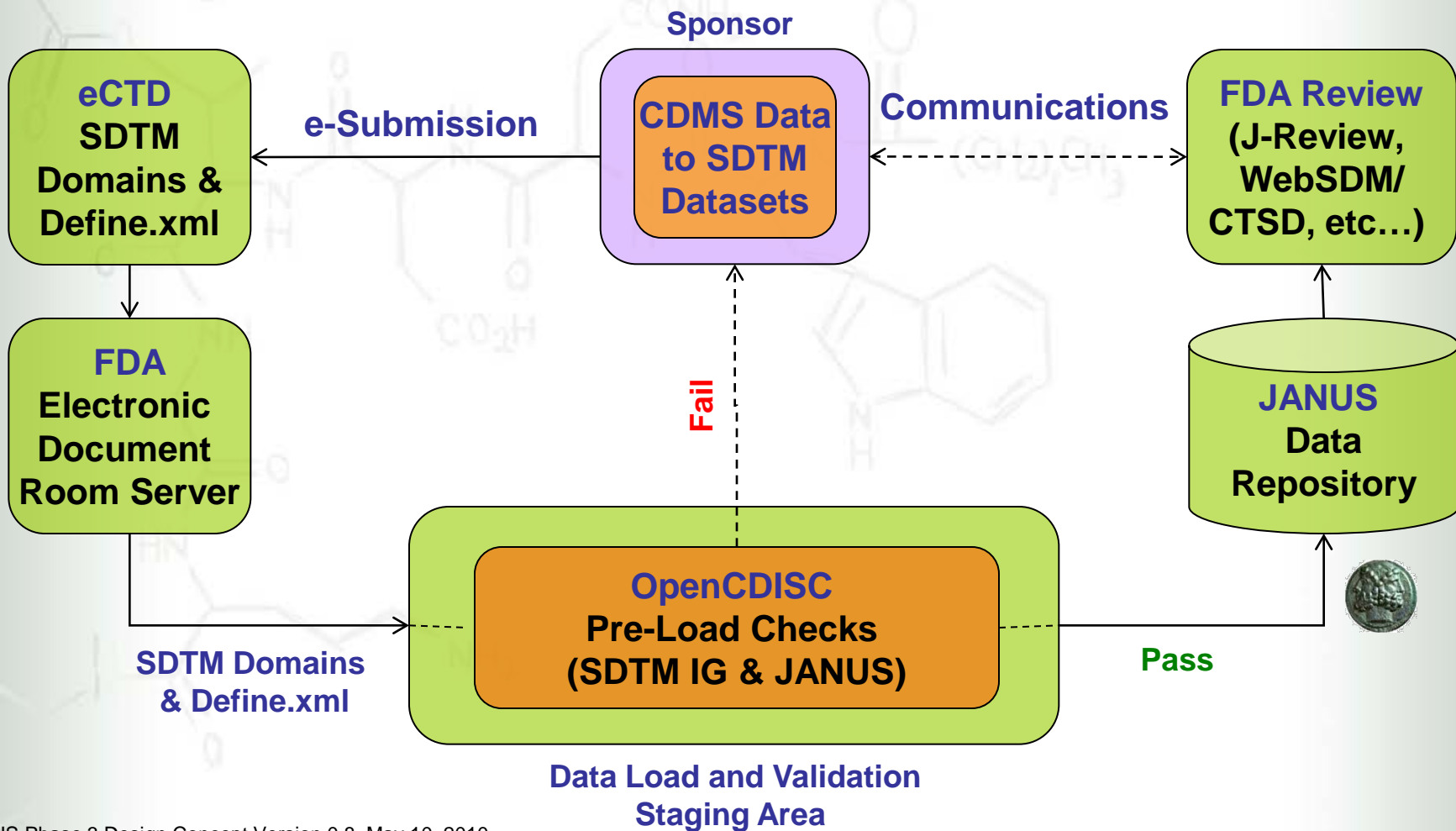
Latest Development at FDA - JANUS Phase 2 (CDER)



Latest Development at FDA - JANUS Phase 2 (CDER)



Latest Development at FDA - JANUS Phase 3 (CDER) *



* JANUS Phase 3 Design Concept Version 0.8, May 10, 2010

Latest Development at FDA - JANUS Phase 3 (CDER) *

- The major enhancements to the current Janus design are to
 - replace the proprietary WebSDM validation tool with the open source OpenCDISC Validator to validate the SDTM data files
 - replace the Janus 1.0 data model with the Janus 2.0 data model.
- Include Open Source Software (No proprietary software)
 - Public access to software implementing the validation tool will allow stakeholders to validate clinical trials data before submitting it to the FDA
 - the public to contribute to the development of the tool
 - stakeholders can see how the CDISC rules are implemented and can provide valuable input

* JANUS Phase 3 Design Concept Version 0.8, May 10, 2010

Side-By-Side Comparison of Validation Tools

No.	Items ⁺	In-house SAS Macro Soln	SAS-CST	OpenCDISC
13.	Is being used by FDA to check your submitted SDTM domains and Define.xml?			✓

Summary	In-house SAS Macro Soln	SAS-CST	OpenCDISC
Total number of check marks	11/14	7/14	11/14

What's Cubist Approach Going To Be Now?

- Continue to build/update our in-house SAS macro-based solution:
 - can include all the checks done at FDA (SDTM IG and Janus) and any others that gets added in future
 - customized checks and reports specific to user requirement
 - both for SDTM domains and for user-defined datasets (ADaM eventually)
 - can be utilized at early stages of SDTM conversion
- Use OpenCDISC in addition to our in-house SAS macro-based solution as one more pass of validation.

Conclusions

- FDA and Industry is quickly adopting CDISC SDTM standards
- Validation of SDTM submission is critical for successful load into Janus
- Tools to check the compliance of the SDTM domains as per the SDTM IG could reduce the risk of a delay in the drug submission process
- Validation tools evaluated:
 - WebSDM
 - PROC CDISC
 - In-House SAS Macro-Based Solution
 - SAS CST
 - OpenCDISC Validator

Acknowledgement

- Cubist Pharmaceuticals, Inc.
- SAS Institute
- Max Kanevsky (OpenCDISC)



Contact Information



The Presenter:

Bhavin Busa
Sr. Statistical Programmer
Cubist Pharmaceuticals, Inc.
Bhavin.Busa@cubist.com

Co-author:

Kim Lindfield

Sr. Manager Biostatistics

Kim.Lindfield@cubist.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies.