

Getting the most out of CDASH and SDTM

Automating the
specification and validation
process

Peter Van Reusel
Business Unit Director
CRO Services
Business & Decision Life Sciences

Tel +32 2 774 11 00
Fax +32 2 774 11 99
Mobile +32 476 54 59 17
peter.vanreusel@businessdecision.com

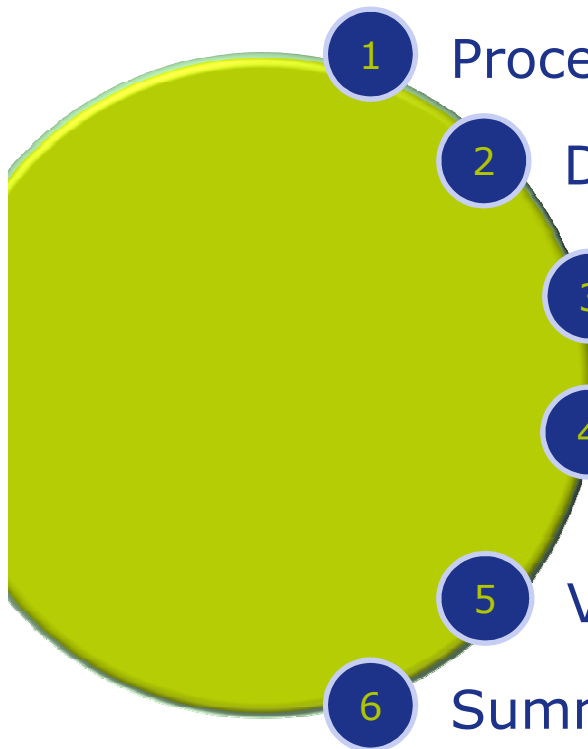
Sint-Lambertusstraat 141 Rue Saint-
Lambert
1200 Brussels

www.businessdecision-lifesciences.com



DRIVE YOUR PERFORMANCE

Agenda

- 
- 1 Process Overview
 - 2 Data Standards Library
 - 3 Study Metadata Repository
 - 4 Generating study metadata specifications and reports
 - 5 Verifying study data and metadata
 - 6 Summary & Conclusion

Agenda

1

Process Overview

2

Data Standards Library

3

Study Metadata Repository

4

Generating study metadata specifications and reports

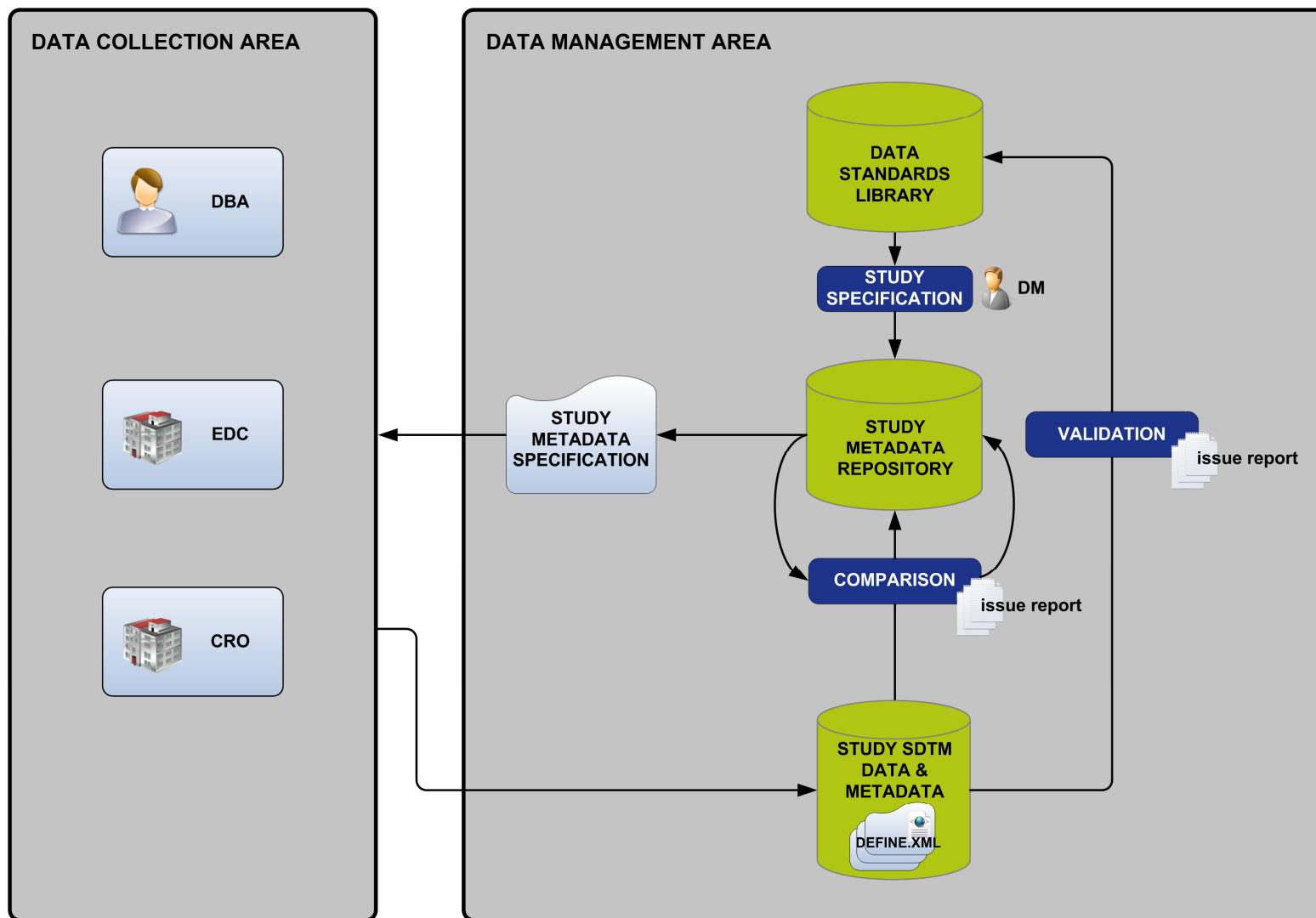
5

Verifying study data and metadata

6

Summary & Conclusion

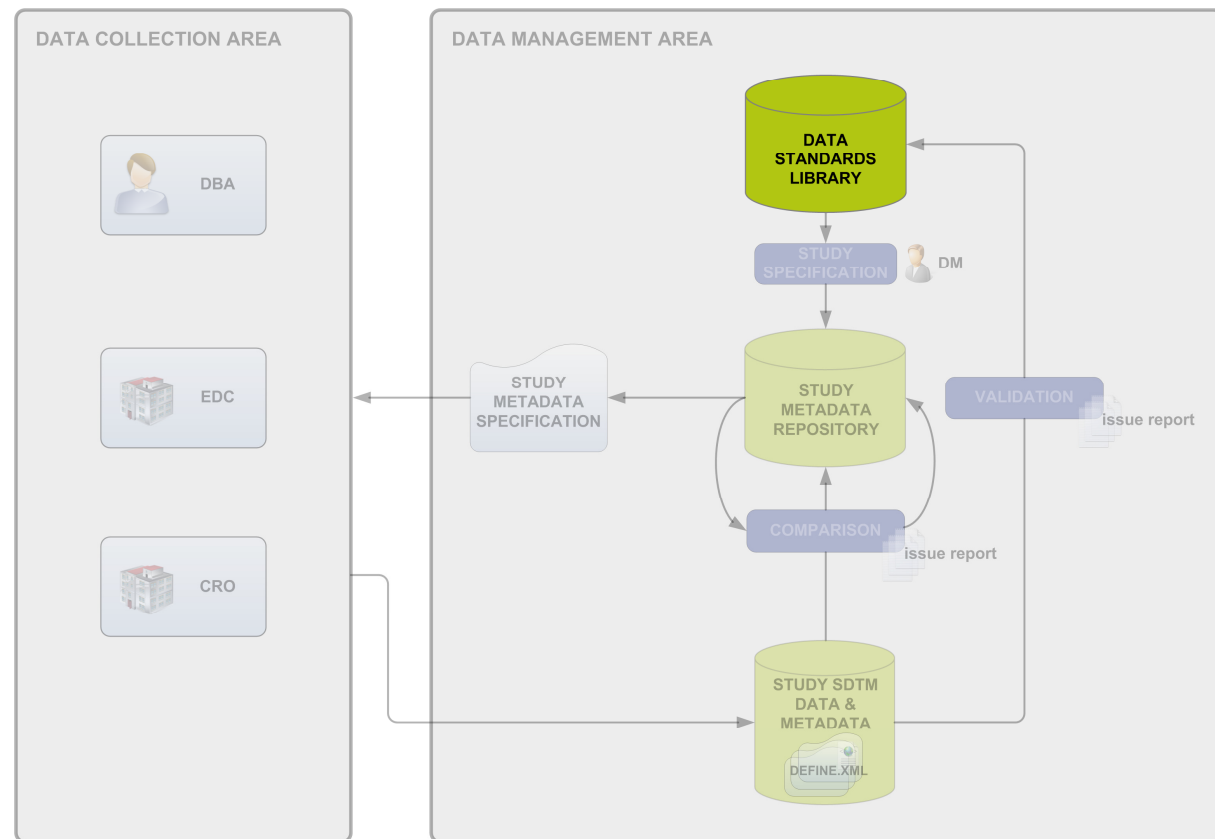
Process Overview



Process Overview

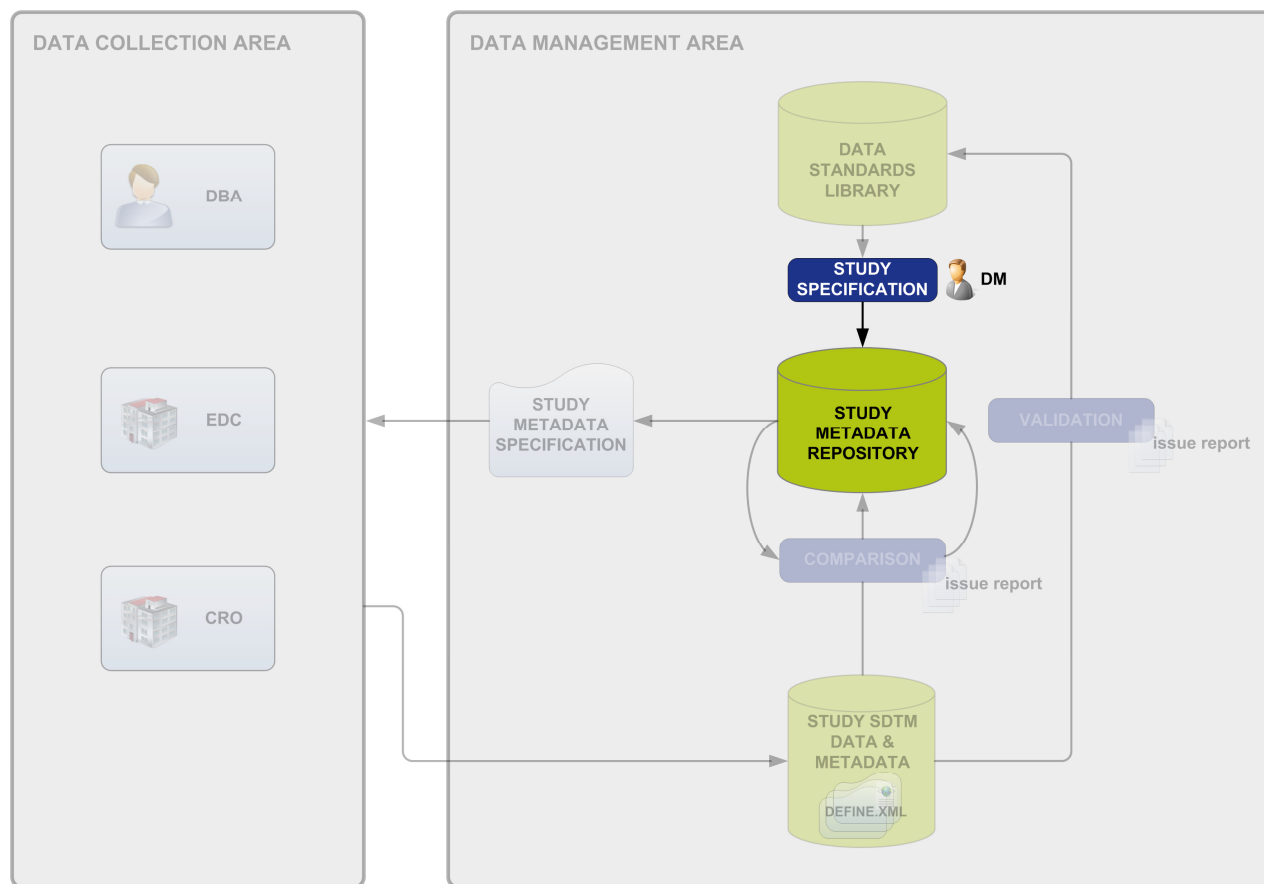
- The Data Standards Library contains :
 - Standard CRF templates (CDASH)
 - Metadata definitions (SDTM, Therapeutic area)

→ *used to
create study
metadata*



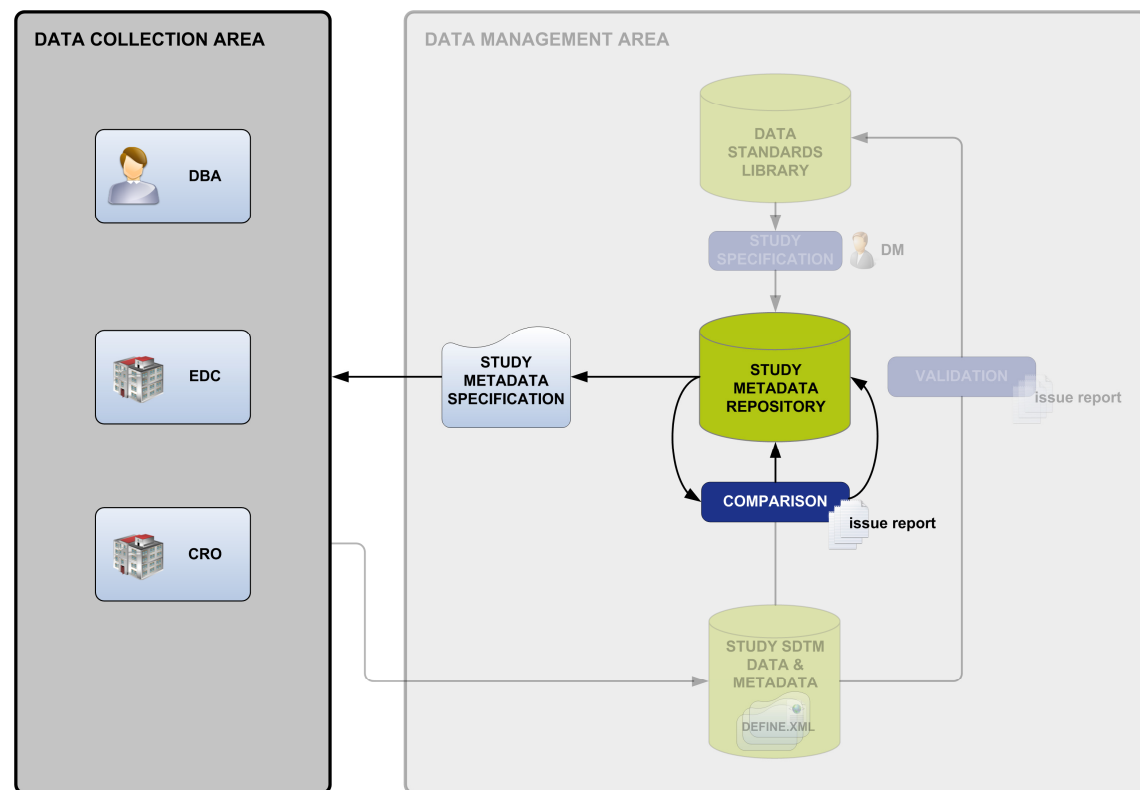
Process Overview

- Generating study metadata by selecting CRF templates from the Data Standards Library



Process Overview

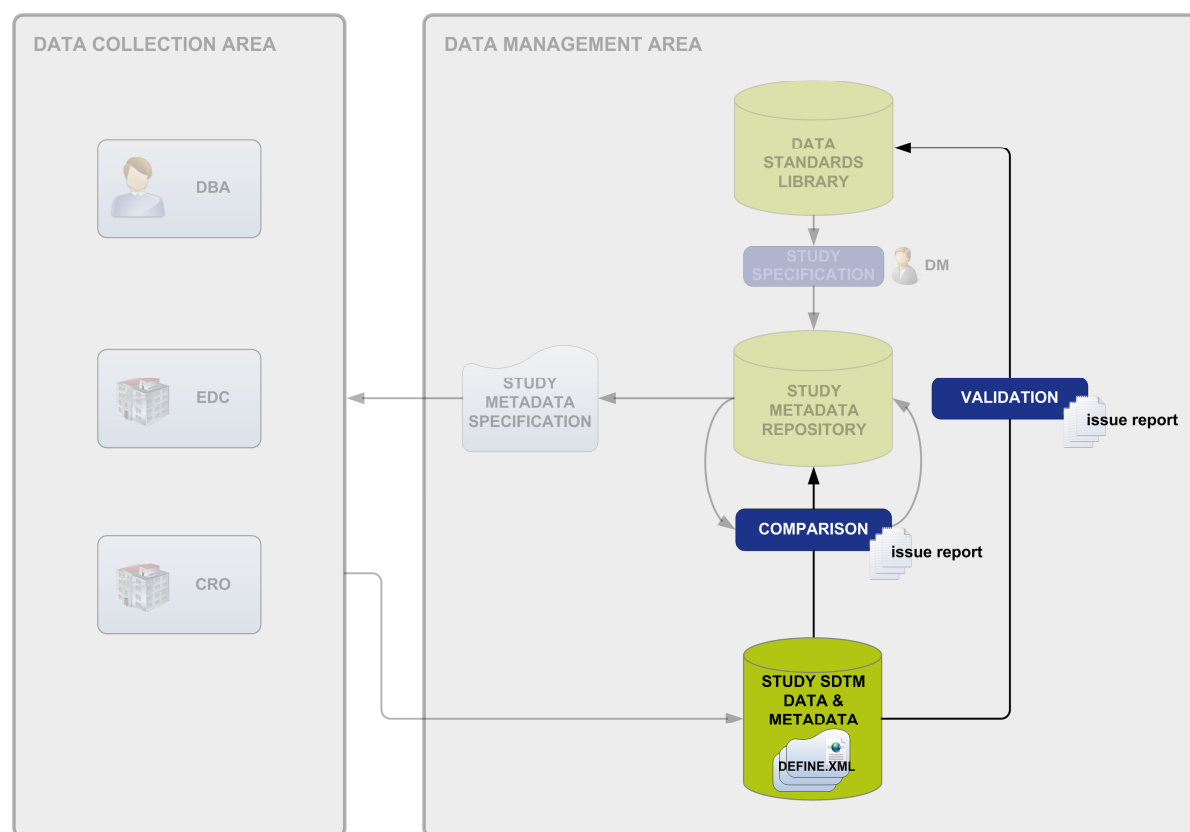
- Study Metadata Repository is used to measure the consistency of metadata
- Study metadata is sent to external or internal partners for database build



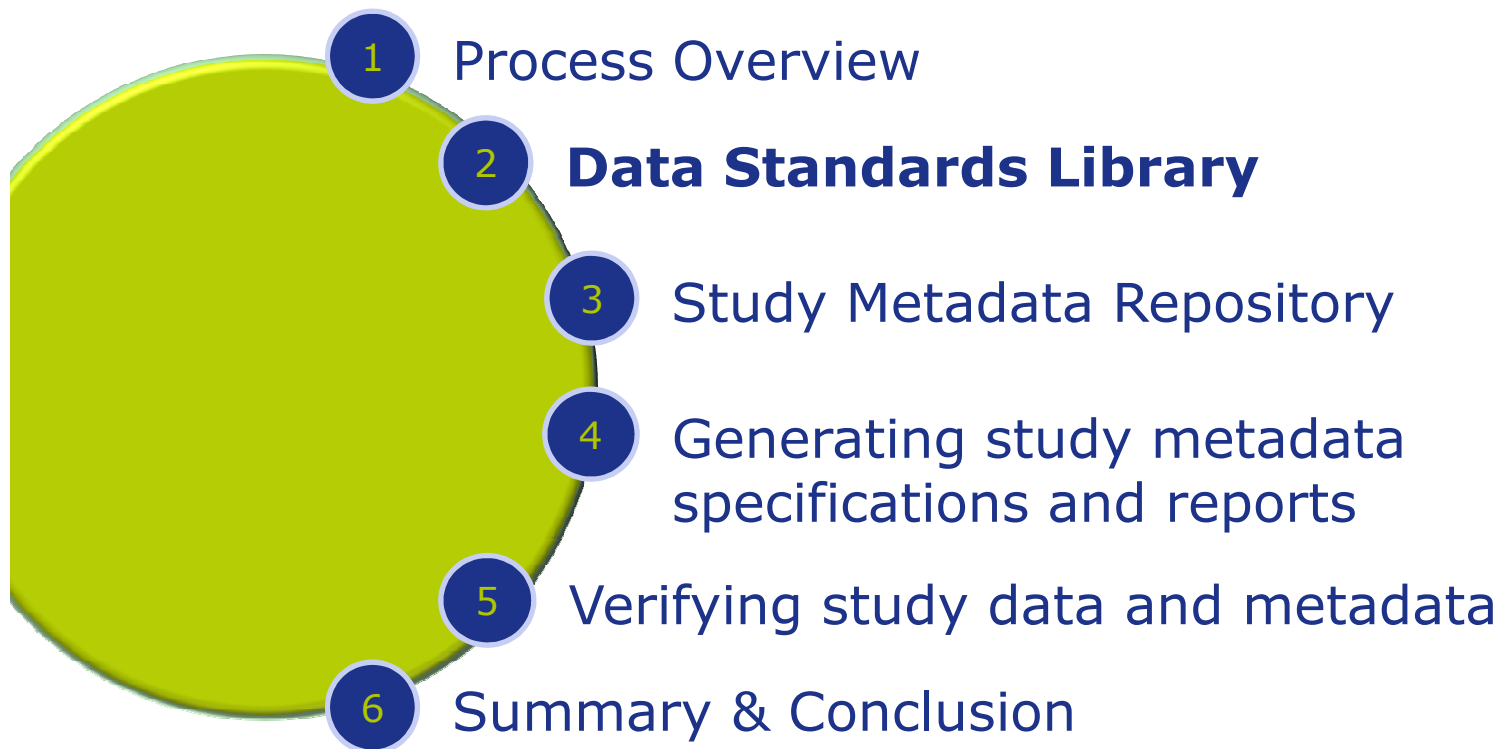
Process Overview

- After study build, data and metadata will be:
 - compared against the Study Metadata Repository
 - validated against the Data Standards Library

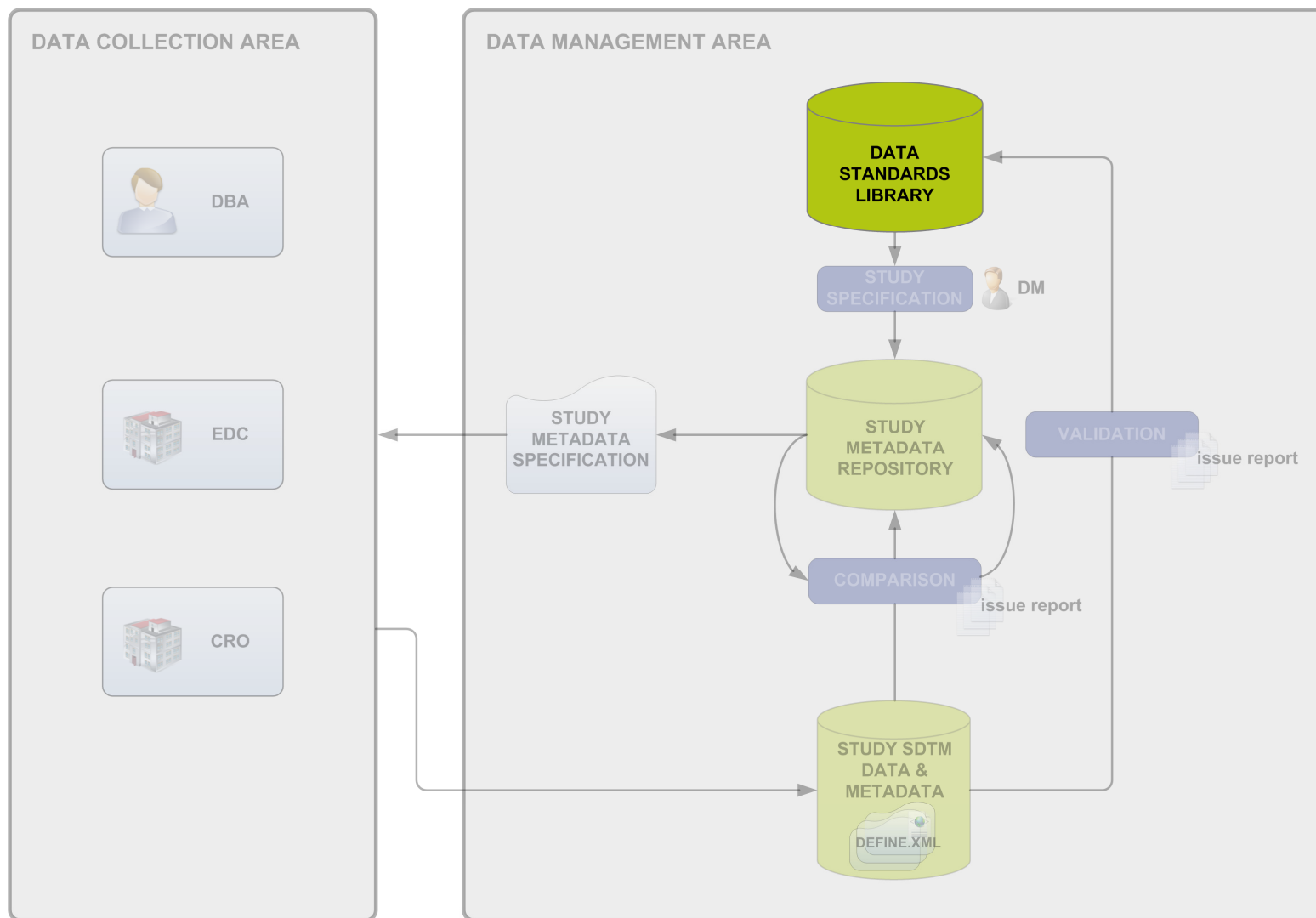
Verification



Agenda

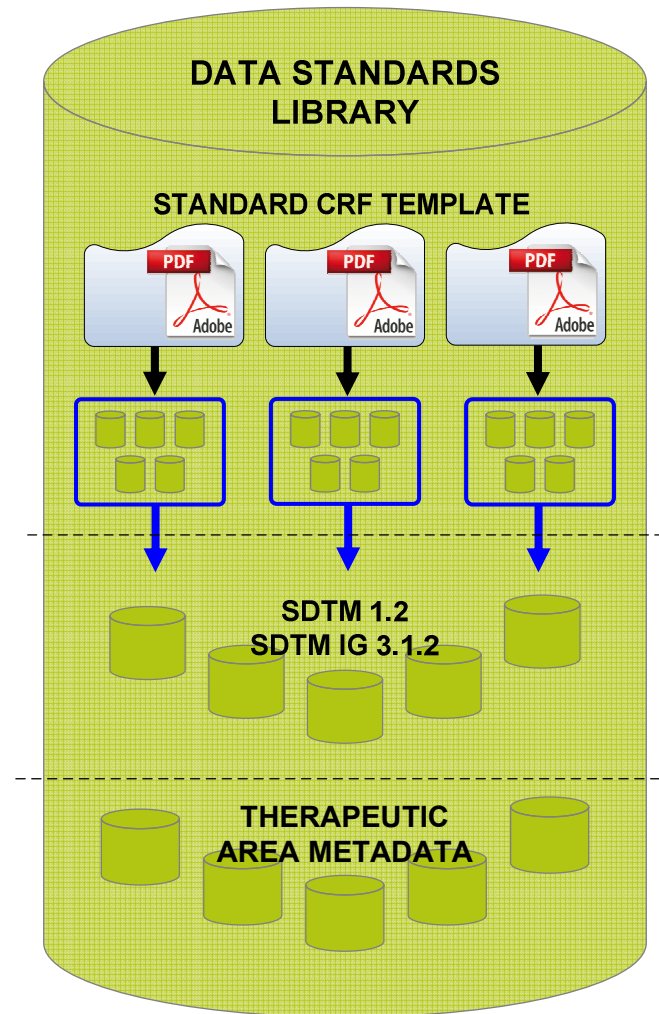


Data Standards Library



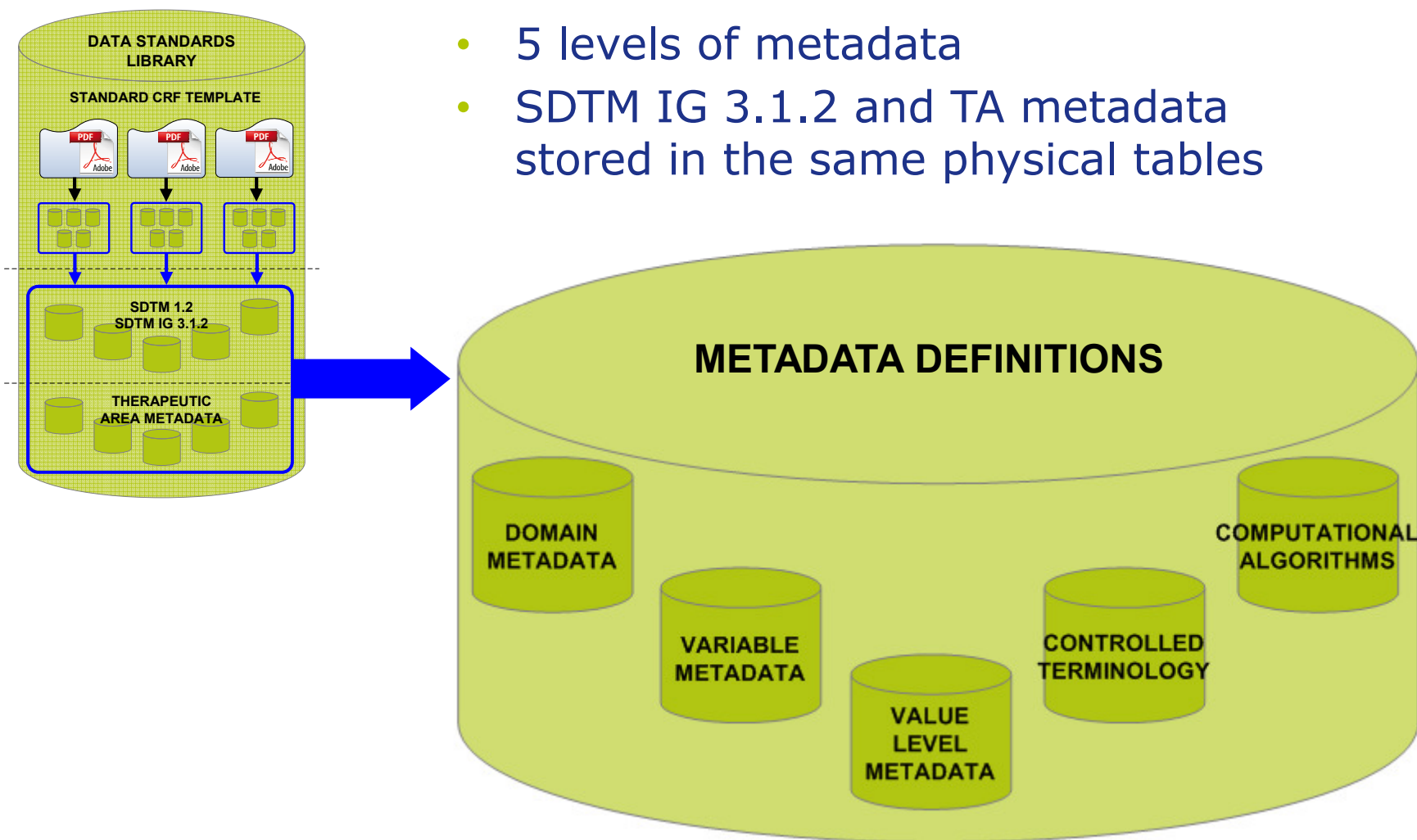
Data Standards Library

- The Data Standards Library contains :
 - Standard CRF Templates:
 - using CDASH metadata
 - with clustered SDTM metadata
 - annotated for CDASH
 - annotated for SDTM
 - Metadata Definitions :
 - SDTM Standards
 - Therapeutic Area Standards



Metadata Definitions

- 5 levels of metadata
- SDTM IG 3.1.2 and TA metadata stored in the same physical tables



Metadata Definitions

Table Metadata

Dataset Name
Dataset Label
Class
Structure
Purpose
Keys

Variable Metadata

Variable Name
Variable Label
Type
Controlled Terminology
Origin
Role
Comment

Value Level Metadata

Source Variable
Value
Label
Type
Controlled Terminology
Origin
Comment

Controlled Terminology

Code Value
Code Text

Computational Algorithms

Reference Name
Computation Method

Metadata Definitions

Table Metadata

Dataset Name
Dataset Label
Class
Structure
Purpose
Keys

Variable Metadata

Variable Name
Variable Label
Type
Controlled Terminology
Origin
Role
Comment

Value Level Metadata

Source Variable
Value
Label
Type
Controlled Terminology
Origin
Comment

Controlled Terminology

Code Value
Code Text

Computational Algorithms

Reference Name
Computation Method

Metadata Definitions

Table Metadata

Dataset Name
Dataset Label
Class
Structure
Purpose
Keys

Dataset	Description	Class	Structure	Purpose	Keys	Location
CM	Concomitant Medications	Interventions	One record per recorded medication occurrence or constant-dosing interval per subject	Tabulation	STUDYID, USUBJID, CMTRT, CMSTDTC	..\Tabulations\cm.xpt
EX	Exposure	Interventions	One record per constant dosing interval per subject	Tabulation	STUDYID, USUBJID, EXSTDTC, EXENDTC, EXDOSE, EXTRT	..\Tabulations\ex.xpt
AE	Adverse Events	Events	One record per adverse event per subject	Tabulation	STUDYID, USUBJID, AEDECOD, AETERM, AESTDTC	..\Tabulations\ae.xpt

Metadata Definitions

Table Metadata

Dataset Name
Dataset Label
Class
Structure
Purpose
Keys

Variable Metadata

Variable Name
Variable Label
Type
Controlled Terminology
Origin
Role
Comment

Value Level Metadata

Source Variable
Value
Label
Type
Controlled Terminology
Origin
Comment

Controlled Terminology

Code Value
Code Text

Computational Algorithms

Reference Name
Computation Method

Metadata Definitions

Variable Metadata

Variable Name
Variable Label
Type
Controlled Terminology
Origin
Role
Comment

Variable	Label	Type	Controlled Terminology	Origin	Role	Comment
STUDYID	Study Identifier	text		Protocol, CRF	IDENTIFIER	The STUDYID variable has a fixed format: 'XXXX-YYYY', where 'XXXX' indicates the 4-digit compound code and the 'YYYY' the 4-digit study code
DOMAIN	Domain Abbreviation	text	DOMAIN	Assigned	IDENTIFIER	
USUBJID	Unique Subject Identifier	text		Protocol	IDENTIFIER	The USUBJID variable has a fixed format: 'XXXX-YYYY-ZZZZZ', where 'XXXX' indicates the 4-digit compound code, 'YYYY' the 4-digit study code and 'ZZZZZ' the 6-digit patient code

Metadata Definitions

Table Metadata

Dataset Name
Dataset Label
Class
Structure
Purpose
Keys

Variable Metadata

Variable Name
Variable Label
Type
Controlled Terminology
Origin
Role
Comment

Value Level Metadata

Source Variable
Value
Label
Type
Controlled Terminology
Origin
Comment

Controlled Terminology

Code Value
Code Text

Computational Algorithms

Reference Name
Computation Method

Metadata Definitions

Value Level Metadata

Source Variable

Value

Label

Type

Controlled Terminology

Origin

Comment

Value Level Metadata (ValueList.VS.VSTESTCD)						
Source Variable	Value	Label	Type	Controlled Terminology	Origin	Comment
VSTESTCD	BMI	BODY MASS INDEX	integer		Derived	See Computational Method: VS.BMI
VSTESTCD	DIABP	DIASTOLIC BLOOD PRESSURE	integer		CRF	
VSTESTCD	HEIGHT	HEIGHT	integer		CRF	
VSTESTCD	PULSE	PULSE RATE	integer		CRF	
VSTESTCD	SYSBP	SYSTOLIC BLOOD PRESSURE	integer		CRF	
VSTESTCD	WEIGHT	WEIGHT	integer		CRF	

Metadata Definitions

Table Metadata

Dataset Name
Dataset Label
Class
Structure
Purpose
Keys

Variable Metadata

Variable Name
Variable Label
Type
Controlled Terminology
Origin
Role
Comment

Value Level Metadata

Source Variable
Value
Label
Type
Controlled Terminology
Origin
Comment

Controlled Terminology

Code Value
Code Text

Computational Algorithms

Reference Name
Computation Method

Metadata Definitions

POSITION, Reference Name (Codelist.POSITION)	
Code Value	Code Text
STANDING	STANDING
SUPINE	SUPINE

SEX, Reference Name (Codelist.SEX)	
Code Value	Code Text
F	Female
M	Male
U	Unknown

Controlled Terminology

Code Value
Code Text

Metadata Definitions

Table Metadata

Dataset Name
Dataset Label
Class
Structure
Purpose
Keys

Variable Metadata

Variable Name
Variable Label
Type
Controlled Terminology
Origin
Role
Comment

Value Level Metadata

Source Variable
Value
Label
Type
Controlled Terminology
Origin
Comment

Controlled Terminology

Code Value
Code Text

Computational Algorithms

Reference Name
Computation Method

Metadata Definitions

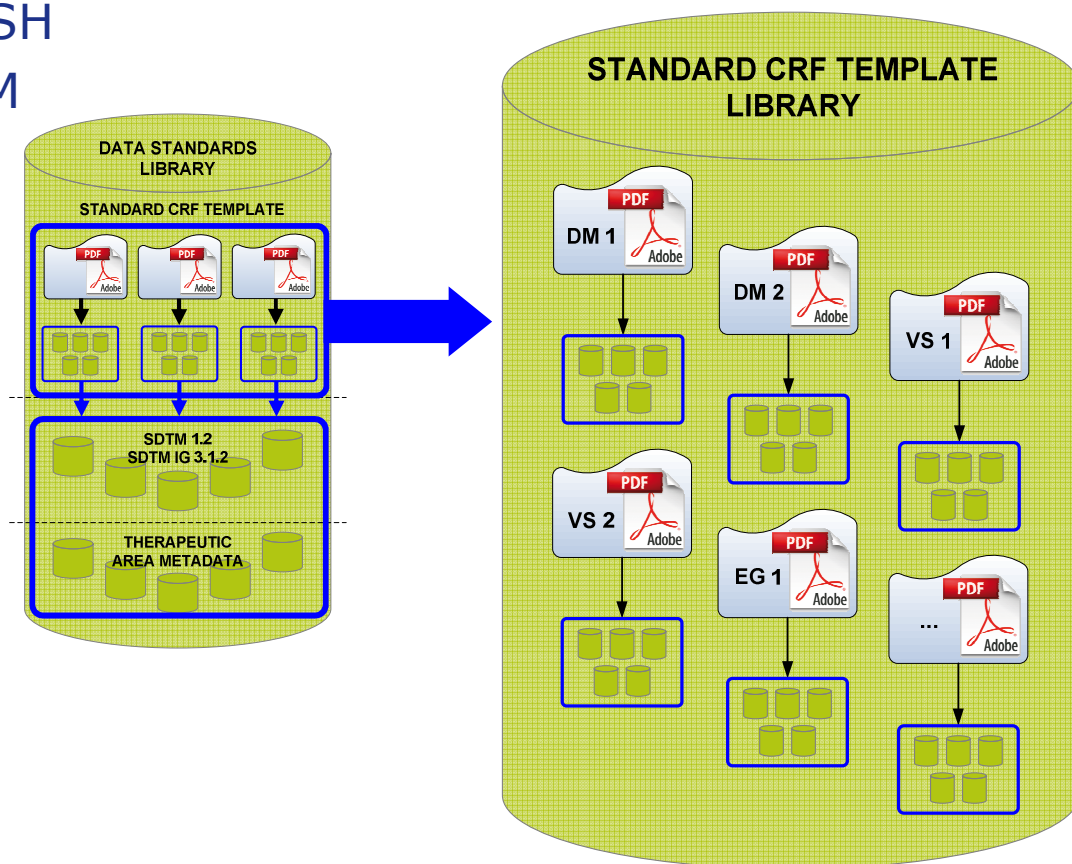
Computational Algorithms (VS.BMI)	
Reference Name	Computation Method
VS.BMI	Equals to WEIGHT DIVIDED BY HEIGHT SQUARED
Computational Algorithms (VS.VSDY)	
Reference Name	Computation Method
VS.VSDY	Equals to (VS.VSDTC - DM.RFSTDTC) + 1 IF VSDTC is on or after RFSTDTC Equals to (VS.VSDTC - DM.RFSTDTC) IF VSDTC precedes RFSTDTC

Computational
Algorithms

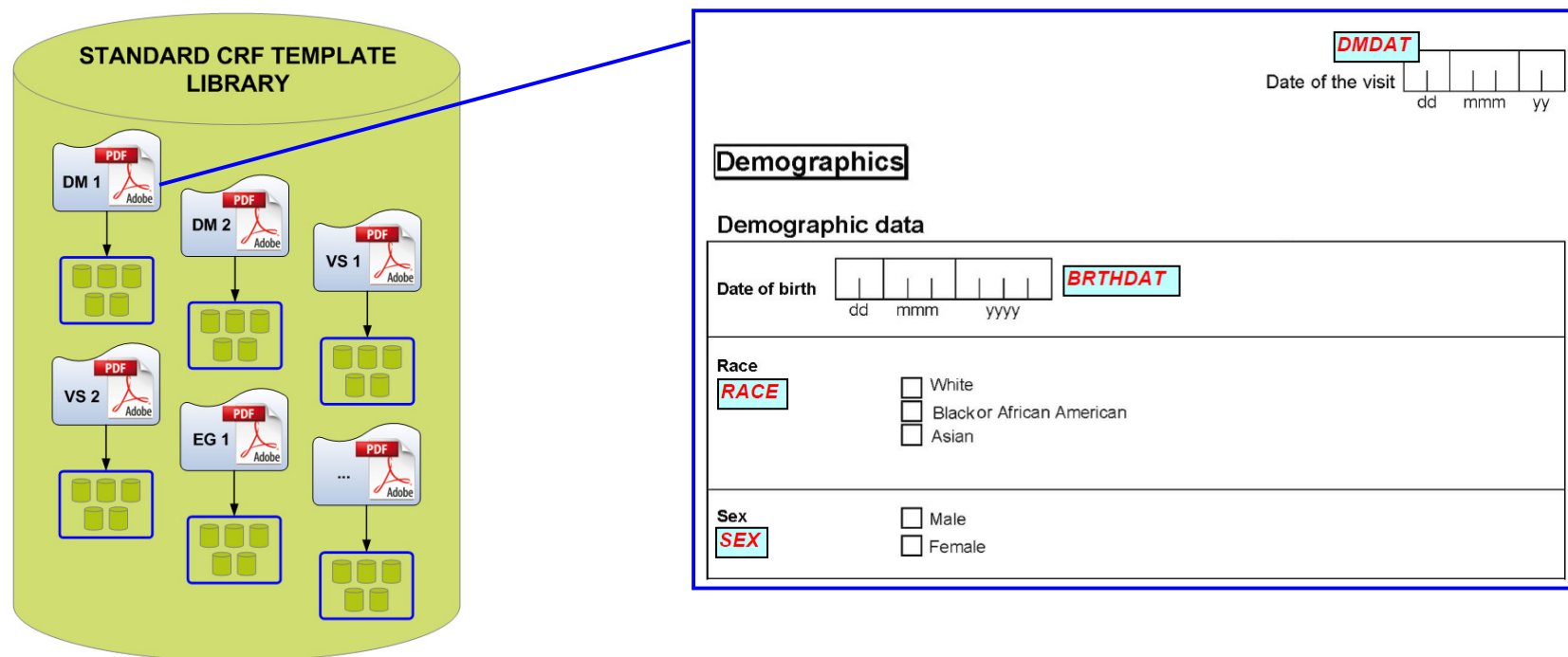
Reference Name
Computation Method

Standard CRF Templates

- Standard CRF Template Library :
 - using CDASH metadata
 - with clustered SDTM metadata
 - annotated for CDASH
 - annotated for SDTM

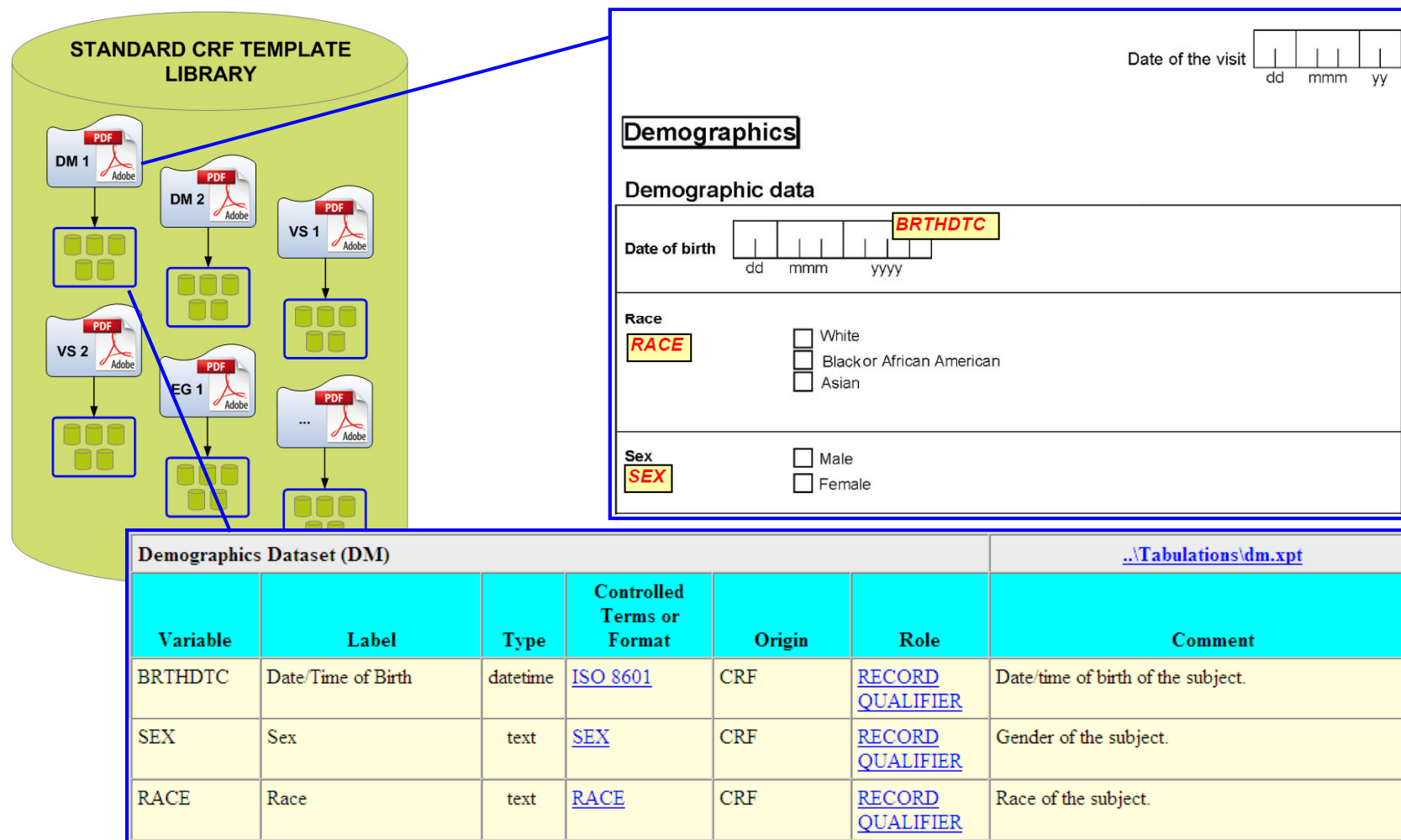


Standard CRF Templates - CDASH

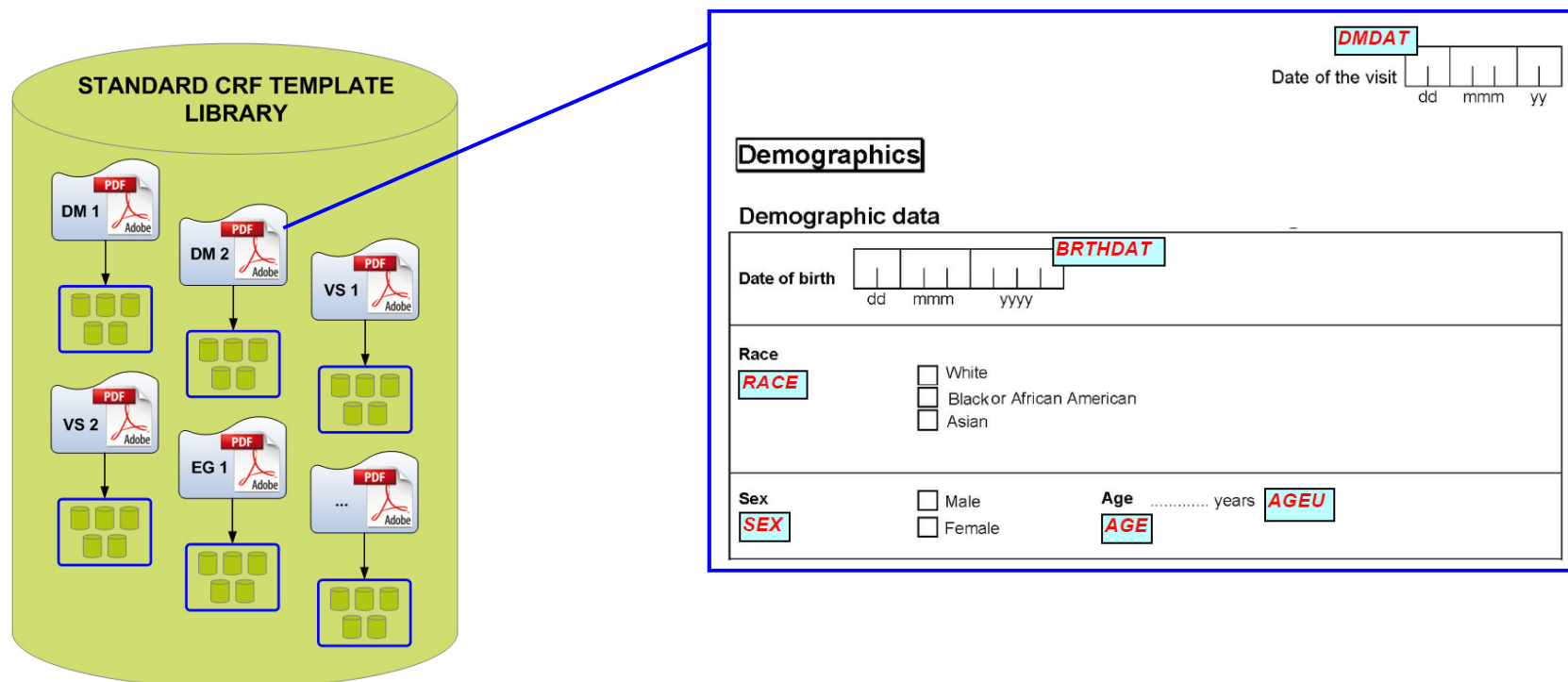


- CRF template using CDASH annotations

Standard CRF Templates - SDTM

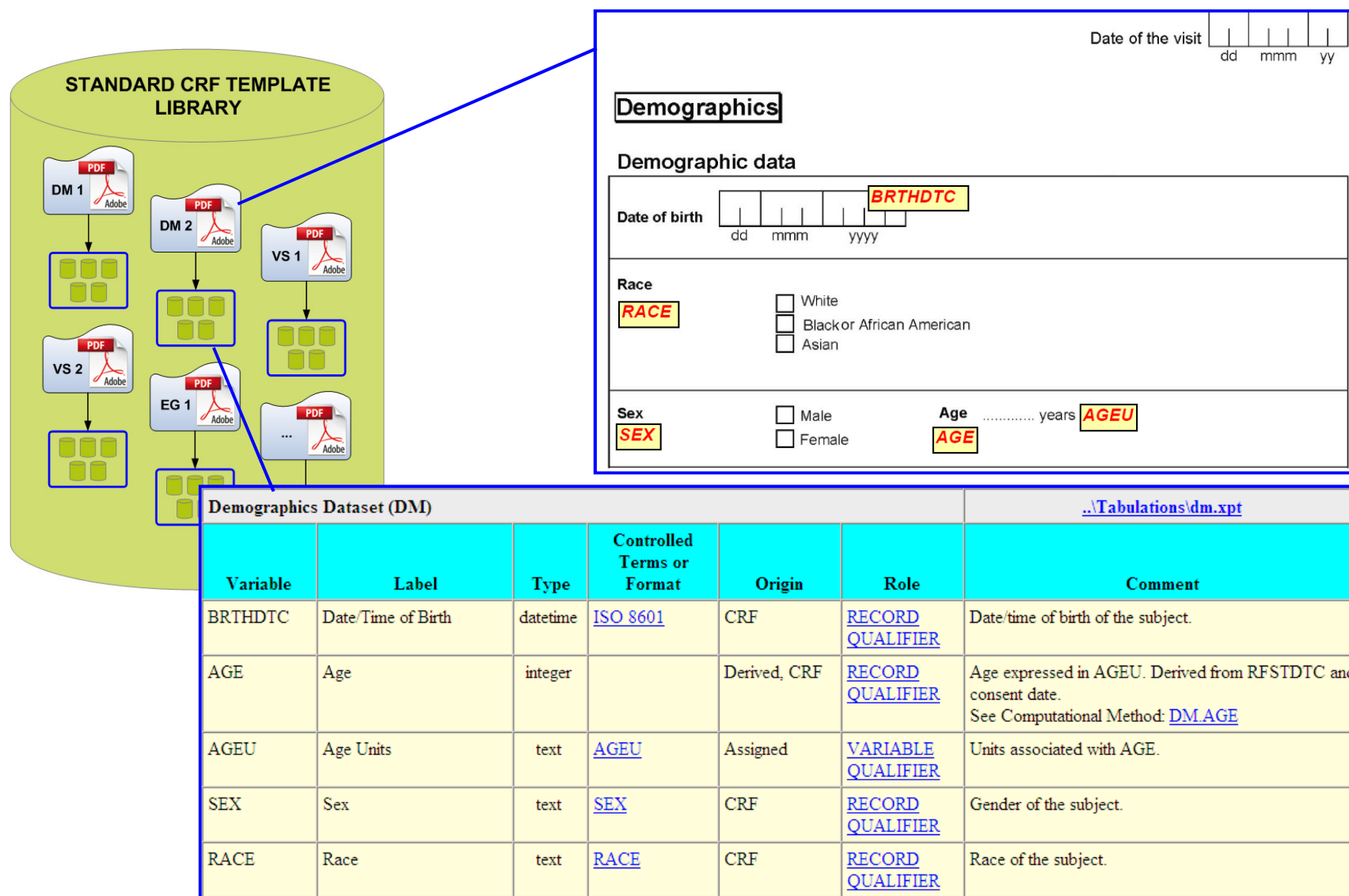


Standard CRF Templates - CDASH

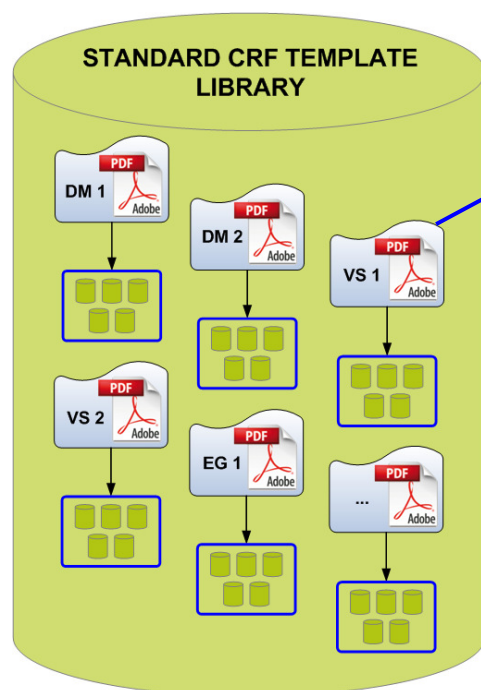


- CRF template using CDASH annotations

Standard CRF Templates - SDTM



Standard CRF Templates - CDASH



Vital Signs

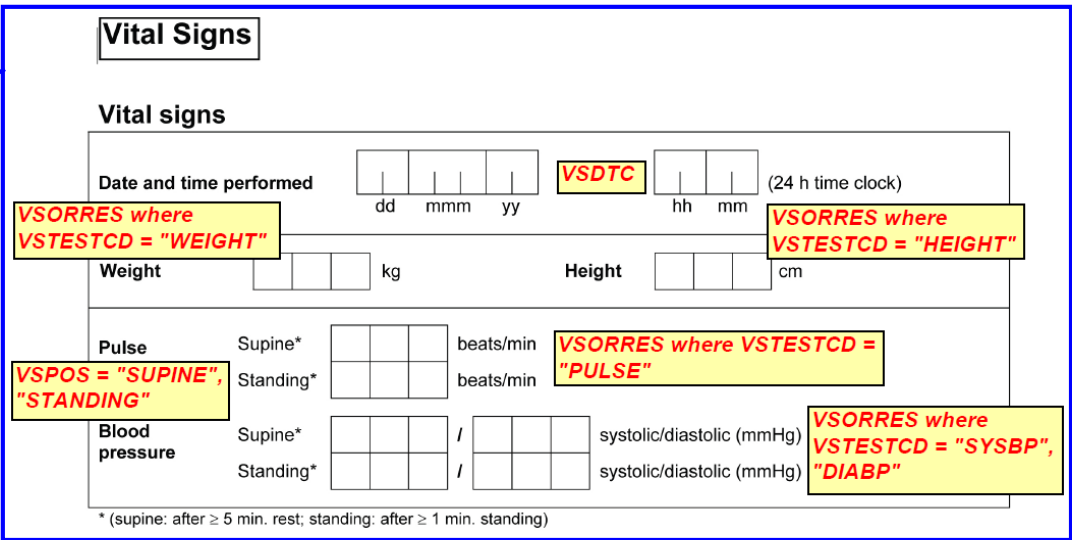
Vital signs

Date and time performed		VSDAT	dd	mmm	yy	VSTIM	hh	mm	(24 h time clock)
VSTEST	Weight	VSORRES		kg	VSTEST	Height		cm	VSORRESU
VSTEST	Pulse	Supine*	VSORRES	beats/min	VSORRESU	Standing*	VSPOS	beats/min	
VSTEST	Blood pressure	Supine*	VSORRES	systolic/diastolic (mmHg)	VSORRESU	Standing*	VSORRES	systolic/diastolic (mmHg)	VSORRESU

*(supine: after ≥ 5 min. rest; standing: after ≥ 1 min. standing)

- CRF template using CDASH annotations

Standard CRF Templates - SDTM

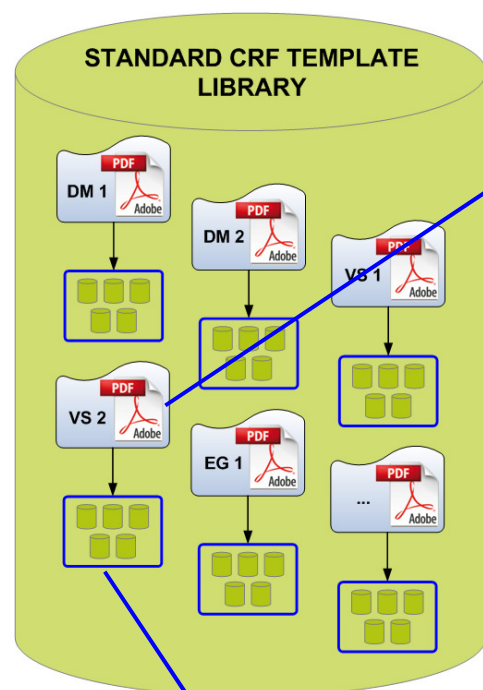


Value Level Metadata (ValueList.VS.VSTESTCD)							
Source Variable	Value	Label	Type	Controlled Terminology	Origin	Role	Comment
VSTESTCD	DIABP	DIASTOLIC BLOOD PRESSURE	integer		CRF		
VSTESTCD	HEIGHT	HEIGHT	integer		CRF		
VSTESTCD	PULSE	PULSE RATE	integer		CRF		
VSTESTCD	SYSBP	SYSTOLIC BLOOD PRESSURE	integer		CRF		
VSTESTCD	WEIGHT	WEIGHT	integer		CRF		

[illegible]

- CRF template using CDASH annotations

Standard CRF Templates - SDTM



Vital Signs

Vital signs

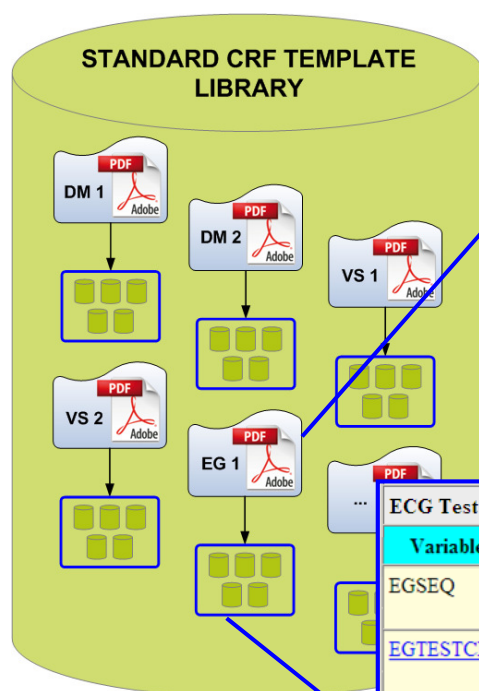
Date and time performed **VSDTC** (24 h time clock)
dd mmm yy hh mm

Pulse beats/min **VSORRES where VSTESTCD = "PULSE"**

Blood pressure / systolic/diastolic (mmHg)
VSORRES where VSTESTCD = "SYSBP", "DIABP"

Value Level Metadata (ValueList.VS.VSTESTCD)							
Source Variable	Value	Label	Type	Controlled Terminology	Origin	Role	Comment
VSTESTCD	DIABP	DIASTOLIC BLOOD PRESSURE	integer		CRF		
VSTESTCD	PULSE	PULSE RATE	integer		CRF		
VSTESTCD	SYSBP	SYSTOLIC BLOOD PRESSURE	integer		CRF		

Standard CRF Templates - SDTM



GENERAL DATA TRANSFER AGREEMENT

TYPE OF DATA: ECG

Version 1.1

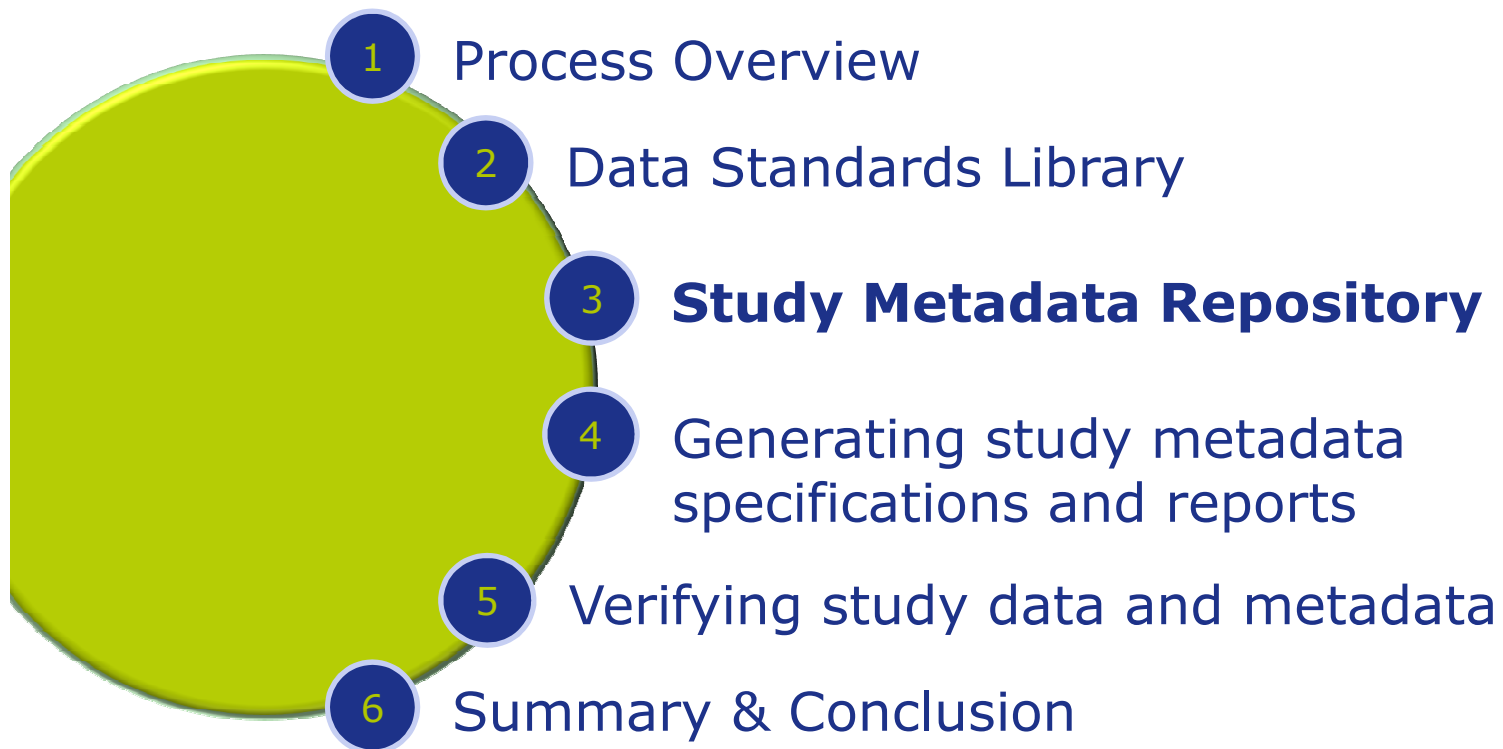
Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format	Core	Length	CDISC Notes
EGTEST	ECG Test or Examination Name	Char		Req		40Verbatim name of the test or examination used to obtain the measurement or finding. The value in EGTEST cannot be longer than 40 characters. Examples: Summary (Mean) PR Duration, Summary (Mean) QT Duration
EGCAT	Category for ECG	Char		Perm		40Used to categorize ECG observations across subjects. Examples: MEASUREMENT, FINDING, INTERVAL
EGSCAT	Subcategory for ECG	Char		Perm		40A further categorization of the ECG.
EGPOS	ECG Position of Subject	Char	POSITION	Perm		20Position of the subject during a measurement or examination. Examples: SUPINE, STANDING, SITTING.
EGORRES	Result or Finding in Original Units	Char		Exp		200Result of the ECG measurement or finding as originally received or collected. Examples of expected values are 62 or 0.151 when the result is an interval or measurement, or "ATRIAL FIBRILLATION" or "QT PROLONGATION" when the result is a finding.
EGORRESU	Original Units	Char	UNIT	Perm		20Original units in which the data were collected. The unit for EGORRES. Examples: sec or msec.
EGSTRESC	Character Result/Finding in Std Format	Char		Exp		200Contains the result value for all findings, copied or derived from EGORRES in a standard format or standard units.

ECG Test Results Dataset (EG)

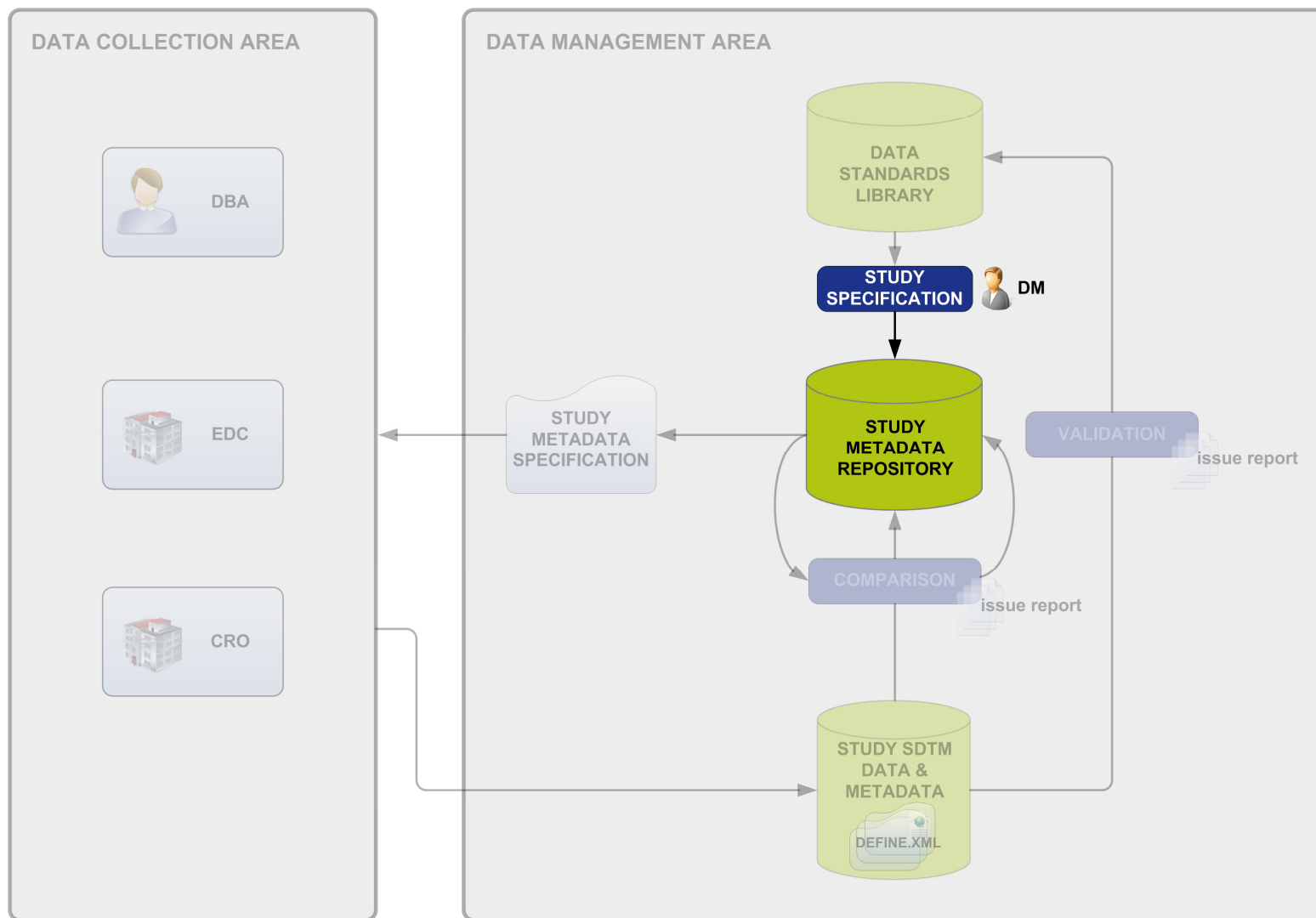
...\\tabulations\sdm\eg.xpt

Variable	Label	Type	Controlled Terminology	Origin	Role	Comment
EGSEQ	Sequence Number	integer		Derived	IDENTIFIER	
EGTESTCD	ECG Test or Examination Short Name	text		Assigned	TOPIC	
EGTEST	ECG Test or Examination Name	text	EGTEST	Derived	SYNONYM QUALIFIER	
EGORRES	Result or Finding in Original Units	text		eDT, CRF	RESULT QUALIFIER	-ORRES value can be empty if associated with a populated --STAT and/or -REASND variable or if associated/linked with record(s) in the corresponding SUPPQUAL or the CO domain. All other observations with empty -ORRES fields are removed from the domain.

Agenda

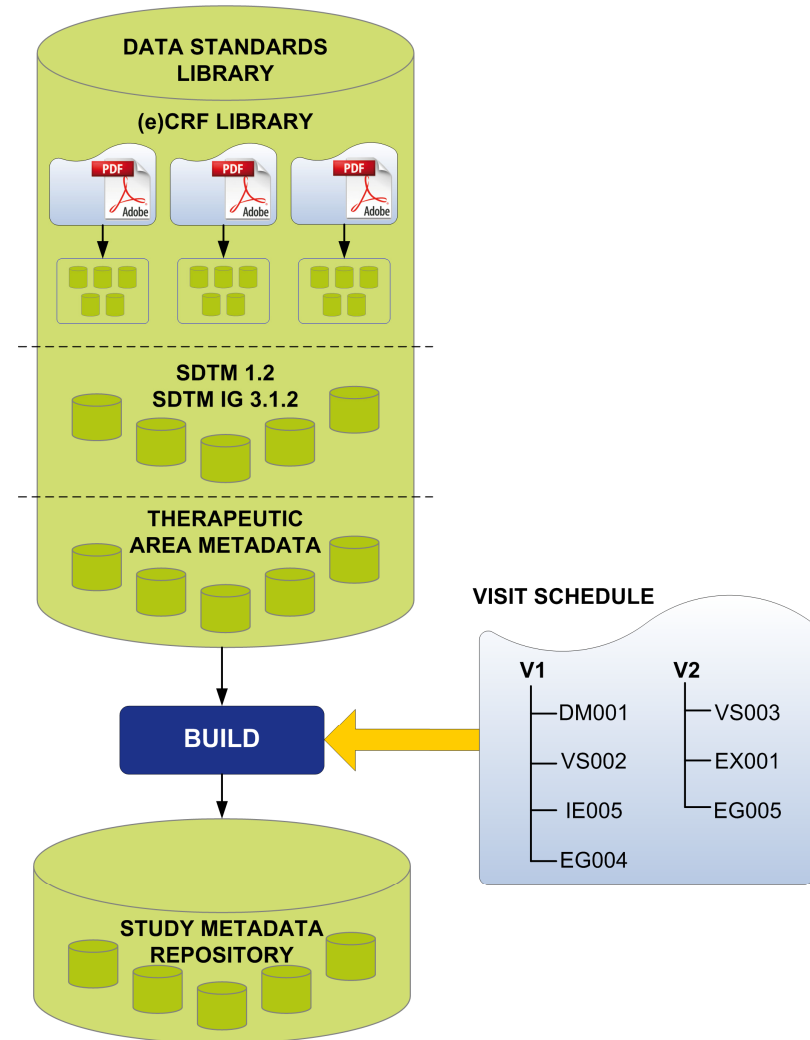


Study Specifications



Generation of Study Specifications

- Generation of study metadata requires :
 - visit schedule and CRF templates
 - study parameters

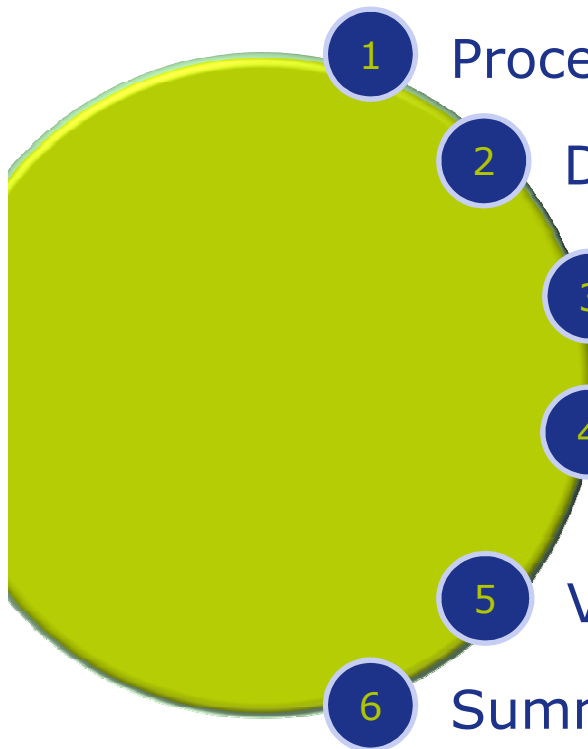


Study Metadata Repository

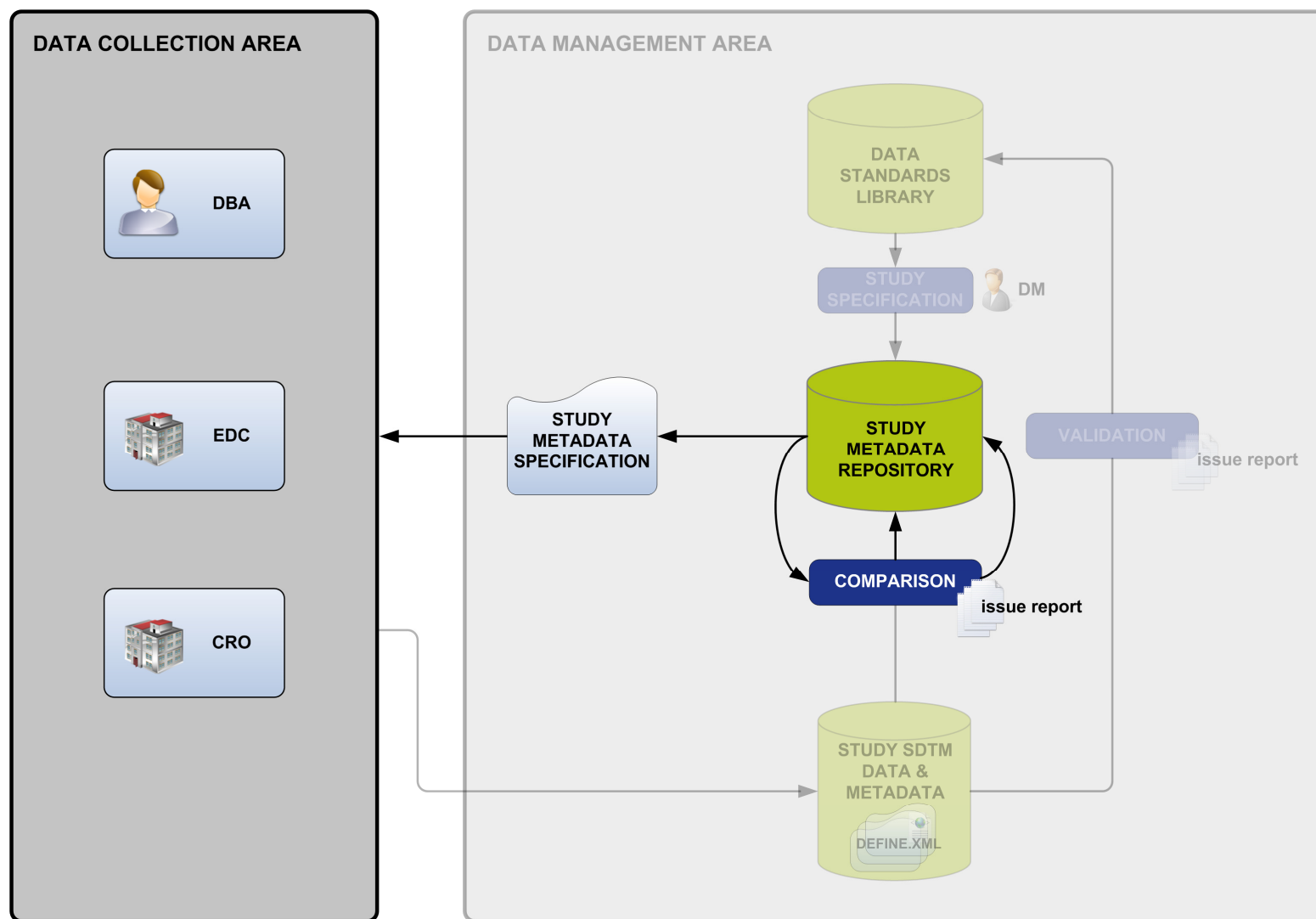
- The Study Metadata Repository contains :

STUDYID	PAR_NAME	PAR_VALUE	PAR_ID	PAR_DESCRIPTION
ABC-0001	StudyOID	ABC-0001	Header	Study Identifier(Required Internal ODM key used to uniquely identify the entity)
ABC-0001	StudyName	ABC-0001	Header	Study Name (Part of the Globalvariables in ODM markup (required section))
ABC-0001	StudyDescription	ABC-0001	Header	Sudy title (Part of the Globalvariables in ODM markup (required section))
ABC-0001	ProtocolName	ABC-0001	Header	Name of the Protocol (Part of the Globalvariables in ODM markup (required section))
ABC-0001	ExternalDictionary	MedDRA	Footer	Name of the External Dictionary (Optional)
ABC-0001	DictionaryVersion	13.0	Footer	Version of the External Dictionary (Optional)
ABC-0001	DictionaryOID	MEDDRA	Footer	Unique Identifier for an External Dictionary (Optional)
ABC-0001	DictionaryDesc	ADVERSE EVENT DICTIONARY	Footer	Description of External Dictionary
ABC-0001	ExternalDictionary	WHO Drug	Footer	Name of the External Dictionary (Optional)
ABC-0001	DictionaryVersion	WHO-DD 10.MAR	Footer	Version of the External Dictionary (Optional)
ABC-0001	DictionaryOID	WHODRUG	Footer	Unique Identifier for an External Dictionary (Optional)
ABC-0001	DictionaryDesc	DRUG DICTIONARY	Footer	Description of External Dictionary
ABC-0001	DictionaryOID	ISO 8601	Footer	Unique Identifier for an External Dictionary (Optional)
ABC-0001	DictionaryDesc	ISO 8601	Footer	Description of External Dictionary
ABC-0001	DictionaryOID	ISO 3166	Footer	Unique Identifier for an External Dictionary (Optional)
ABC-0001	DictionaryDesc	ISO 3166	Footer	Description of External Dictionary

Agenda

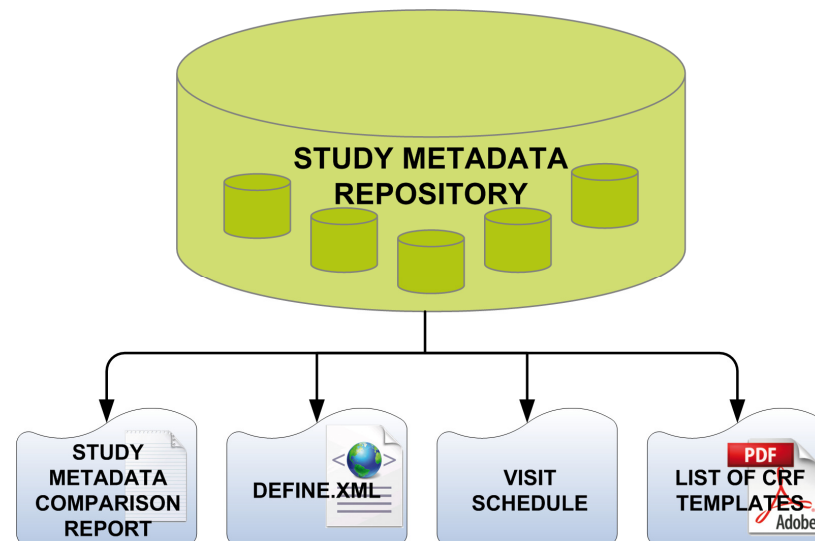
- 
- 1 Process Overview
 - 2 Data Standards Library
 - 3 Study Metadata Repository
 - 4 **Generating study metadata specifications and reports**
 - 5 Verifying study data and metadata
 - 6 Summary & Conclusion

Study Metadata Specifications and Reports



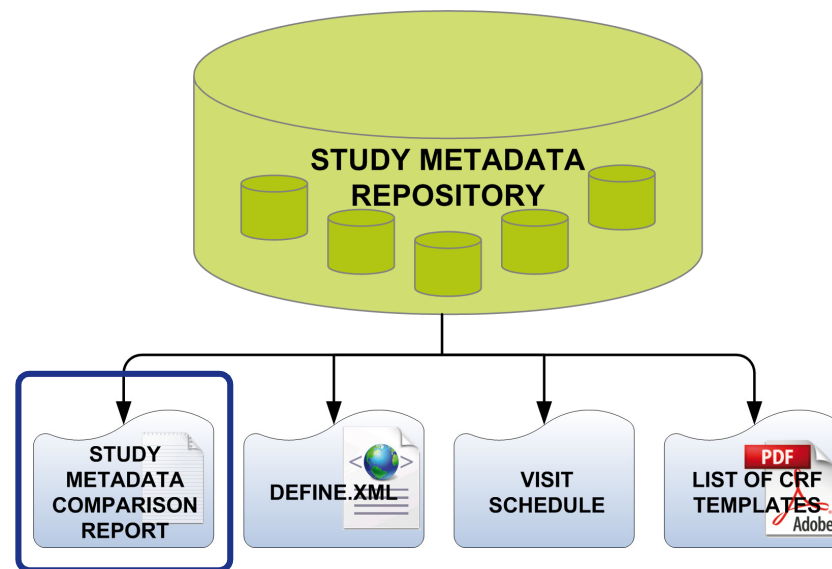
Study Metadata Specifications and Reports

- Comparison reports
 1. study metadata vs. Data Standards Library
 2. study metadata across the Study Metadata Repository
- Study metadata specifications
 1. 5 levels of metadata
 2. define.xml
 3. list of CRF templates
 4. visit schedule



Study Metadata Comparison Reports

- Comparison reports
 1. To compare study metadata against the Data Standards Library
 2. To measure metadata consistency between studies



Comparison done across all the levels of metadata existing in the Study Metadata Repository

1. Comparison with the Data Standards Library

- 3 different statuses :
 - **Ok** : metadata present in Data Standards Library
 - **Study Specific** : metadata not present in the Data Standards Library
 - **Difference** : difference in the metadata value

Status	Dataset	Variable	Length	A1000-0000	Global Standards
Difference	CM	USUBJID	40	X	
Difference	CM	USUBJID	20		X
Study specific	CM	VISIT	40	X	
Study specific	CM	VISITNUM	8	X	
Ok	CO	CODTC	20	X	X

2. Comparison across studies

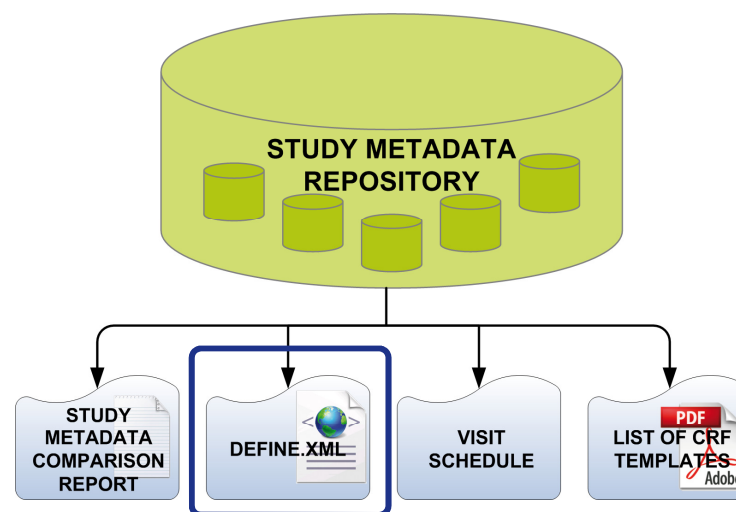
- 2 different statuses:
 - Ok** : no metadata conflict
 - Difference** : difference in metadata value between studies

Status	Dataset	Label	A1000_0000	A1000_0001	A1000_0002
Ok	AE	Adverse Events	X	X	X

Status	Reference	Computational algorithm	A1000_0000	A1000_0001	A1000_0002
Difference	DM.AGE	Equal to (DS.DSSTDTC-DM.BRTHDTC)/365.25 where DSTERM = 'INFORMED CONSENT OBTAINED'		X	
Difference	DM.AGE	Equals to (DS.DSSTDTC-DM.BRTHDTC)/365.25 where DSTERM = 'INFORMED CONSENT OBTAINED'	X		X
Difference	DM	Demographic		X	
Ok	DS	Disposition	X	X	X

Study Specifications for Data Collection Area

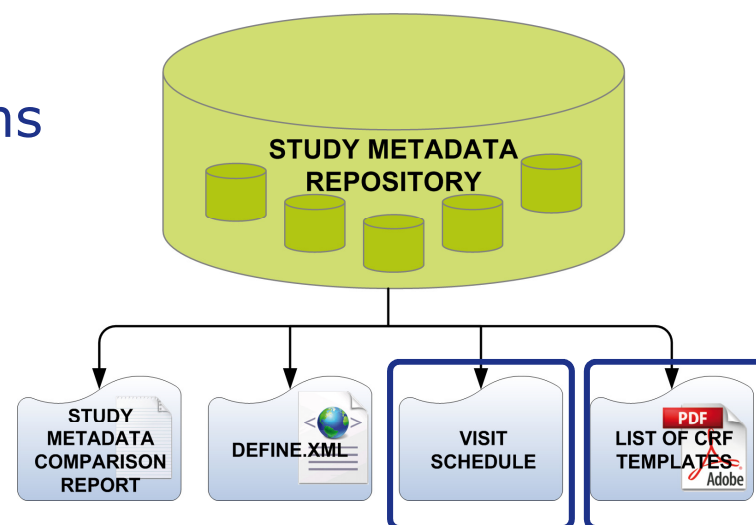
- Study specifications
 - 5 tables, one for each level of metadata
 - Define.xml



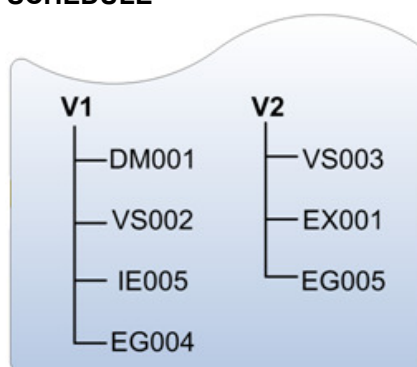
Demographics Dataset (DM)						..\Tabulations\dm.xpt
Variable	Label	Type	Controlled Terms or Format	Origin	Role	Comment
STUDYID	Study Identifier	text		Protocol, CRF	IDENTIFIER	Unique identifier for a study within the submission.
DOMAIN	Domain Abbreviation	text	DOMAIN	Assigned	IDENTIFIER	Two-character abbreviation for the domain most relevant to the observation.
USUBJID	Unique Subject Identifier	text		Derived	IDENTIFIER	Unique subject identifier within the submission. This is the concatenation of STUDYID, SITEID and SUBJID.
SUBJID	Subject Identifier for the Study	text		CRF	TOPIC	Unique subject identifier within the study.

Study Specifications for Data Collection Area

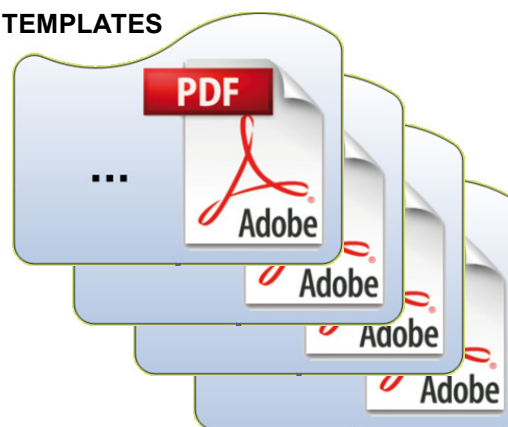
- Study metadata specifications
 3. Visit schedule
 4. CRF templates



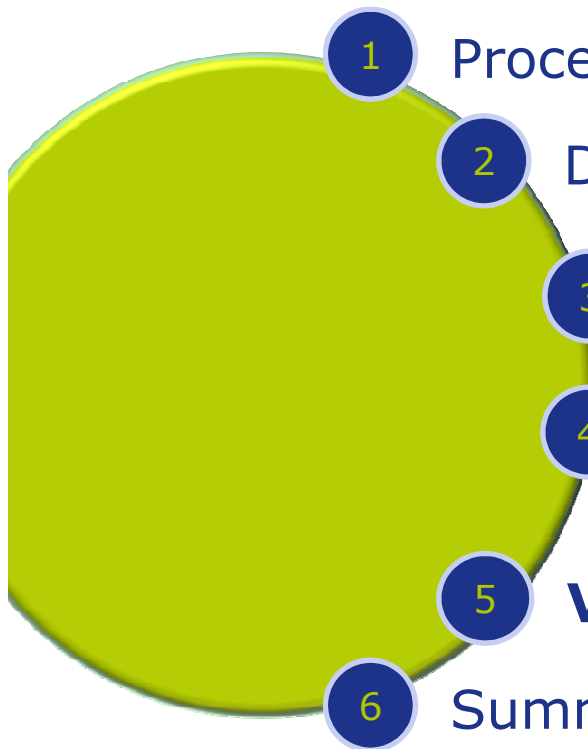
VISIT SCHEDULE



CRF TEMPLATES

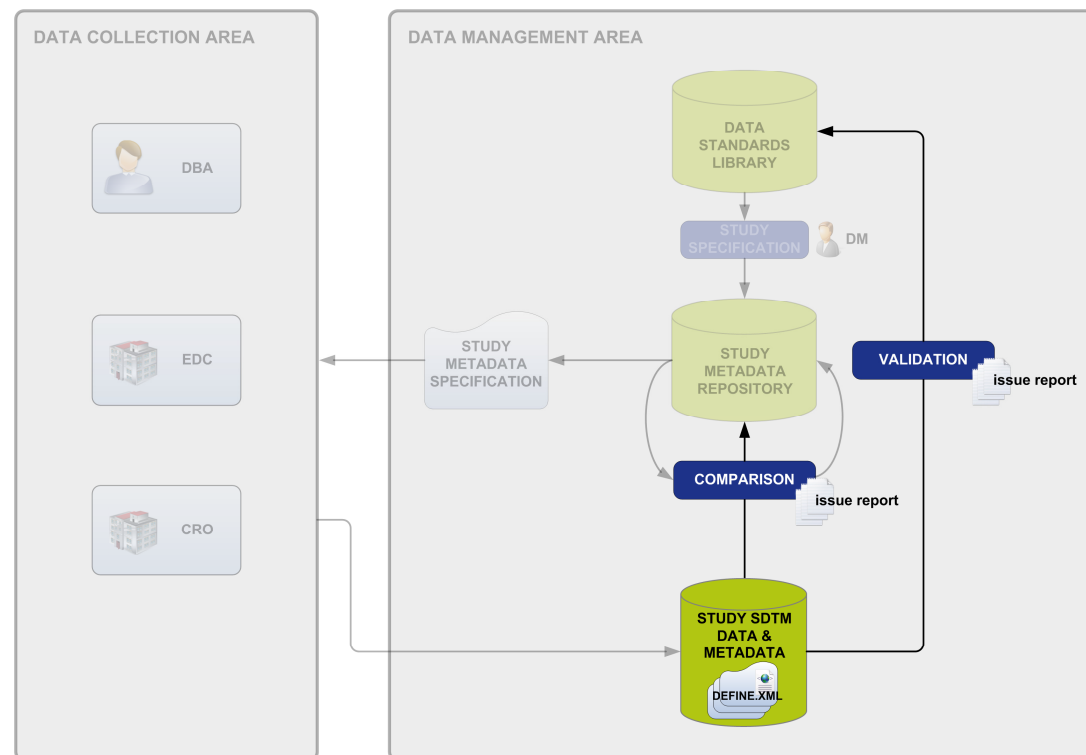


Agenda

- 
- 1 Process Overview
 - 2 Data Standards Library
 - 3 Study Metadata Repository
 - 4 Generating study metadata specifications and reports
 - 5 **Verifying study data and metadata**
 - 6 Summary & Conclusion

Data and Metadata Verification

- The data and metadata are verified by a two-step process :
 - Comparison against the Study Metadata Library
 - Validation against the Data Standards Library

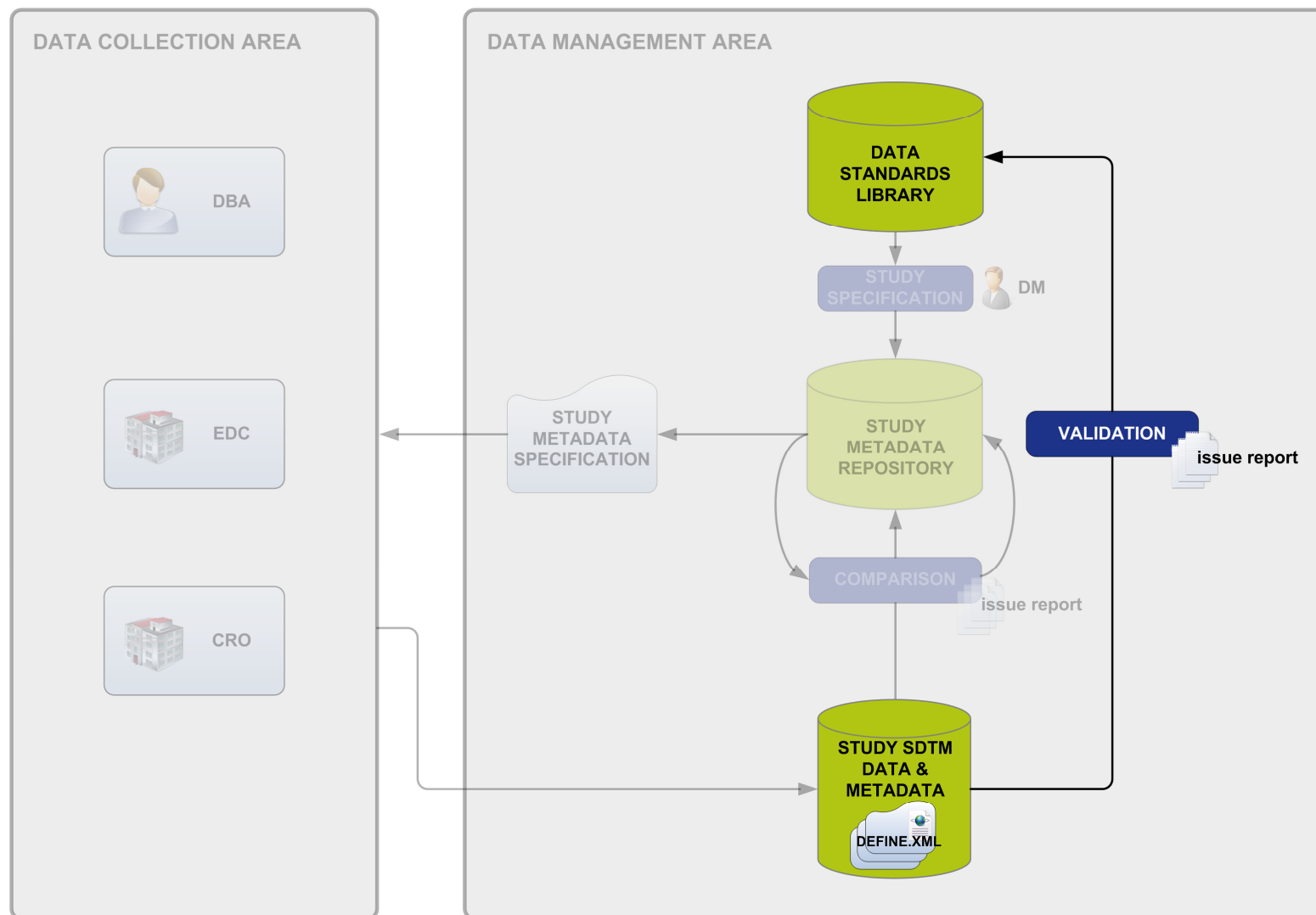


Comparison of study data and metadata

- 4 different statuses :
 - Ok** : no conflict in metadata
 - Difference** : difference in the metadata value
 - Missing** : metadata not in study data
 - Addition** : metadata not in study specifications

Status	Dataset	Variable	Label	Specifications	Study
Difference	AE	AEACN	Action Taken with Study Trtment		X
Difference	AE	AEACN	Action Taken with Study Treatment	X	
Ok	AE	AEBODSYS	Body System or Organ Class	X	X
Missing	AE	AECAT	Category for Adverse Event	X	
Addition	AE	AECOM	Comment about Adverse Event		X
Ok	AE	AECONTRT	Concomitant or Additional Trtmnt Given	X	X

Validation of study data and metadata

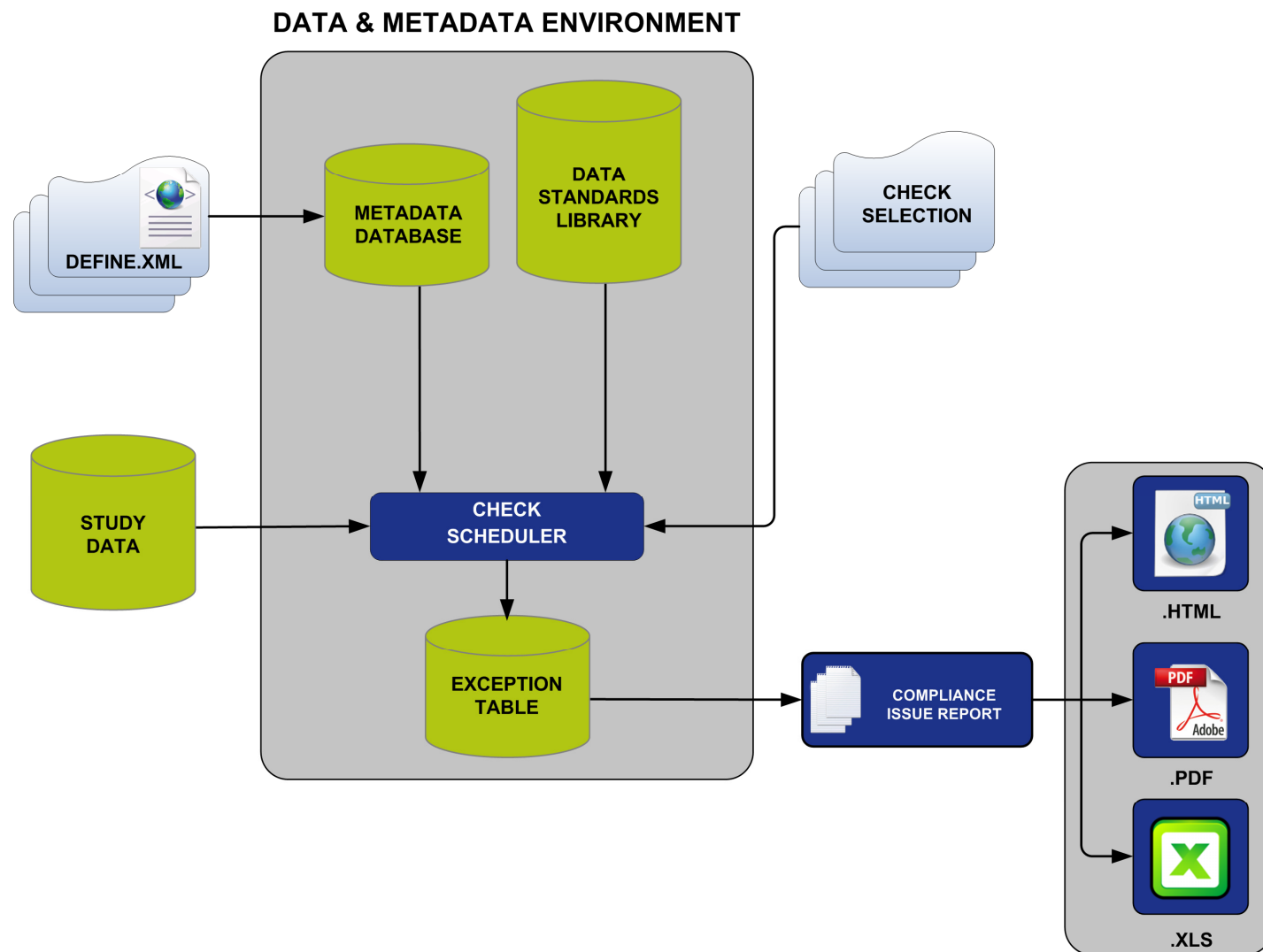


Validation - Checks



Business & Decision Life Sciences								
Data Model Version	Check	Sub CheckID	Error_Message	Type	Applies to	Severity	WebSDM	WebSDM Rule Description
SDTMIG V3.1.2	1001		DATE VARIABLE does not have a ISO 8601 format.	Data	All observation classes, TD, Special Purpose	Medium	IR5002	Identifies records where the DATE variable does not conform to the ISO 8601 standard.
SDTMIG V3.1.2	1002		The STANDARD UNIT is not consistent per LAB TEST or EXAMINATION.	Data	Findings	High	IR5006	Identifies Measurement, Test or Examination records where the STANDARD UNIT is not consistent with the standard units value (Standard Units).
SDTMIG V3.1.2	1003	1	The AE is serious, but none of the serious criteria is answered 'Y'.	Data	Events	Medium	IR5008	Identifies records where Seriousness is 'Y' but none of the following criteria are answered 'Y': Congenital Anomaly or Birth Defect, Results in Death, Requires or Prohibits Further Medical or Surgical Intervention, or Other Medically Important Seriousness.
SDTMIG V3.1.2	1003	2	The AE is not serious, but one of the serious criteria is answered 'Y'.	Data	Events	Medium	IR5008	Identifies records where Seriousness is not 'Y' but one of the following criteria is answered 'Y': Congenital Anomaly or Birth Defect, Results in Death, Requires or Prohibits Further Medical or Surgical Intervention, or Other Medically Important Seriousness.
SDTMIG V3.1.2	1004	1	A STANDARD RESULT is completed, but the STATUS is completed.	Data	Findings	Medium		
SDTMIG V3.1.2	1004	2	No STANDARD RESULT is completed, but STATUS is not completed.	Data	Findings	Low		
SDTMIG V3.1.2	1005		START DATE/TIME OF OBSERVATION falls after the END DATE/TIME OF OBSERVATION.	Data	All observation classes, Special Purpose	High	IR5101	Identifies records that violate the constraint that the START DATE/TIME OF OBSERVATION is not later than the END DATE/TIME OF OBSERVATION.
SDTMIG V3.1.2	1006		The OCCURRENCE is completed differently from 'Y' or 'N'.	Data	Events, Interventions	High	IR5106	Identifies records that violate the constraint that the OCCURRENCE is completed with 'Y' or 'N'.
SDTMIG V3.1.2	1007		The LENGTH of the --TEST variable is more than 40 characters.	Data	Findings	High	IR5113	Identifies records that violate the constraint that the length of the --TEST variable is not more than 40 characters.
SDTMIG V3.1.2	1008	1	The --TESTCD starts with a number.	Data	Findings	High	IR5114	Identifies records that violate the constraint that the --TESTCD does not start with a number or contain special characters.
SDTMIG V3.1.2	1008	2	The --TESTCD contains special characters.	Data	Findings	High	IR5114	Identifies records that violate the constraint that the --TESTCD does not start with a number or contain special characters.
SDTMIG V3.1.2	1009		An ORIGINAL RESULT is completed but the STANDARD RESULT (CHARACTERISTIC) is missing.	Data	Findings	High	IR5135	Identifies records that violate the constraint that if the ORIGINAL RESULT is completed, the STANDARD RESULT (CHARACTERISTIC) must also be completed. Identifies user-defined code lists.

Validation - Flowchart



Validation - Reports



- 3 reports are created:
 - Exception log**
 - Exception summary
 - Exception table

LIST_OF_CHECKS	ERROR_LEVEL	COMPLETION TIME	ELAPSED_TIME
Check1002	The Check Completed Successfully	2011-03-02T11-39-37	.
Check1005	The Check Completed Successfully	2011-03-02T11-39-37	0:00:00
Check1007	The Check Completed Successfully	2011-03-02T11-39-38	0:00:01
Check1008	The Check Completed Successfully	2011-03-02T11-39-38	0:00:00
Check1009	The Check Completed Successfully	2011-03-02T11-39-38	0:00:00
Check1011	The Check Completed Successfully	2011-03-02T11-39-41	0:00:03
Check1012	The Check Completed Successfully	2011-03-02T11-39-48	0:00:07
Check1102	The Check Completed Successfully	2011-03-02T11-39-49	0:00:01
Check1103	The Check Completed Successfully	2011-03-02T11-39-49	0:00:00
Check1122	The Check Completed Successfully	2011-03-02T11-39-49	0:00:00
Check1132	The Check Completed Successfully	2011-03-02T11-39-49	0:00:00
Check1242	The Check Completed Successfully	2011-03-02T11-39-50	0:00:01
Check1243	The Check Completed Successfully	2011-03-02T11-39-51	0:00:01
Check1249	The Check Completed Successfully	2011-03-02T11-39-51	0:00:00
Check1251	The Check Completed Successfully	2011-03-02T11-39-51	0:00:00
Check1414	The Check Completed Successfully	2011-03-02T11-39-52	0:00:01

Validation - Reports



- 3 reports are created:
 - Exception log
 - Exception summary**
 - Exception table

LIST OF CHECKS	EXCEPTIONS_COUNT	CHECK TYPE
Check1002	5	DATA
Check1005	8	DATA
Check1007	2	DATA
Check1008	1	DATA
Check1009	3	DATA
Check1010	1	METADATA
Check1011	8	DATA
Check1012	44	DATA
Check1102	1	DATA
Check1103	3	DATA
Check1122	1	DATA
Check1132	2	DATA
Check1242	4	DATA
Check1243	1	DATA
Check1249	4	DATA
Check1251	1	DATA
Check1414	2	DATA

Validation - Reports



- 3 reports are created:
 - Exception log
 - Exception summary
 - Exception table**

Reported Exception Record Identifiers Exception Attributes

CHECKID	ERRORMSG	SEVERITY	STUDYID	DOMAIN	USUBJID	KEY1CD	KEY1	VAR1CD	VAR1	VAR2CD	VAR2	
1002	EXAMINATION. The STANDARD UNIT is not consistent per TEST or	High	1000-0000	LB				LBTESTCD	LIPASE	LBSTRESU	%	
1002	EXAMINATION. The STANDARD UNIT is not consistent per TEST or	High	1000-0000	LB				LBTESTCD	LIPASE	LBSTRESU	mmol/L	
1002	EXAMINATION. The STANDARD UNIT is not consistent per TEST or	High	1000-0000	LB				LBTESTCD	ALKP	LBSTRESU	U	
1002	EXAMINATION. The STANDARD UNIT is not consistent per TEST or	High	1000-0000	LB				LBTESTCD	ALKP	LBSTRESU	U/L	
1002	EXAMINATION. The STANDARD UNIT is not consistent per TEST or	High	1000-0000	LB				LBTESTCD	LIPASE	LBSTRESU	I/L	
1005	DATE/TIME OF OBSERVATION. START DATE/TIME OF OBSERVATION falls after the END	High	1000-0000	AE	1000-0000-00019	AESQ	1	AESTDTC	1998-12-01T17:00	AEENDTC	1998-12-01T16:30	
1005	DATE/TIME OF OBSERVATION. START DATE/TIME OF OBSERVATION falls after the END	High	1000-0000	AE	1000-0000-00037	AESQ	1	AESTDTC	1998-12-15T07:40	AEENDTC	1998-11-25T08:00	
1005	DATE/TIME OF OBSERVATION. START DATE/TIME OF OBSERVATION falls after the END	High	1000-0000	SE	1000-0000-00004	SESEQ	2	SESTDTC	1998-11-17T08:06:00	SEENDTC	1998-11-17T07:07:00	
1005	DATE/TIME OF OBSERVATION. START DATE/TIME OF OBSERVATION falls after the END	High	1000-0000	SE	1000-0000-00009	SESEQ	2	SESTDTC	1998-11-24T08:00:00	SEENDTC	1997-11-24T08:01:00	
1005	DATE/TIME OF OBSERVATION. START DATE/TIME OF OBSERVATION falls after the END	High	1000-0000	SV	1000-0000-00022			SVSTDTC	1998-12-03	SVENDTC	1998-12-02	
1005	DATE/TIME OF OBSERVATION. START DATE/TIME OF OBSERVATION falls after the END	High	1000-0000	SV	1000-0000-00010			SVSTDTC	1998-11-24	SVENDTC	1998-11-22	
1005	DATE/TIME OF OBSERVATION. START DATE/TIME OF OBSERVATION falls after the END	High	1000-0000	SV	1000-0000-00014			SVSTDTC	1998-11-16	SVENDTC	1998-10-16	
1005	DATE/TIME OF OBSERVATION. START DATE/TIME OF OBSERVATION falls after the END	High	1000-0000	SV	1000-0000-00035			SVSTDTC	1998-12-17	SVENDTC	1997-12-17	
1007	characters. The LENGTH of the --TEST variable is more than 40	High	1000-0000	LB	1000-0000-00001	LBSEQ	35	LBTEST	ERY. MEAN CORPUSCULAR HB	LENGTH		43
1007	characters. The LENGTH of the --TEST variable is more than 40	High	1000-0000	LB	1000-0000-00030	LBSEQ	54	LBTEST	CONCENTRATION MEAS	LENGTH		41
1008	The --TESTCD starts with a number.	High	1000-0000	LB	1000-0000-00001	LBSEQ	3	LBTESTCD	PARTIAL THROMBOPLASTINE TIME			
1009	RESULT (CHARACTERISTIC) is missing.	High	1000-0000	IE	1000-0000-00006	IESEQ	1	IEORRES	2AMYL	IESTRESC		
1009	RESULT (CHARACTERISTIC) is missing.	High	1000-0000	SC	1000-0000-00031	SCSEQ	1	SCORRES	N	SCSTRESC		
1009	RESULT (CHARACTERISTIC) is missing.	High	1000-0000	SC	1000-0000-00014	SCSEQ	1	SCORRES	R-K	SCSTRESC		
1009	RESULT (CHARACTERISTIC) is missing.	High	1000-0000	SC	1000-0000-00014	SCSEQ	1	SCORRES	W-B	SCSTRESC		
1009	RESULT (CHARACTERISTIC) is missing.	High	1000-0000	SC	1000-0000-00022	SCSEQ	1	SCORRES	M-U	SCSTRESC		

LIFE SCIENCES

Agenda

- 
- 1 Process Overview
 - 2 Data Standards Library
 - 3 Study Metadata Repository
 - 4 Generating study metadata specifications and reports
 - 5 Verifying study data and metadata
 - 6 **Summary & Conclusion**

Summary

- *A multi layered data standards library (CDASH and SDTM) that facilitates building the study metadata*
- *The use of the study metadata repository to generate study specifications and to control consistencies across studies*
- *Validate data and metadata against the data standards library*
- *Compare data and metadata against the study metadata repository*
- *Complex data management processes can be automated through workflows*
- *Similar process can be designed for ADaM*

Conclusion

- *Maximize re-usability and increase efficiency of data management processes*
- *Building a foundation facilitating future efficiencies (eProtocol, Data as a Clinical Data Repository, ...)*
- *Move data standards upfront to:*
 - *Drive the study specification process and to reduce study specification cycle times*
 - *Significantly increase quality of data and metadata*
 - *Reduce re-work and optimizes resources utilization and cost*
- *This process also allows the use of therapeutic area standards*
- *Disconnects standard definitions from technical implementation through a common metadata layer*

