

CDISC (CDASH/SDTM) integration into OC/RDC

Generation of real-time SDTM
datasets and metadata
through a generic SDTM
converter mechanism

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

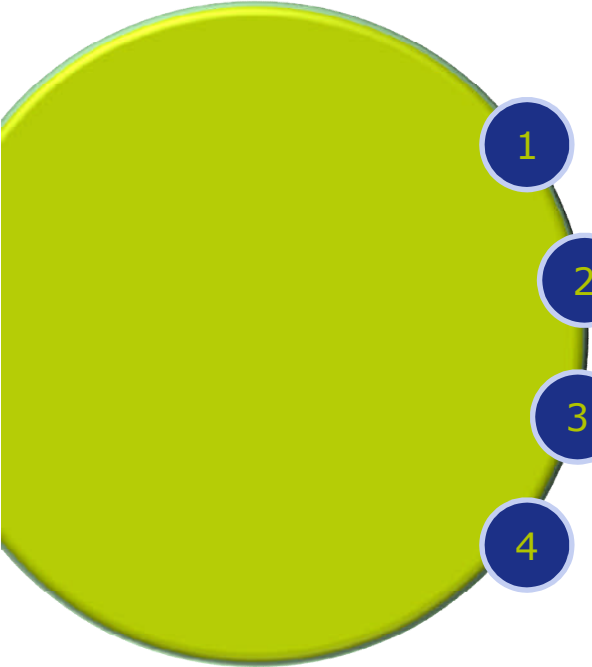
Sint-Lambertusstraat 141 Rue Saint-
Lambert
1200 Brussels

www.businessdecision-lifesciences.com



DRIVE YOUR PERFORMANCE®

Agenda

- 
- 1 Introduction
 - 2 CDASH/SDTM in OC/RDC
 - 3 SDTM Converter Mechanism
 - 4 Conclusion

Agenda

1

Introduction

2

CDASH/SDTM in OC/RDC

3

SDTM Converter Mechanism

4

Conclusion

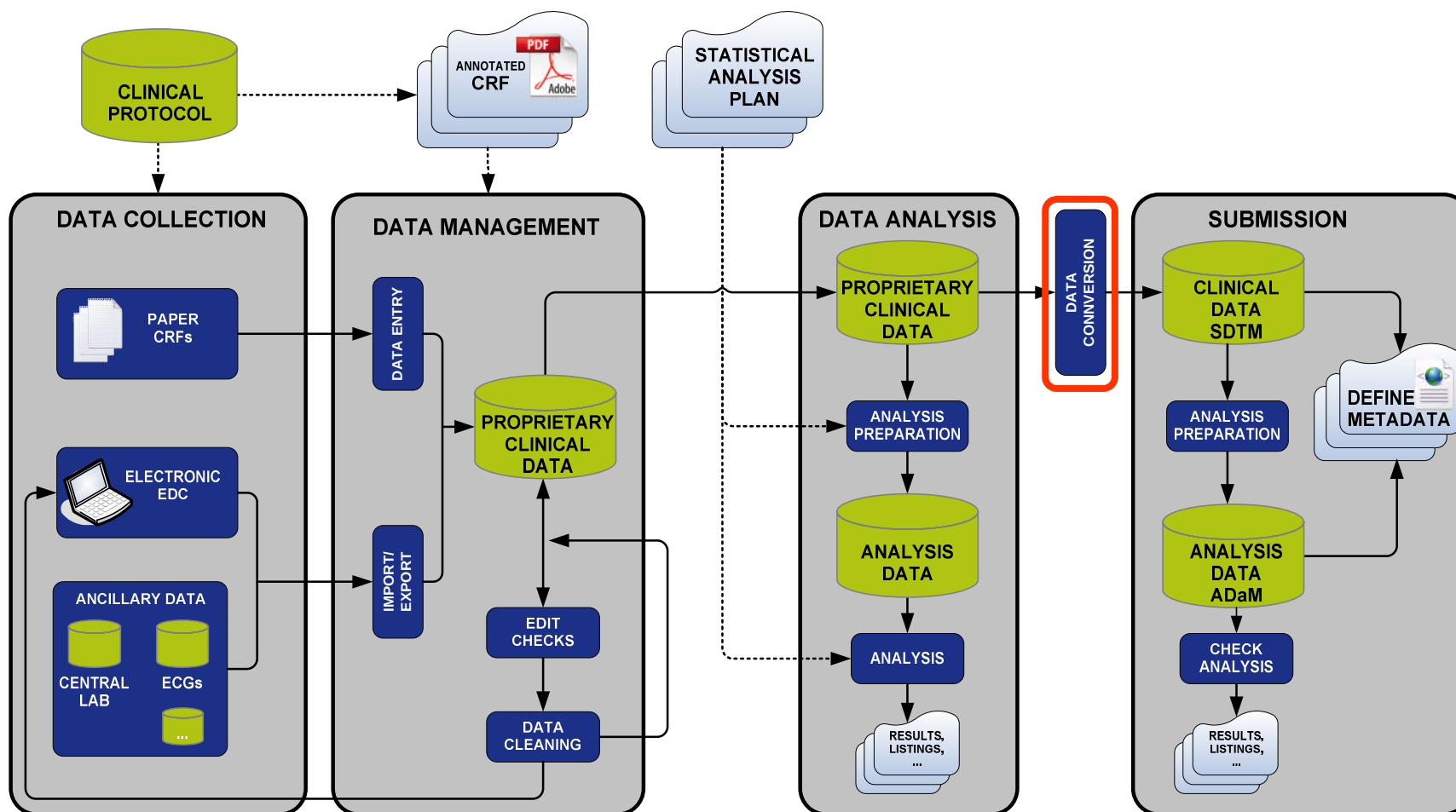
CDISC organization

- Clinical Data Interchange Standards Consortium
 - CDISC started in 1999
 - Is a non-profit organization
 - Data models are created by industry volunteers
 - Mission: to develop & support global, platform-independent data standards that enable information system interoperability to improve medical research & related areas of healthcare

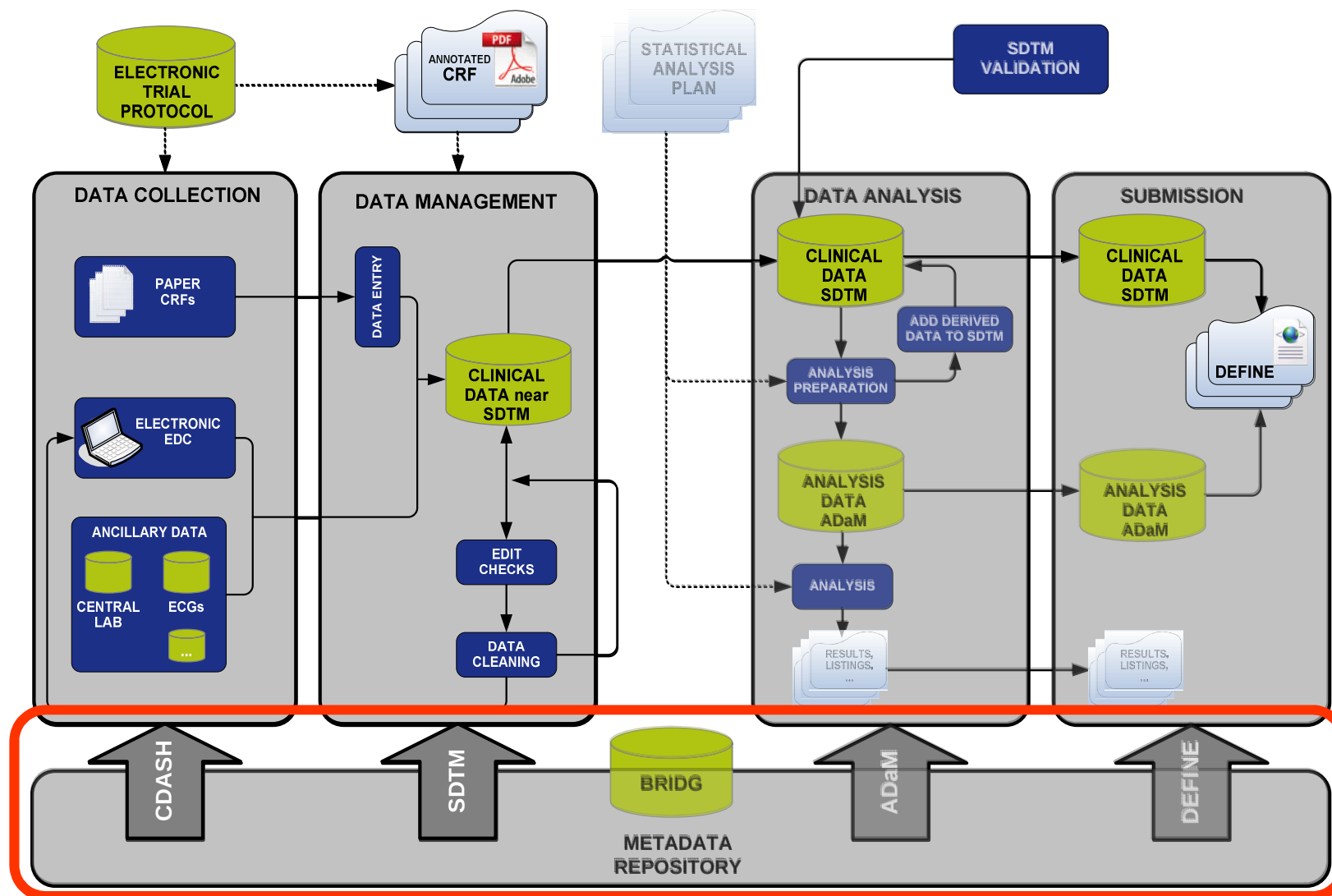
CDISC models

B R I D G / H L 7	Acronym	Title	Use
	CDASH	Clinical Data Acquisition Standards Harmonization	Data Collection
	SDTM	Study Data Tabulation Model	Content
	ADaM	Analysis Dataset Model	Content
	Define.xml CRT-DDS	Case Report Tabulation Data Definition Specification	Metadata
	ODM	Operational Data Model	Format
	LAB	Clinical Laboratory Model	Content
	PRG	Protocol Representation Group	Content

Near Term Approach: Legacy Data Conversion



Longer Term Approach: Global CDISC Integration



Agenda

1

Introduction

2

CDASH/SDTM in OC/RDC

3

SDTM Converter Mechanism

4

Conclusion

CDISC integration objectives

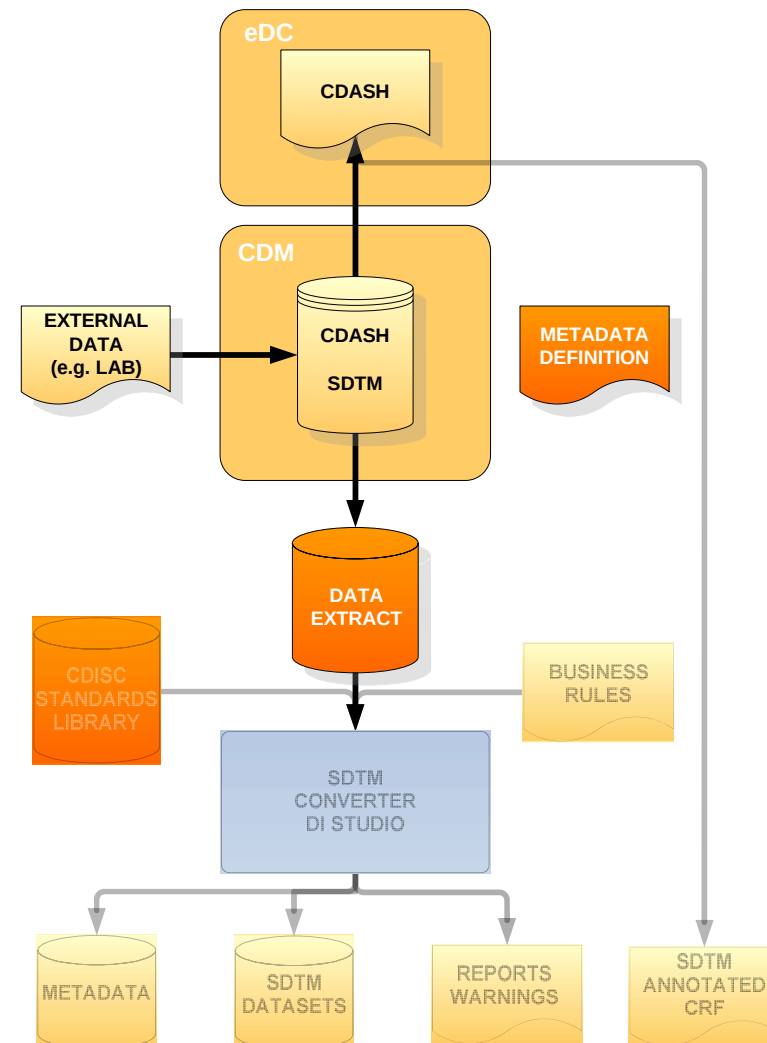
- To integrate global CDISC standards in OC/RDC
 - Use CDISC nomenclature across the CDM system
 - Ensure metadata consistency
 - Use of CDISC controlled terminology
 - Avoid additional data conversion steps
 - Generate SDTM data as of study start
 - Production of real-time SDTM deliverables
 - Facilitate generation of ADaM datasets based on SDTM
 - Allow CDISC consistency checking early in the process

CDISC integration prototype scope

Current scope:	
• Special purpose domains: • General class domains: – Findings: – Events: – Interventions: • SUPPQUAL: • External data upload: • New domain:	CO, DM VS AE EX SUPPAE, -BA, -DM, -EX, -LB LB BA
Outside the scope:	
• Trial Design • RELREC	

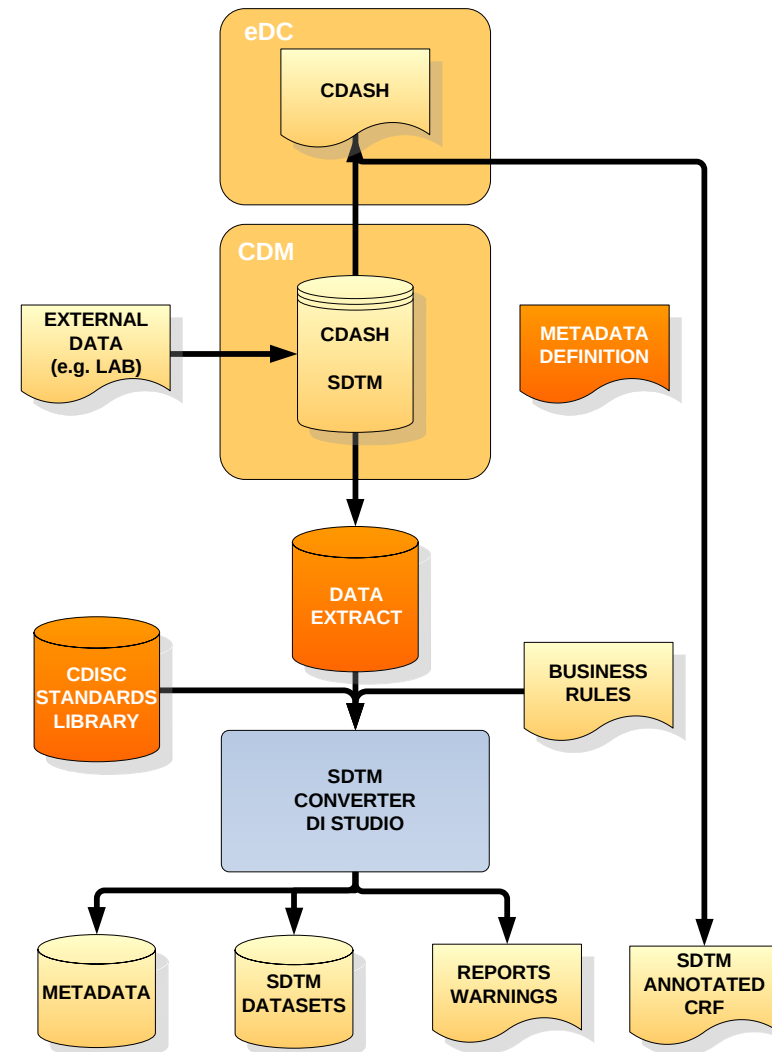
OC/RDC system set-up

- Metadata definition:
 - Create metadata specifications
- OC level:
 - Define DCM, DCI and DCI books
 - External data upload
- RDC level:
 - Patient data entry
- OC level:
 - OC extract creation and output



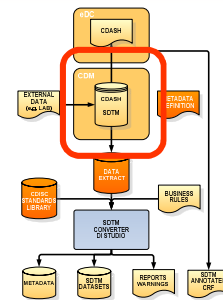
Additional CDISC integration set-up

- Metadata definition:
 - Create metadata specifications
- OC level:
 - Use CDASH/SDTM variables
 - Define DCM, DCI and DCI books
 - External data upload
- RDC level:
 - Patient data entry
- OC level:
 - OC extract creation and output
- SDTM converter level:
 - CDISC (CDASH/SDTM) standards library and business rules
 - SDTM converter mechanism
 - Generation of SDTM datasets, metadata and reports/warnings



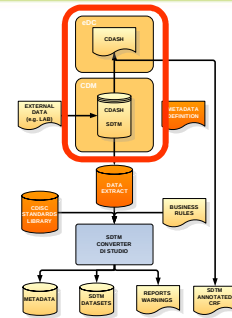
OC level: Use CDASH/SDTM variables

- Define in Global Library:
 - DVGs (SDTM controlled terminology/SAS formats)
 - Questions (CDASH/SDTM variables)
 - Questions groups (SAS datasets)



OC level: Define DCM, DCI & DCI books

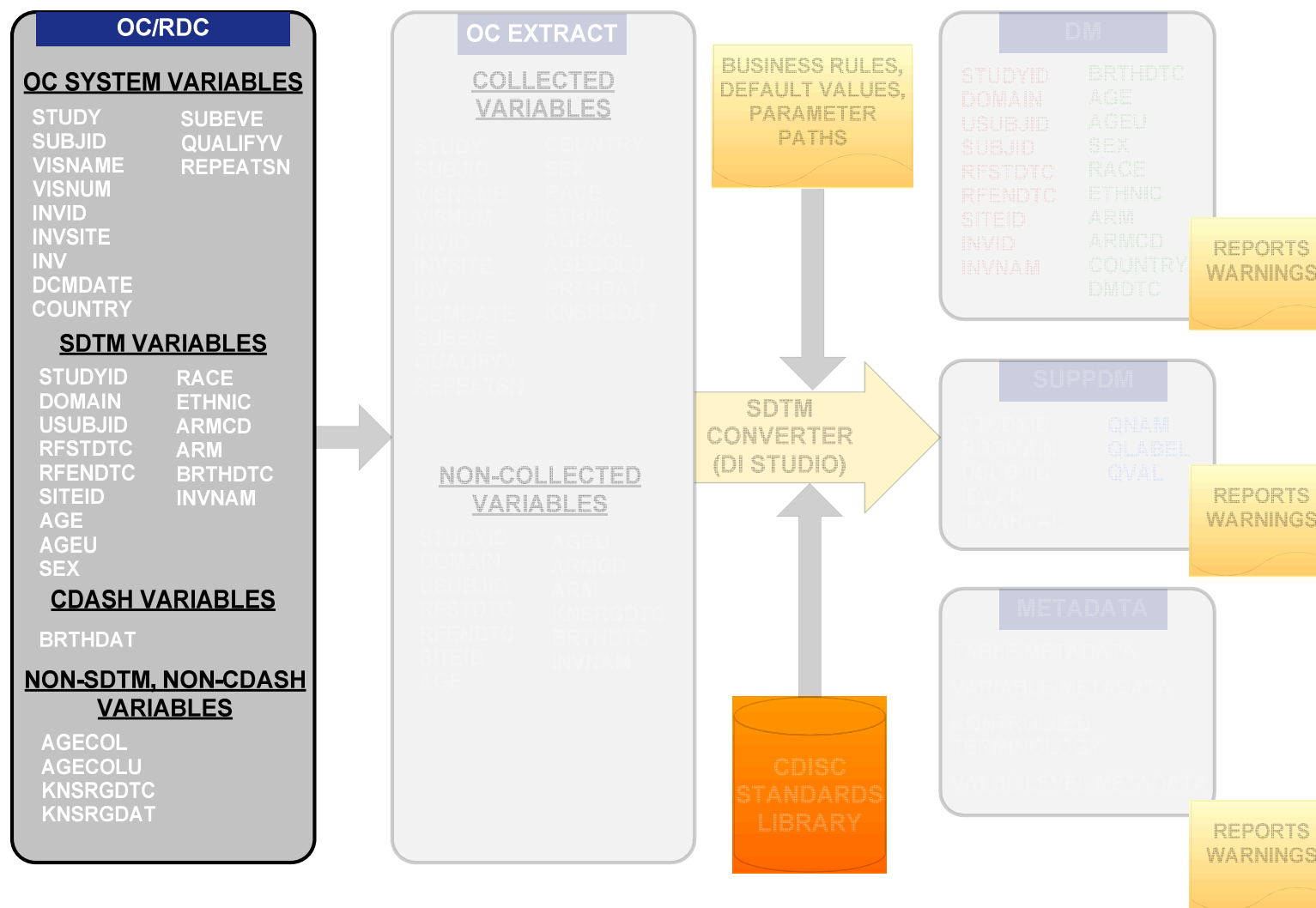
- Define at the study level:
 - DCM (collection of QGs)
 - DCI (collection of DCMs)
 - DCI books
- Create default view definitions (O*C extracts)



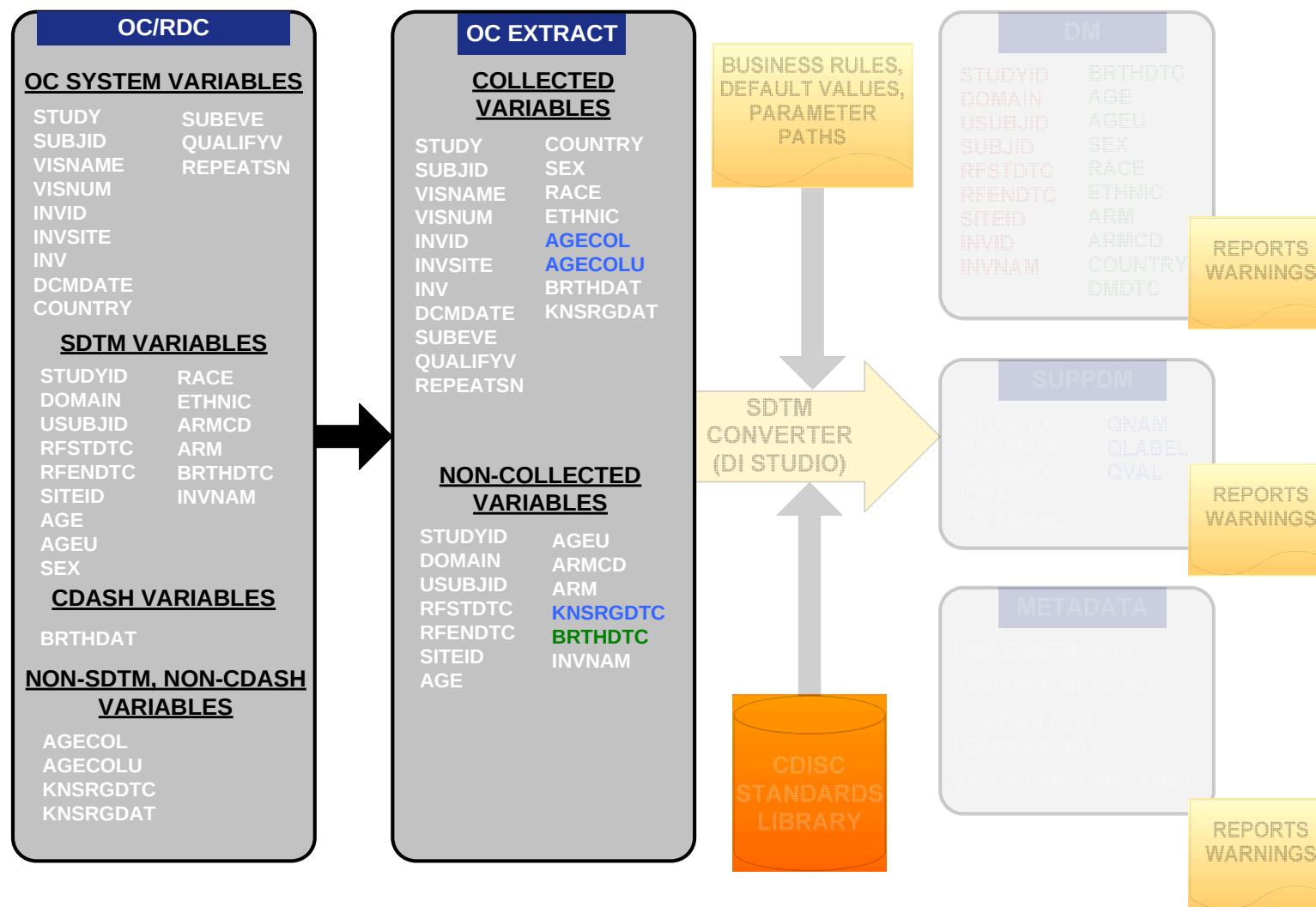
OC level: set-up rules & naming conventions

- Kept to a minimum
- Use CDASH/SDTM variables ONLY except –
 - Rename &DOMAIN.TESTCD → &DOMAIN.TSTC
 - Rename &DOMAIN.ORRES → &DOMAIN.ORES
 - Use of extended attribute DVG_LONG_VALUE in &DOMAIN.TSTC (Testcode) to populate TEST
- All Non CDASH/Non SDTM variables populate SUPP datasets
- Potential use of extended attributes for numeric values (to produce SDTM+)

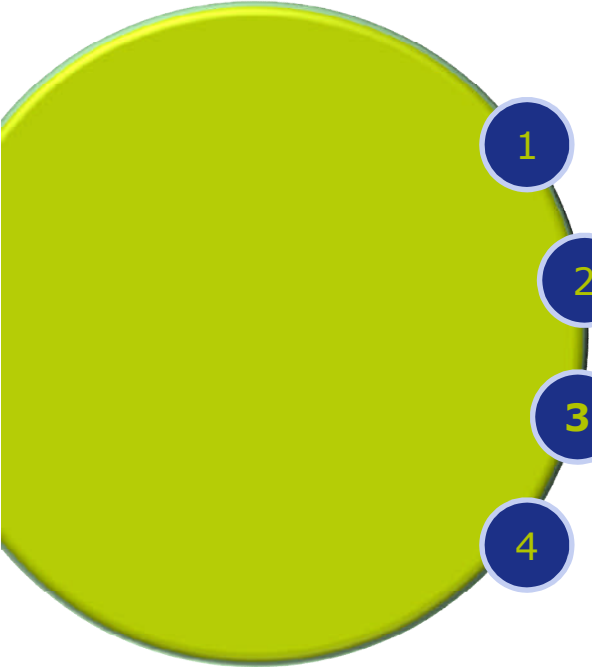
Data flow for e.g. Demography



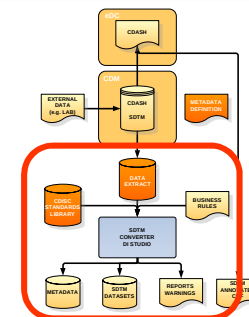
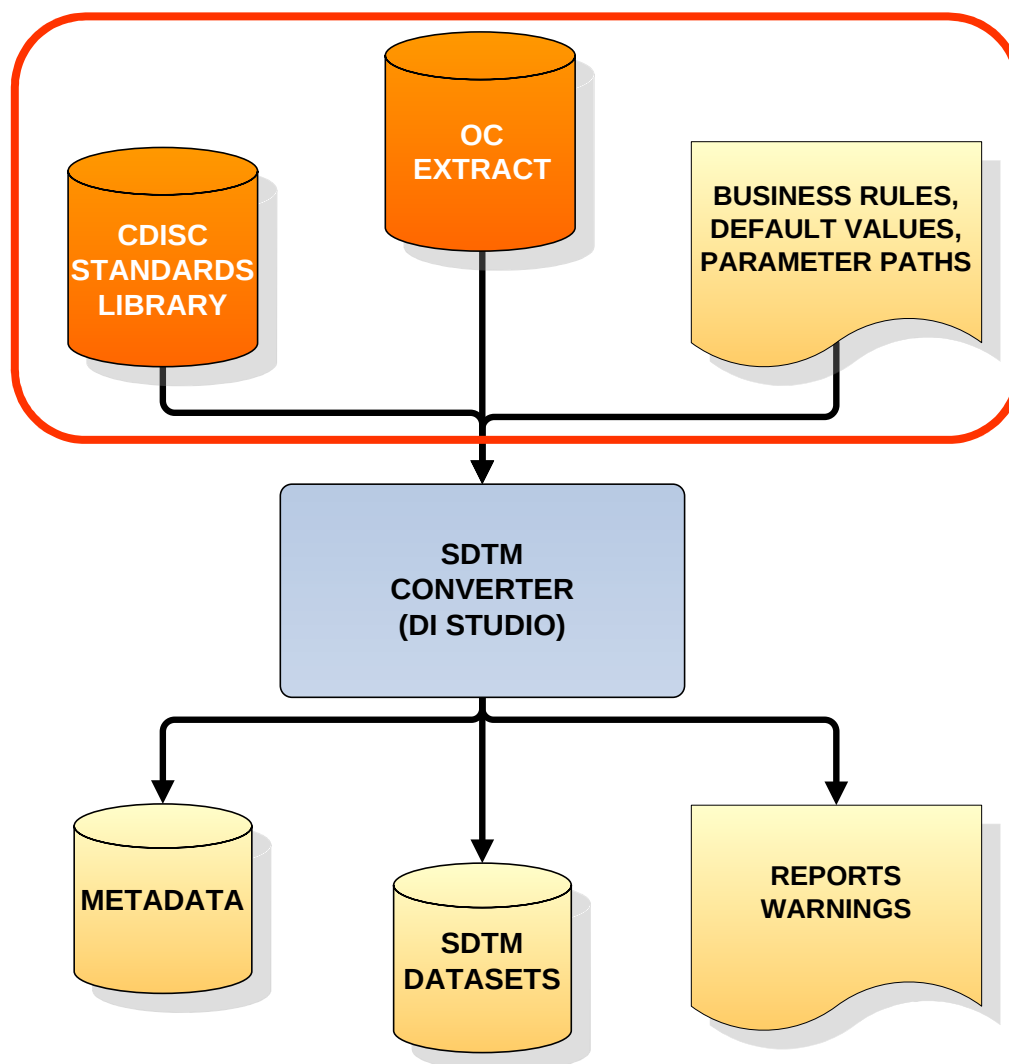
Data flow for e.g. Demography



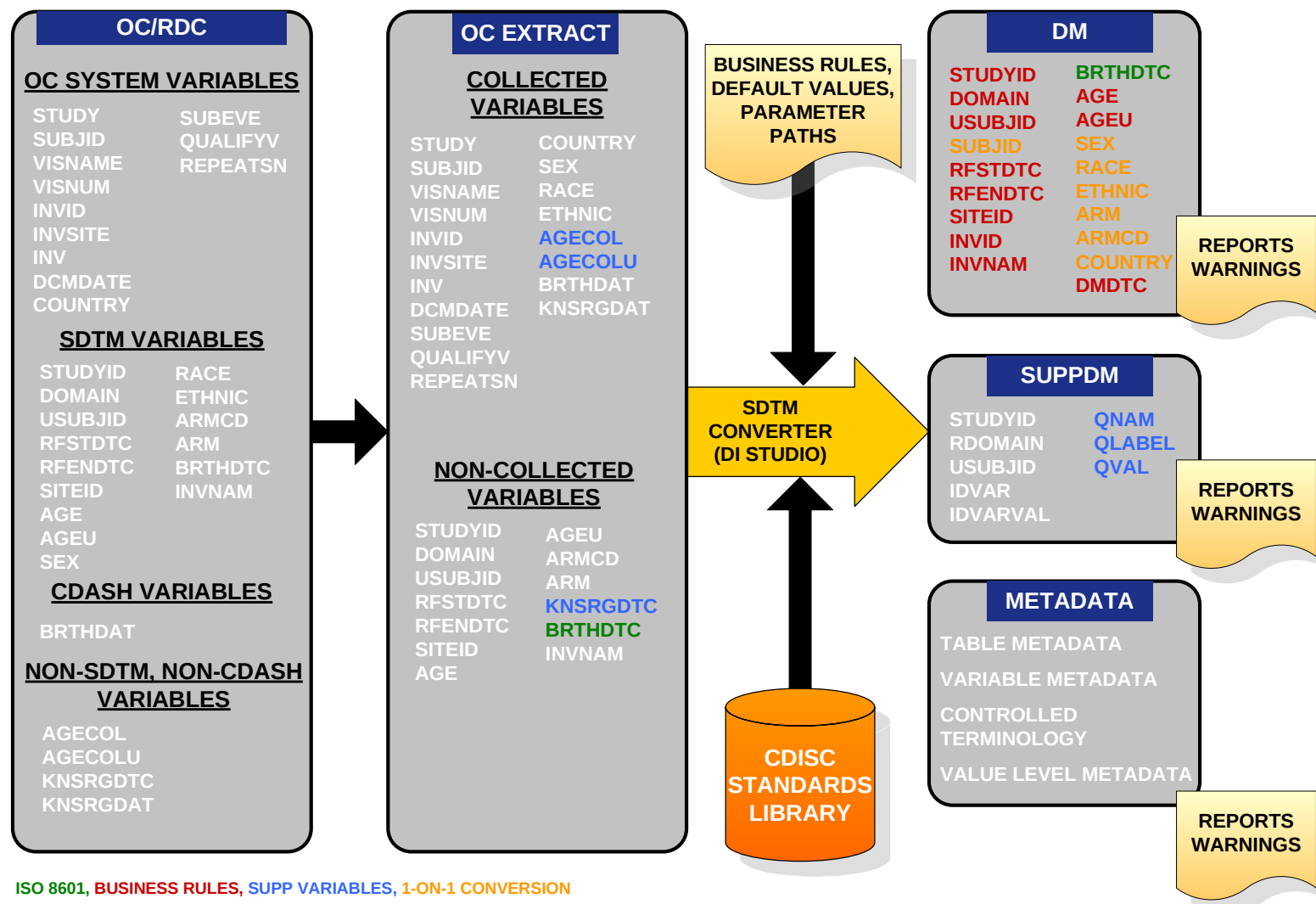
Agenda

- 
- 1 Introduction
 - 2 CDASH/SDTM in OC/RDC
 - 3 SDTM Converter Mechanism**
 - 4 Conclusion

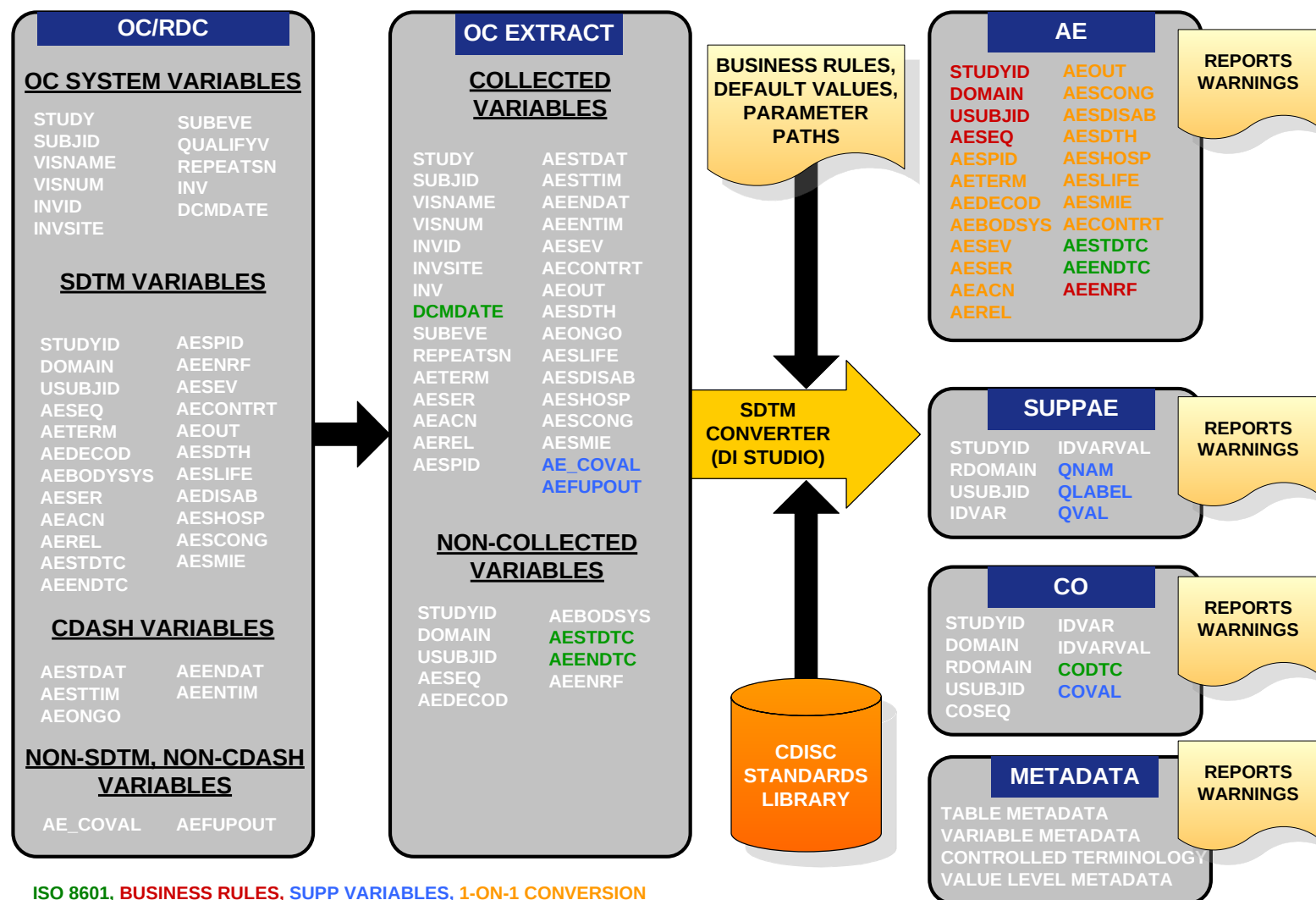
SDTM Conversion – High Level Perspective



Data flow for Demography

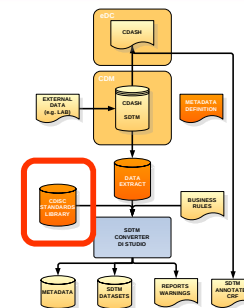


Data flow for Adverse Experiences



SDTM converter input: CDISC standards library

- Used by SDTM converter to interpret standard source or system variables i.e., SDTM or CDASH or OC system



Metadata SDTM variables



Metadata CDASH variables



Metadata OC system variables

- Used by SDTM converter to interpret standard source or system variables i.e., SDTM, CDASH or OC system



SDTM Global Library	Domain	Column	Type	Length	Label	Role	Core	Position	Version
CDASH Global Library	IDENTIFIERS	STUDYID	C	20	Study Identifier	Identifier	Req	1	3.1.1
OC System Variable	IDENTIFIERS	DOMAIN	C	2	Domain Abbreviation	Identifier	Req	2	3.1.1
	IDENTIFIERS	USUBJID	C	20	Unique Subject Ide...	Identifier	Req	3	3.1.1
	IDENTIFIERS	--SEQ	N	8	Sequence Number	Identifier	Req	4	3.1.1
	IDENTIFIERS	--GRPID	C	20	Group ID	Identifier	Perm	5	3.1.1
	IDENTIFIERS	--REFID	C	20	Reference ID	Identifier	Perm	6	3.1.1

	Domain	CDASH Variable
SDTM Global Library		
CDASH Global Library	AE	AEACN
OC System Variable	AE	AEACNOTH
	AE	AEBODSYS
	AE	AECAT
	AE	AECNTRT

SDTM Global Library	Domain	Non Standard Variable
CDASH Global Library	ALL	STUDY
OC System Variable	ALL	VISNAME
	ALL	VISNUM
	ALL	SUBVIS
	ALL	REPEATSN
	ALL	SITFID

SDTM converter input: Short list of business rules

**Defines
generic
SDTM
conversion
choices:**

Business Rules

USUBJID

- ☒ STUDYID || SUBJID
- ☐ STUDYID || SITEID || SUBJID
- ☐ INCLUDE CUSTOM MACRO

RFSTDTC

- ☐ MIN(EX.STDTC)
- ☐ MIN(DS.DSSTDTC) where MIN(VISITNUM)
- ☒ MIN(DS.DSSTDTC) where VISIT = "RANDOMIZATION"
- ☐ INCLUDE CUSTOM MACRO

RFENDTC

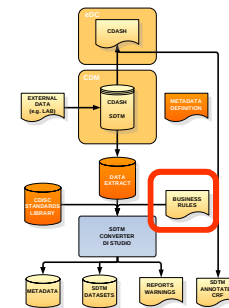
- ☐ MAX(EX.STDTC)
- ☒ MAX(DS.DSSTDTC) where MAX(VISITNUM)
- ☐ MIN(DS.DSSTDTC) where VISIT = "COMPLETED" OR "DISCONTINUED"
- ☐ INCLUDE CUSTOM MACRO

--DTC

- ☒ --DAT/--TIM
- ☐ --DAT
- ☐ VISDAT

NEXT

Business & Decision
Life Sciences



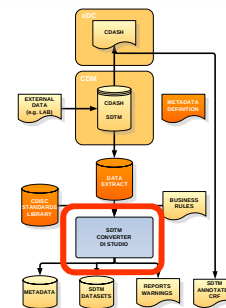
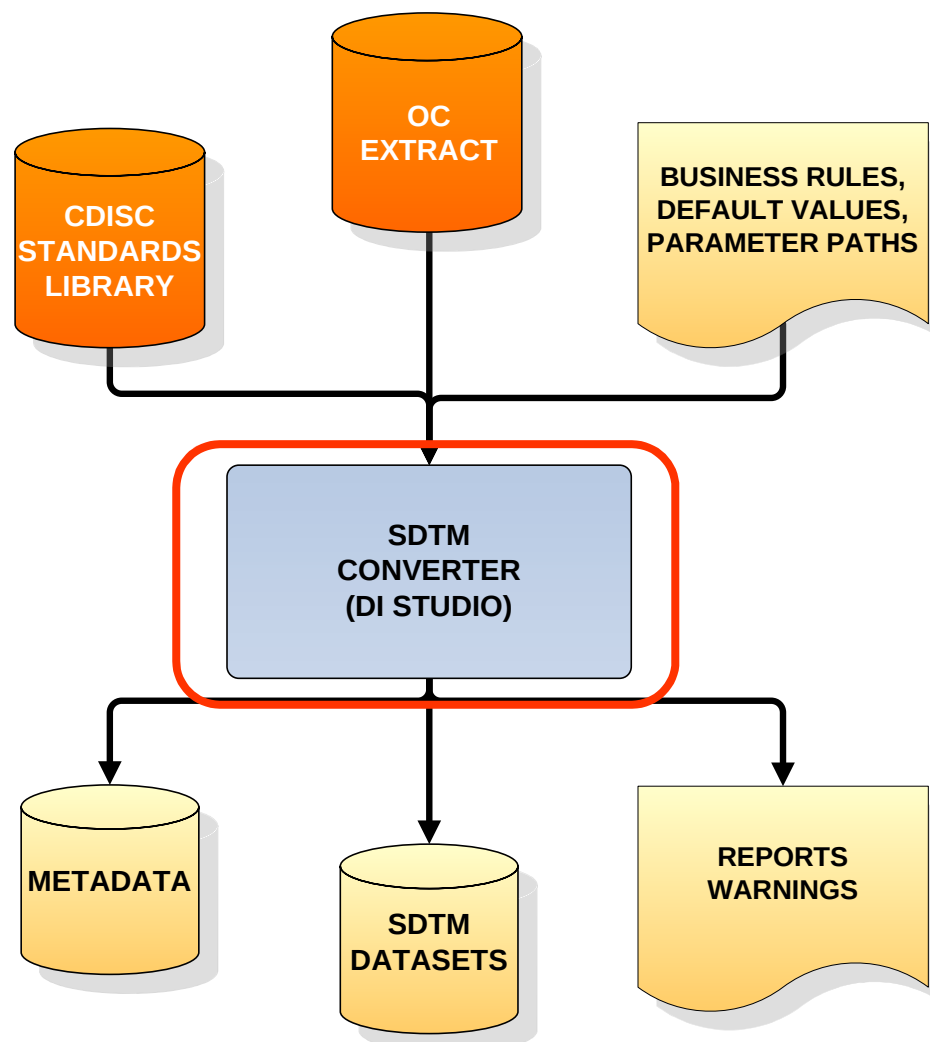
LIFE SCIENCES

SDTM converter input: Short list of default values

- A default (and editable) value table is used by the SDTM converter

DCM Name	QG	Domain	Variable	Default Value
EX1EX1	EX1	EX	EXTRT	"DRUG A"
EX1EX1	EX1	EX	EXDOSE	20
EX2EX2	EX2	EX	EXTRT	"DRUG B"
EX2EX2	EX2	EX	EXDOSE	40
.....				

SDTM conversion – high level perspective



SDTM converter mechanism

- What is it?
 - A mechanism based on Base SAS code (DI Studio)
- List of programs:
 - ISO 8601 data transformation
 - Automatic transpose for findings
 - Global transformations
 - Domain specific transformations
 - Generate SEQ variable
 - Extract TEST from DVG for TESTCD
 - Create Supplementary Datasets
 - Create Comment Dataset
 - Create SDTM Dataset
 - ...

ISO 8601 date transformation

- Rule - Converts variables ending with DAT & TIM in ISO 8601 date format

INPUT DATASET

AESTDAT	AESTTIM	AEENDAT	AEENTIM
20090101	1200	20090601	
200901	1200	20090601	1200
2009			1400

OUTPUT DATASET

AESTDTC	AEENDTC
2009-01-01T12:00	2009-06-01
2009-01--T12:00	2009-09-01T12:00
2009	-----T14:00

Automatic transpose for findings

- Rule - Transposes finding variables TESTCD and corresponding results. Puts records of variables ending with the same number together

INPUT DATASET

VSCAT	VSTSTC1	VSORES1	VSTSTC2	VSORES2
SUPINE	SYSBP	120	DIABP	96
SITTING	SYSBP	110	DIABP	89

OUTPUT DATASET

VSCAT	VSTESTCD	VSORRES
SUPINE	SYSBP	120
SUPINE	DIABP	96
SITTING	SYSBP	110
SITTING	DIABP	89

Global transformations

- Rule - Compute values based on rules in business rule sheet

USUBJID



STUDY||SUBJID



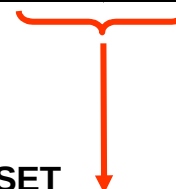
STUDY||SITEID||SUBJID



include custom macro

INPUT DATASET

STUDYID	SUBJID	USUBJID
ABC	1	
ABC	2	
ABC	3	



OUTPUT DATASET

STUDYID	SUBJID	USUBJID
ABC	1	ABC-1
ABC	2	ABC-2
ABC	3	ABC-3

Domain specific transformations

- Rule - Populate variables with values specified in default value sheet

DCM Name	QG	Domain	Variable	Default Value
EX1EX1	EX1	EX	EXTRT	"DRUG A"
EX1EX1	EX1	EX	EXDOSE	20
EX2EX2	EX2	EX	EXTRT	"DRUG B"
EX2EX2	EX2	EX	EXDOSE	40

INPUT DATASET - EX1

USUBJID	EXTRT	EXDOSE	EXDOSU	EXDOSFRM
AAA-1				
AAA-2				

INPUT DATASET - EX2

USUBJID	EXTRT	EXDOSE	EXDOSU	EXDOSFRM
AAA-1				
AAA-2				

OUTPUT DATASET - EX

USUBJID	EXTRT	EXDOSE	EXDOSU	EXDOSFRM
AAA-1	DRUG A	20	mg	INJECTION
AAA-1	DRUG B	40	mg	TABLET
AAA-2	DRUG A	20	mg	INJECTION
AAA-2	DRUG B	40	mg	TABLET

Generate SEQ variable

- Rule – Generate SEQ variable based on Keys specified in Table Level Metadata in the CDISC Standards Library

TABLE LEVEL METADATA

Dataset	Description	Class	Structure	Purpose	Keys
AE	Adverse Events	Events	One record per adverse event per subject	Tabulation	STUDYID, USUBJID, AETERM, AESTDTC
LB	Laboratory Test Results	Findings	One record per lab test per time point per visit per subject	Tabulation	STUDYID, USUBJID, LBTESTCD, VISITNUM, LBDTC
DM	Demographics	Special Purpose	One record per subject	Tabulation	STUDYID, USUBJID

- Rule - For custom domains use the following:
 - Event: --TERM, -- STDTC
 - Intervention: --TRT, --STDTC
 - Finding: --TESTCD, VISITNUM, TPTNUM, --DTC

Extract TEST from DVG for TESTCD

- Rule - Populates TEST by extracting DVG long value of TESTCD

- Vertical structure:

INPUT DATASET	
VSTSTC	VSTSTCL
SYSBP	SYSTOLIC BLOOD PRESSURE
DIABP	DIASTOLIC BLOOD PRESSURE
PULSE	PULSE RATE

OUTPUT DATASET	
VSTESTCD	VSTEST
SYSBP	SYSTOLIC BLOOD PRESSURE
DIABP	DIASTOLIC BLOOD PRESSURE
PULSE	PULSE RATE

- Horizontal structure:

INPUT DATASET			
VSTSTC1	VSTSTC1L	VSTSTC2	VSTSTC2L
SYSBP	SYSTOLIC BLOOD PRESSURE	DIABP	DIASTOLIC BLOOD PRESSURE

OUTPUT DATASET	
VSTESTCD	VSTEST
SYSBP	SYSTOLIC BLOOD PRESSURE
DIABP	DIASTOLIC BLOOD PRESSURE

Create Supplementary Datasets

- Rule - Non-CDASH/Non-SDTM/Non-OC System variables to populate supplementary datasets

INPUT DATASET	
AESEQ	AEFUPOUT LABEL
	AE FOLLOW-UP TO OUTCOME
1	FOLLOW-UP SUFFICIENT
2	LOST TO FOLLOW-UP
3	FOLLOW-UP SUFFICIENT

OUTPUT DATASET					
RDOMAIN	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL
AE	AESEQ	1	AEFUPOUT	AE FOLLOW-UP TO OUTCOME	FOLLOW-UP SUFFICIENT
AE	AESEQ	2	AEFUPOUT	AE FOLLOW-UP TO OUTCOME	LOST TO FOLLOW-UP
AE	AESEQ	3	AEFUPOUT	AE FOLLOW-UP TO OUTCOME	FOLLOW-UP SUFFICIENT

Create Comment Dataset

- Rule - Collect comments in OC in &DOMAIN._COVAL variable

INPUT DATASET

AESEQ	AE_COVAL
1	Comment1
2	Comment2
3	Comment3

OUTPUT DATASET

RDOMAIN	IDVAR	IDVARVAL	COVAL
AE	AESEQ	1	Comment1
AE	AESEQ	2	Comment2
AE	AESEQ	3	Comment3

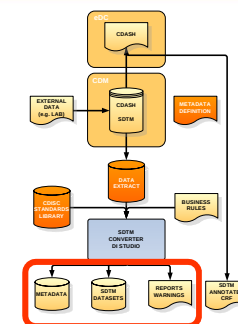
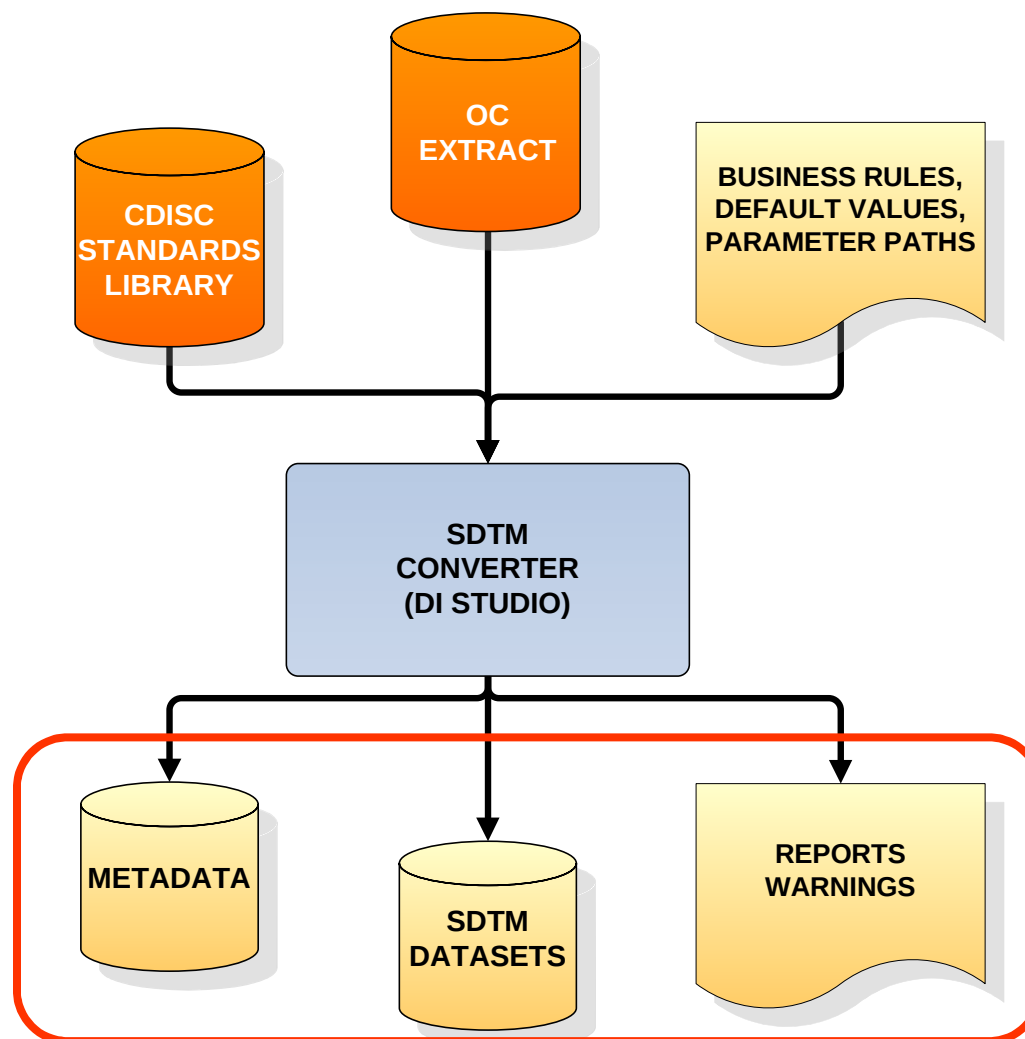
Create SDTM Dataset

- Rule – Use variable level attributes from the CDISC standards library

VARIABLE LEVEL METADATA

SASTableName	SASColumnName	SASColumnType	SASColumnLength	SASLabel
AE	STUDYID	C	20	Study Identifier
AE	DOMAIN	C	2	Domain Abbreviation
AE	USUBJID	C	20	Unique Subject Identifier

SDTM conversion – high level perspective



SDTM converter output: Reports and Warnings

- Data profile reports of the SDTM generated datasets.

DATAPATH.DM

Created: 19JUN09:14:43:12
obs : 3
Keys :

#	Variable	Label	Type	Length	+	Contents
1	STUDYID	Study Identifier	CHAR	20	+	1 distinct values, 0 missing values, all shown
2	DOMAIN	Domain Abbreviation	CHAR	2	+	1 distinct values, 0 missing values, all shown
3	USUBJID	Unique Subject Identifier	CHAR	20	+	3 distinct values, 0 missing values, all shown
4	SUBJID	Subject Identifier for the Study	CHAR	20	+	3 distinct values, 0 missing values, all shown
5	RFSTDTC	Subject Reference Start Date/Time	CHAR	20	+	3 distinct values, 0 missing values, all shown
6	RFENDTC	Subject Reference End Date/Time	CHAR	20	+	3 distinct values, 0 missing values, all shown
7	SITEID	Study Site Identifier	CHAR	20	+	1 distinct values, 0 missing values, all shown
8	INVID	Investigator Identifier	CHAR	20	+	1 distinct values, 0 missing values, all shown
9	INVNAM	Investigator Name	CHAR	20	+	1 distinct values, 0 missing values, all shown
10	BRTHDTC	Date/Time of Birth	CHAR	20	+	3 distinct values, 0 missing values, all shown
11	AGE	Age in AGEU at RFSTDTC	NUM	8	+	3 distinct values, 0 missing values, all shown Mean: 57 Minimum: 44 Maximum: 79 Missing values: 0
12	AGEU	Age Units	CHAR	10	+	1 distinct values, 0 missing values, all shown
13	SEX	Sex	CHAR	1	+	2 distinct values, 0 missing values, all shown
14	RACE	Race	CHAR	20	+	3 distinct values, 0 missing values, all shown
15	ETHNIC	Ethnicity	CHAR	12	+	2 distinct values, 0 missing values, all shown
16	ARMCD	Planned Arm Code	CHAR	8	+	All values are missing
17	ARM	Description of Planned Arm	CHAR	20	+	All values are missing
18	COUNTRY	Country	CHAR	20	+	1 distinct values, 0 missing values, all shown
19	DMDTC	Date/Time of Collection	CHAR	20	+	3 distinct values, 0 missing values, all shown

SDTM converter output: Reports and Warnings

- Warning reports: includes basic checks for conformance with the CDISC domain models:
 - Including all required and expected variables as columns in a domain
 - Using CDISC-specified standard variable names
 - Using CDISC-specified data types for all variables
 - Etc.
- These conformance checks are a subset of the complete set of CDISC SDTM compliance checks

SDTM converter output: Metadata

Table Metadata

Dataset Name
Dataset Label
Class
Structure
Purpose
Keys

Variable Metadata

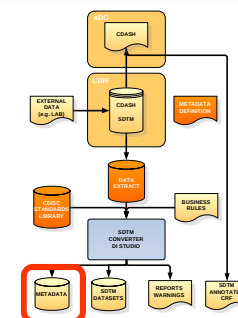
Variable Name
Variable Label
Type
Length
Controlled Terminology
Role

Controlled Terminology

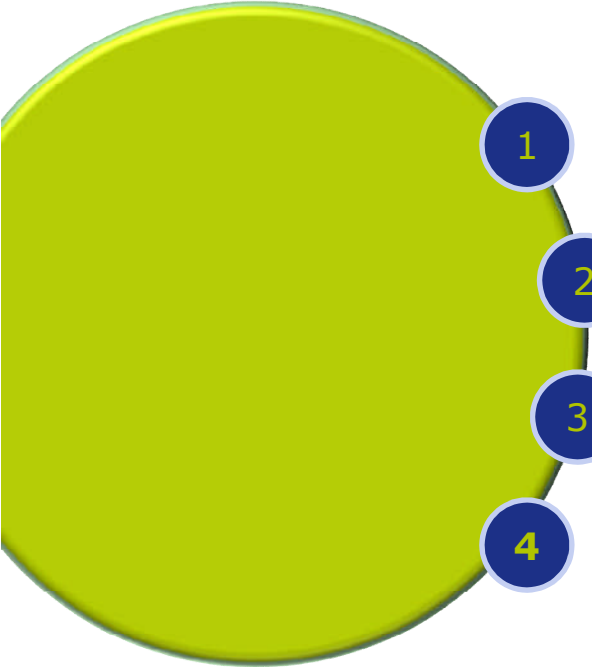
Code Value
Code Text

Value Level Metadata

Source Variable
Value
Label
Type
Controlled Terminology



Agenda

- 
- 1 Introduction
 - 2 CDASH/SDTM in OC/RDC
 - 3 SDTM Converter Mechanism
 - 4 **Conclusion**

Conclusion

- Maximize use of CDASH and SDTM variables
- Use CDISC controlled terminology
- Through smart set-up in OC, automate SDTM generation through a generic SDTM converter mechanism
 - No SDTM conversion required in OC/RDC; occurs outside OC/RDC
 - No additional programming is required within OC
 - Real-time SDTM datasets and metadata, ready for analysis



LIFE SCIENCES

Thank you for
your attention

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®