CDISC 360:

Evolving our standards towards end to end automation

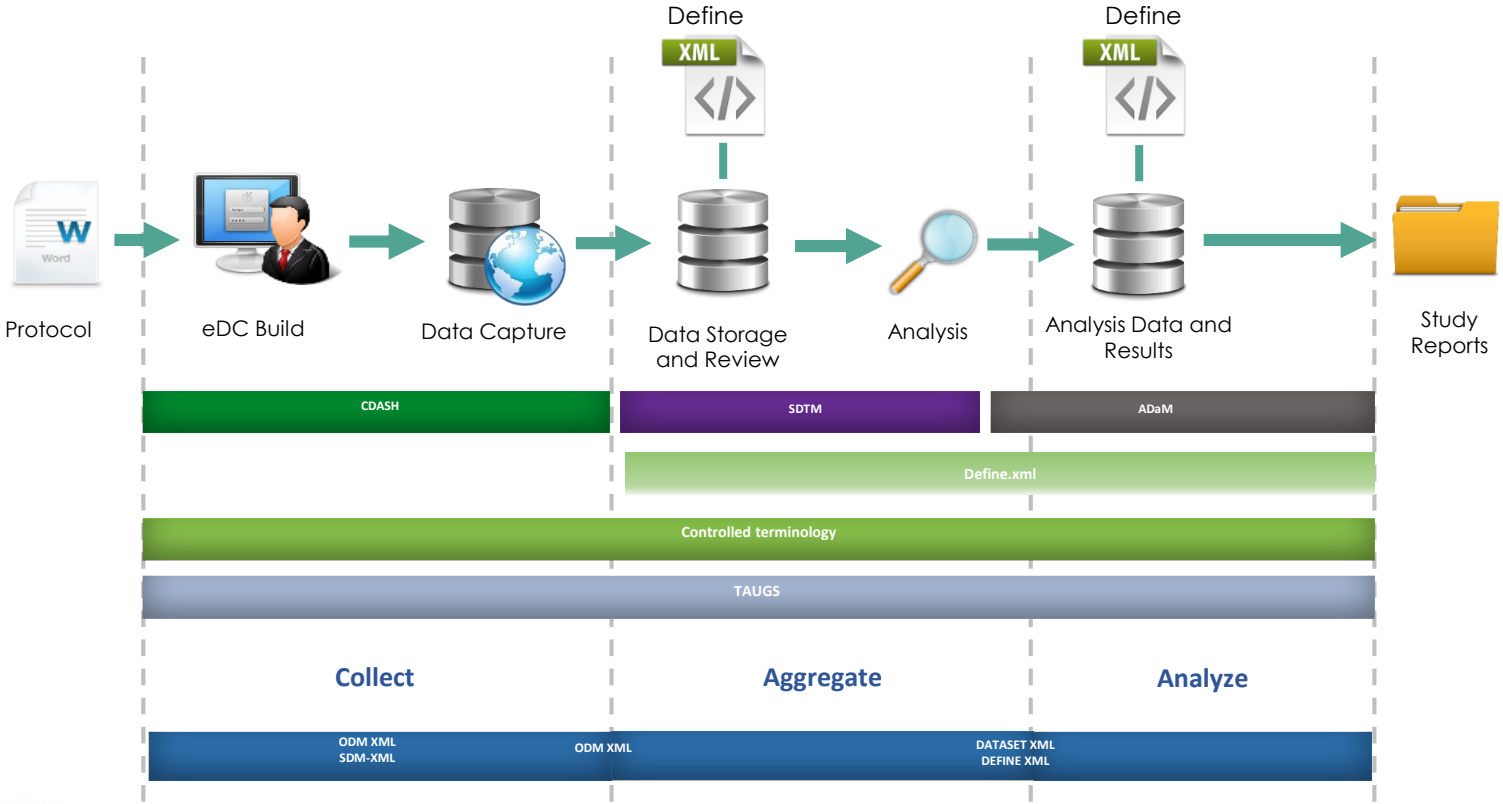
Peter van Reusel
Sam Hume
Barry Cohen

Agenda

1. Where are we today
2. Approach
3. Logistics
4. Relationship to Other Initiatives
5. Expected outcomes

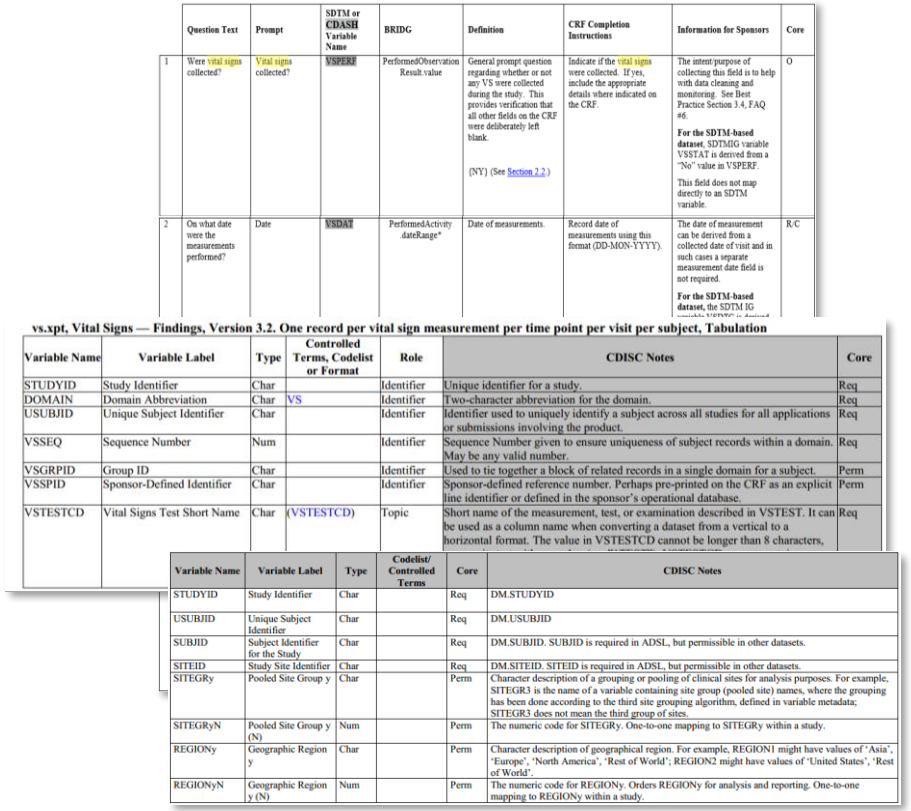
1. Where are we today

Today we are here



Defined structures

- CDISC Foundational models provide much needed structure
  - Normative Content
  - 2 dimensional (tables, columns)
  - Standard to represent data
- The information itself is not defined
  - We do not need new structures
  - We need to define
    - Entities
    - Semantics (meaning)
    - Relationships between information
    - Rules in the data lifecycle



	Question Text	Prompt	SDTM or CDASH Variable Name	BRIDG	Definition	CRF Completion Instructions	Information for Sponsors	Core
1	Were vital signs collected?	Vital signs collected?	VSPERF	PerformedObservation Result.value	General prompt question regarding whether or not any VS were collected during the study. This provides verification that all other fields on the CRF were deliberately left blank. (NY) (See Section 2.2.)	Indicate if the vital signs were collected. If yes, include the appropriate details where indicated on the CRF.	The intent/purpose of collecting this field is to help with data cleaning and monitoring. See Best Practice Section 3.4, FAQ #6. For the SDTM-based dataset, SDTMIG variable VSSTAT is derived from a "No" value in VSPERF. This field does not map directly to an SDTM variable.	O
2	On what date were the measurements performed?	Date	VSDAT	PerformedActivity .dateRange*	Date of measurements.	Record date of measurements using this format (DD-MON-YYYY).	The date of measurement can be derived from a collected date of visit and in such cases a separate measurement date field is not required. For the SDTM-based dataset, the SDTM IG	R/C

vs.xpt, Vital Signs — Findings, Version 3.2. One record per vital sign measurement per time point per visit per subject, Tabulation

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format	Role	CDISC Notes	Core
STUDYID	Study Identifier	Char		Identifier	Unique identifier for a study.	Req
DOMAIN	Domain Abbreviation	Char	VS	Identifier	Two-character abbreviation for the domain.	Req
USUBJID	Unique Subject Identifier	Char		Identifier	Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product.	Req
VSSEQ	Sequence Number	Num		Identifier	Sequence Number given to ensure uniqueness of subject records within a domain. May be any valid number.	Req
VSGRPID	Group ID	Char		Identifier	Used to tie together a block of related records in a single domain for a subject.	Perm
VSSPID	Sponsor-Defined Identifier	Char		Identifier	Sponsor-defined reference number. Perhaps pre-printed on the CRF as an explicit line identifier or defined in the sponsor's operational database.	Perm
VSTESTCD	Vital Signs Test Short Name	Char	(VSTESTCD)	Topic	Short name of the measurement, test, or examination described in VSTEST. It can be used as a column name when converting a dataset from a vertical to a horizontal format. The value in VSTESTCD cannot be longer than 8 characters,	Req

Variable Name	Variable Label	Type	Codelist/ Controlled Terms	Core	CDISC Notes
STUDYID	Study Identifier	Char		Req	DM.STUDYID
USUBJID	Unique Subject Identifier	Char		Req	DM.USUBJID
SUBJID	Subject Identifier for the Study	Char		Req	DM.SUBJID. SUBJID is required in ADSL, but permissible in other datasets.
SITEID	Study Site Identifier	Char		Req	DM.SITEID. SITEID is required in ADSL, but permissible in other datasets.
SITEGRy	Pooled Site Group y	Char		Perm	Character description of a grouping or pooling of clinical sites for analysis purposes. For example, SITEGR3 is the name of a variable containing site group (pooled site) names, where the grouping has been done according to the third site grouping algorithm, defined in variable metadata; SITEGR3 does not mean the third group of sites.
SITEGRyN	Pooled Site Group y (N)	Num		Perm	The numeric code for SITEGRy. One-to-one mapping to SITEGRy within a study.
REGIONy	Geographic Region y	Char		Perm	Character description of geographical region. For example, REGION1 might have values of 'Asia', 'Europe', 'North America', 'Rest of World'; REGION2 might have values of 'United States', 'Rest of World'.
REGIONyN	Geographic Region y (N)	Num		Perm	The numeric code for REGIONy. Orders REGIONy for analysis and reporting. One-to-one mapping to REGIONy within a study.

Why Change?

Industry needs are maturing

- Machine-readable standards
- Move beyond normative structural description of data
- Provide semantic relations between data – add meaning
- Add process metadata to enable end-to-end automation
- We want non-standard experts to use our standards

2. Approach

What is the CDISC 360 Project?

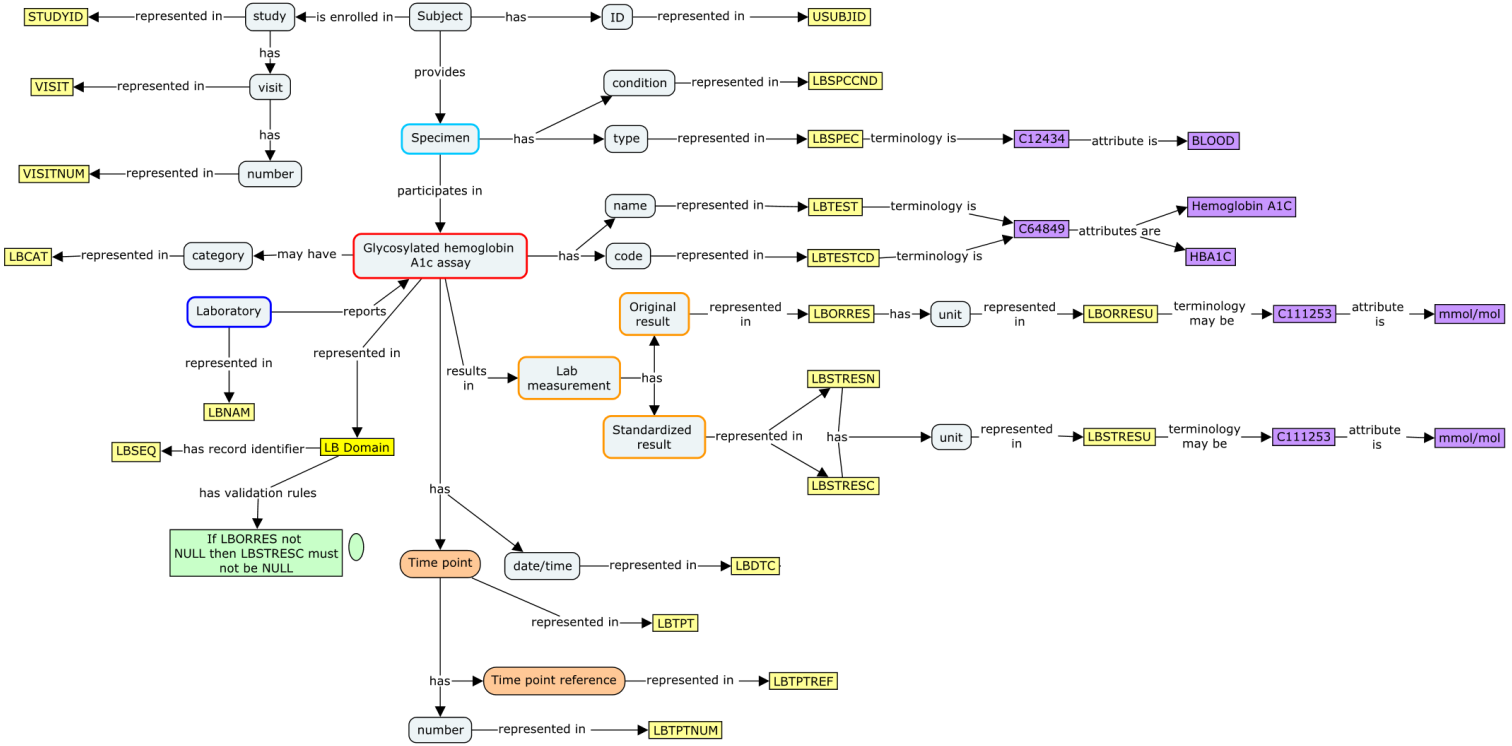
Adding a conceptual layer to standards

- Create and store standards as concepts which create meaning between data
- A serious attempt to store and use data standards as linked metadata
- Add computer readable process metadata which enables end to end automation
- Evolve from normative to informative standards
- CDISC 360 will develop concept-based standard definitions, and test and demonstrate end-to-end automation of study specification and data processing

➔ *Test and demonstrate, but **not building software***

cdisc

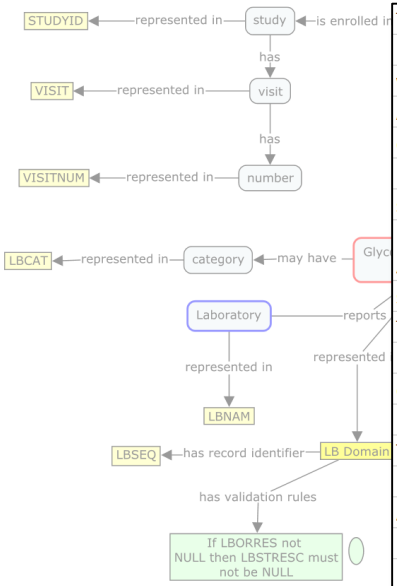
Biomedical Concept



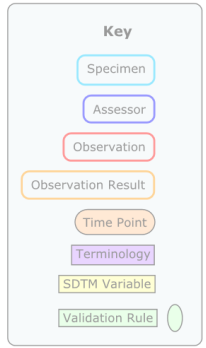
→ *Data is expressed through the CDISC Foundational Models*

→ *Can be mapped to BRIDG Reference Model*

Biomedical Concept

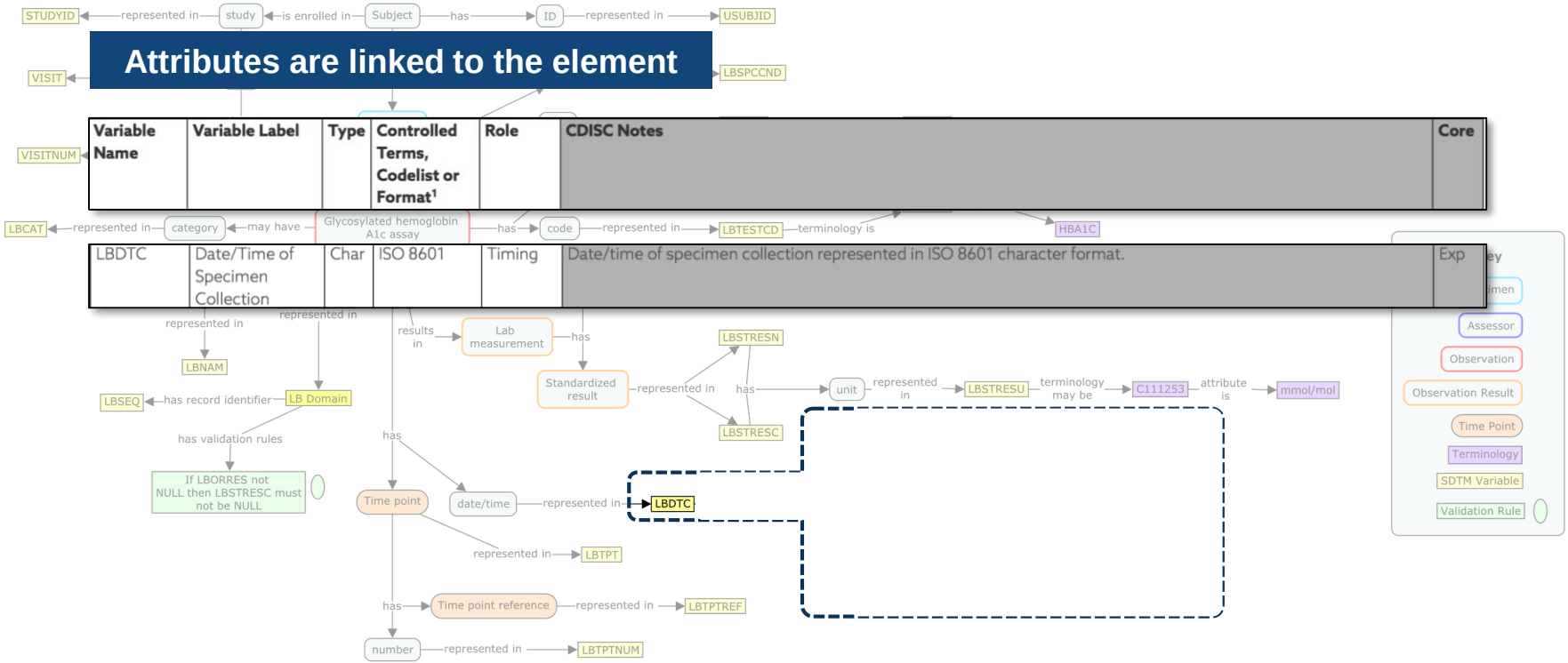


Treatment variable	is	ADHBA1C.TRTP
population flag	based on	TA.ARM/TA.ARMCD
version	is	9.2
Analysis Data Set	includes	ADHBA1C.ITTFL
derivation	uses	Visit Windowing
LB Tabulation Dataset	includes	LB.SPCCND
source	is	LB.TEST/LB.TESTCD
ADHBA1C.AVAL	hasderivation Set to “Y” when ADHBA1C.AVISIT = “Baseline”	
statistical method	is	repeated measures analysis
TV Trial Design Dataset	has	name
LOCF	uses	ADHBA1C.AVAL
derivation	uses	LB.STRESN
DM dataset	includes	TA.ARM/TA.ARMCD
TA.ARM/TA.ARMCD	is either	drug A
Display (TFL)	has	identifier
Analysis Data Set	includes	ADHBA1C.AVAL
metadata	includes	Analysis Purpose
metadata	includes	Documentation
Analysis Data Set	includes	ADHBA1C.DTYPE
ADHBA1C.DTYPE	has	derivation
Display (TFL)	includes one to many	result
metadata	includes	Selection criterion
LB Tabulation Dataset	includes	LB.VISIT

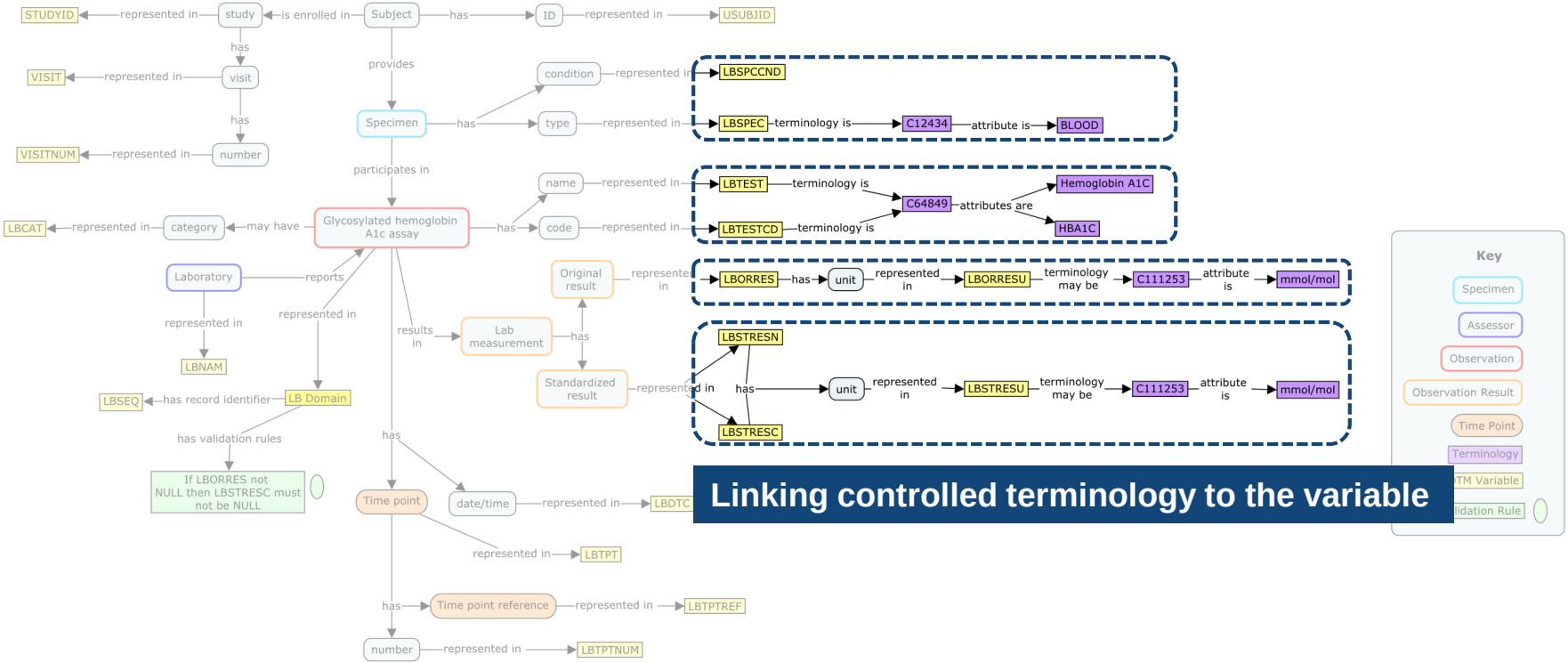


Triple Store

Biomedical Concept

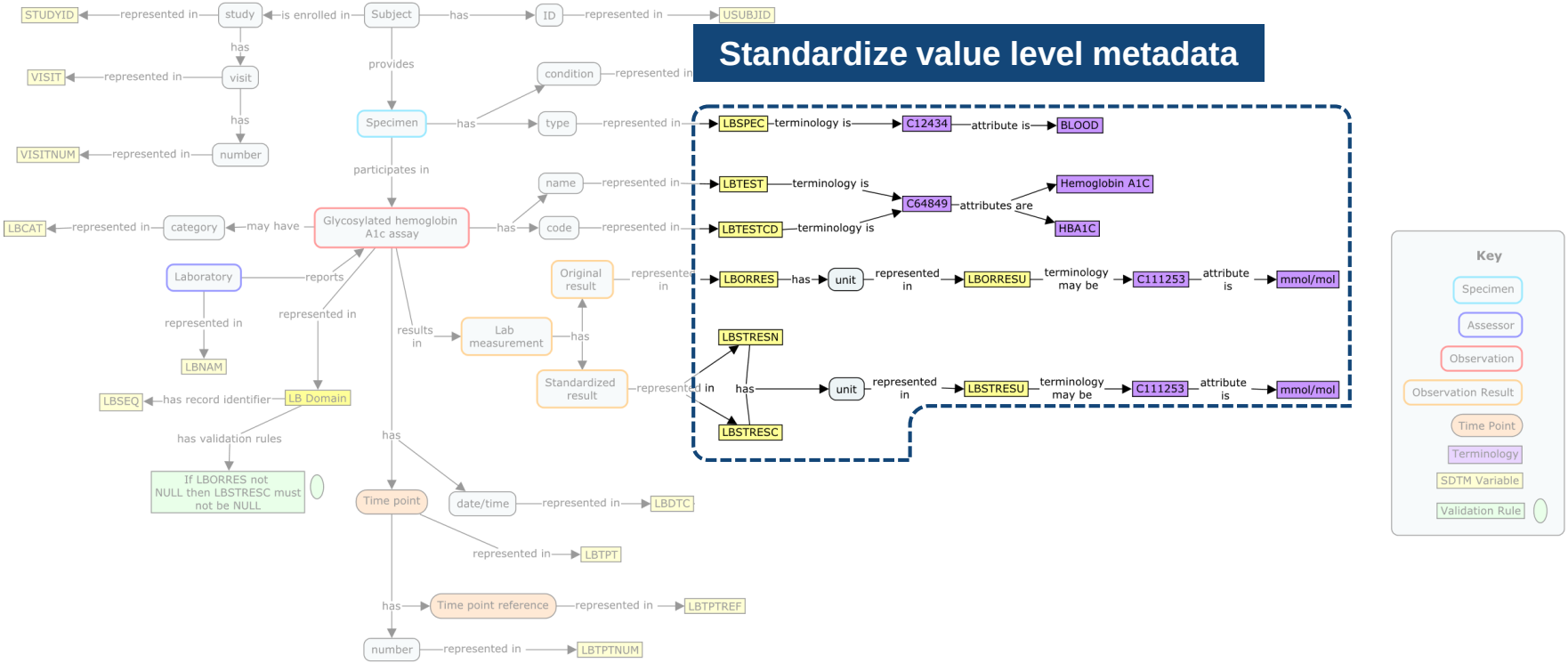


Biomedical Concept

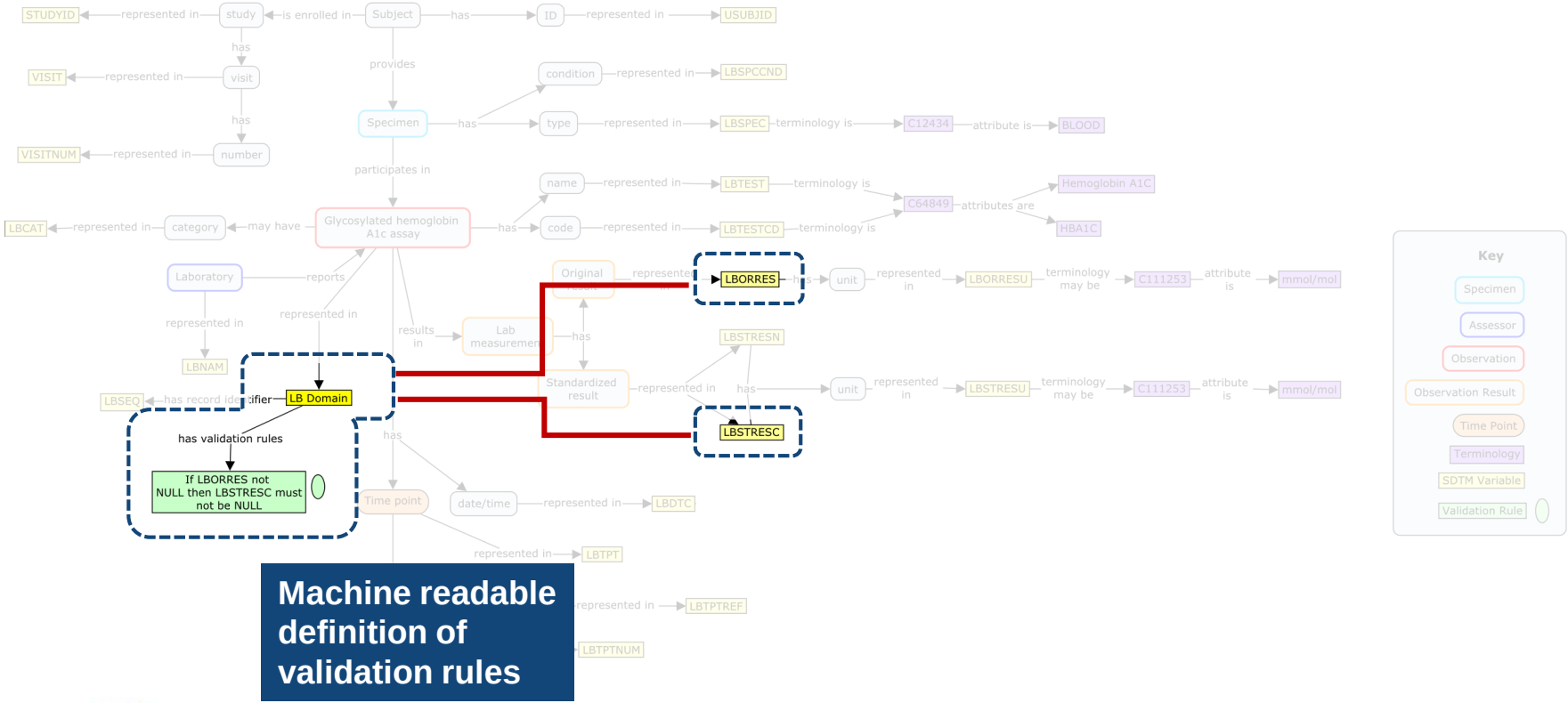


Linking controlled terminology to the variable

Biomedical Concept

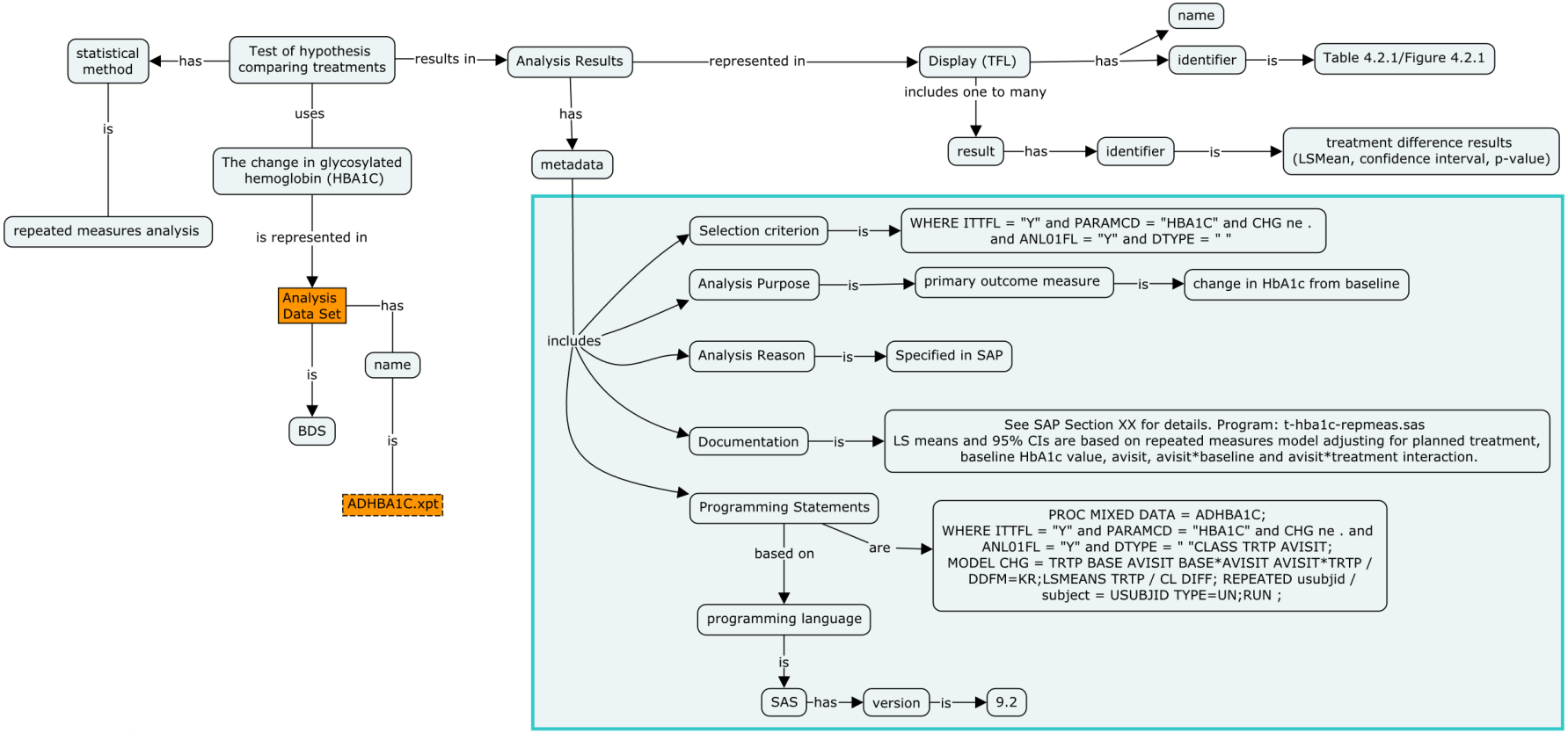


Biomedical Concept

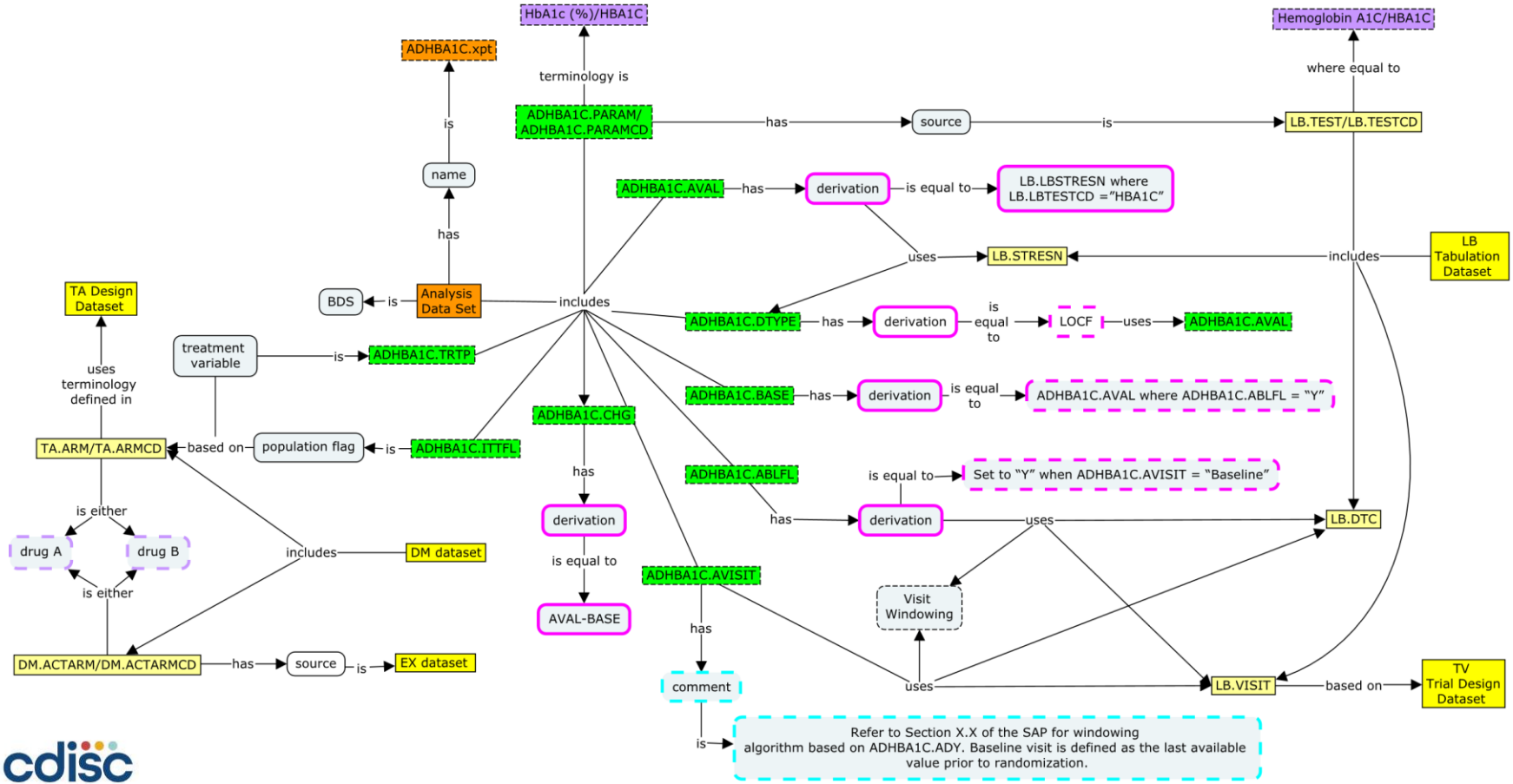


Machine readable definition of validation rules

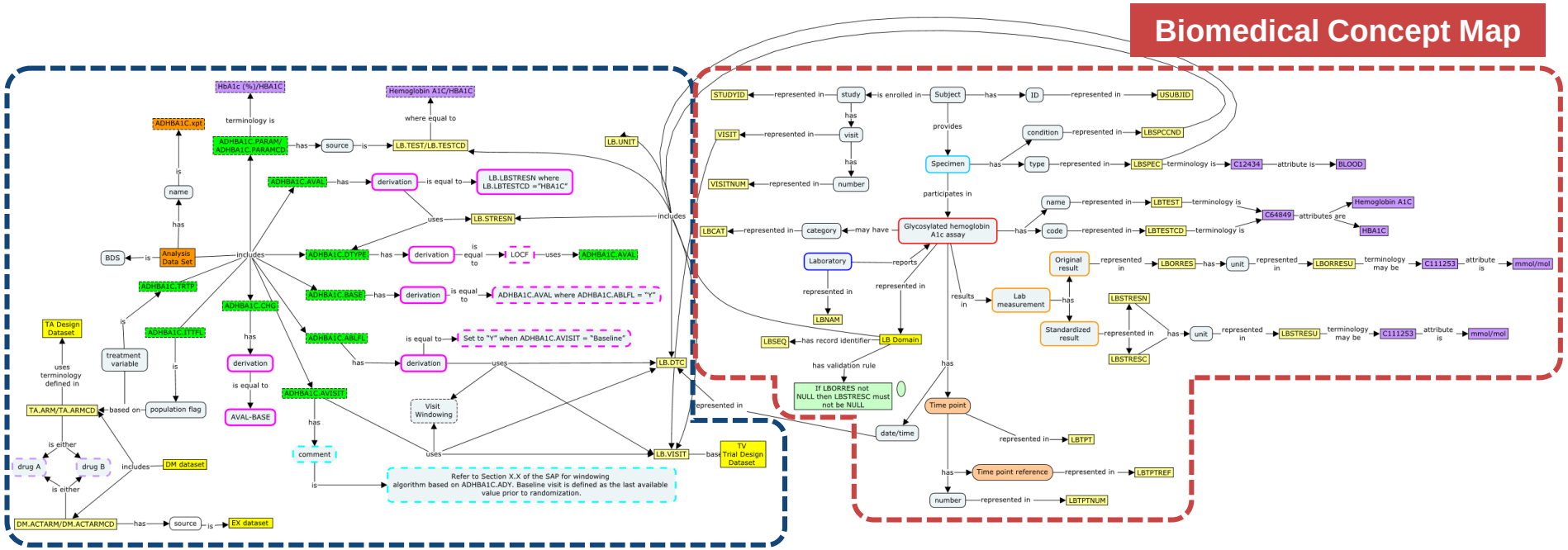
Analysis Result



Analysis Concept



One Model



→ The Biomedical Concept and Analysis Concept are ***ONE MODEL***

The Power of a Conceptuel Model for Data Standards

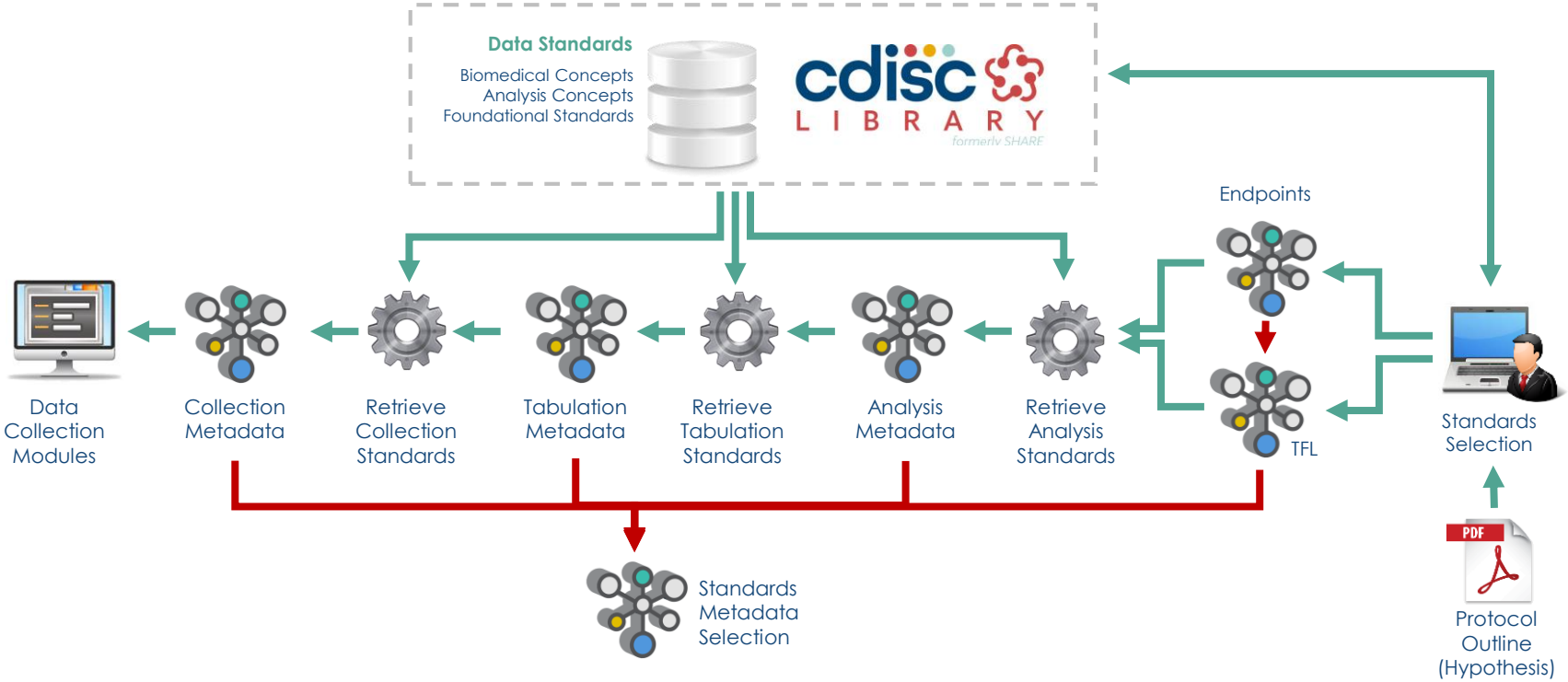
- Linking controlled terminology to the variable – standardize value level metadata
- Machine readable definition of validation rules
- Linking derivations and algorithms to variable(s)
  - Include process metadata (ETL instructions)
- Possibility to standardize Analysis outputs and Collection instruments
  - Combining layout, variables, process information together
- Link Analysis Concepts to Biomedical Concepts
  - Choose an analysis and automatically obtain all related end-to-end metadata

→ ***All of the above:*** *enables automation, increase confidence in results, true analysis traceability*

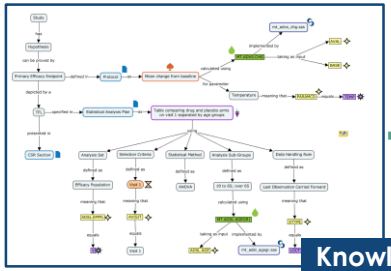
cdisc

Use Case 1 : End to Start Specification

Selecting standards concepts and linked metadata needed for a study



CDISC Library API extension



Knowledge graph

LOAD

Biomedical Concepts
Analysis Concepts
Foundational Standards

cdisc LIBRARY
formerly SHARE

Call

API

Extend Current Set of APIs

Return

API metadata Return

PAA	BBB	X-3	Y+H	DEF	1+2	…

- Dataset metadata
- Concept metadata
- Concept relations
- Controlled terminology
- Process metadata
- Configuration Defaults and options

cdisc

DISCLAIMER NOTE

The following is not a software demonstration

Sole purpose is to illustrate how data standards can enable tools

Welcome

Login: CDarwin

Password: **********

SIGN IN >>

CDASH

DOMAIN
CDASH Variable
Value Level Metadata
Controlled Terminology

SDTM

DOMAIN
SDTM Variable
Value Level Metadata
Controlled Terminology
Computational algorithm

ADaM

DOMAIN
ADaM Variable
ADaM Parameters
Controlled Terminology
Computational algorithm

AE	EX	CM	MH	VS	LB
Adverse Events	Exposure	Concomittant Medication	Medical History	Vital Signs	Laboratory Test Results

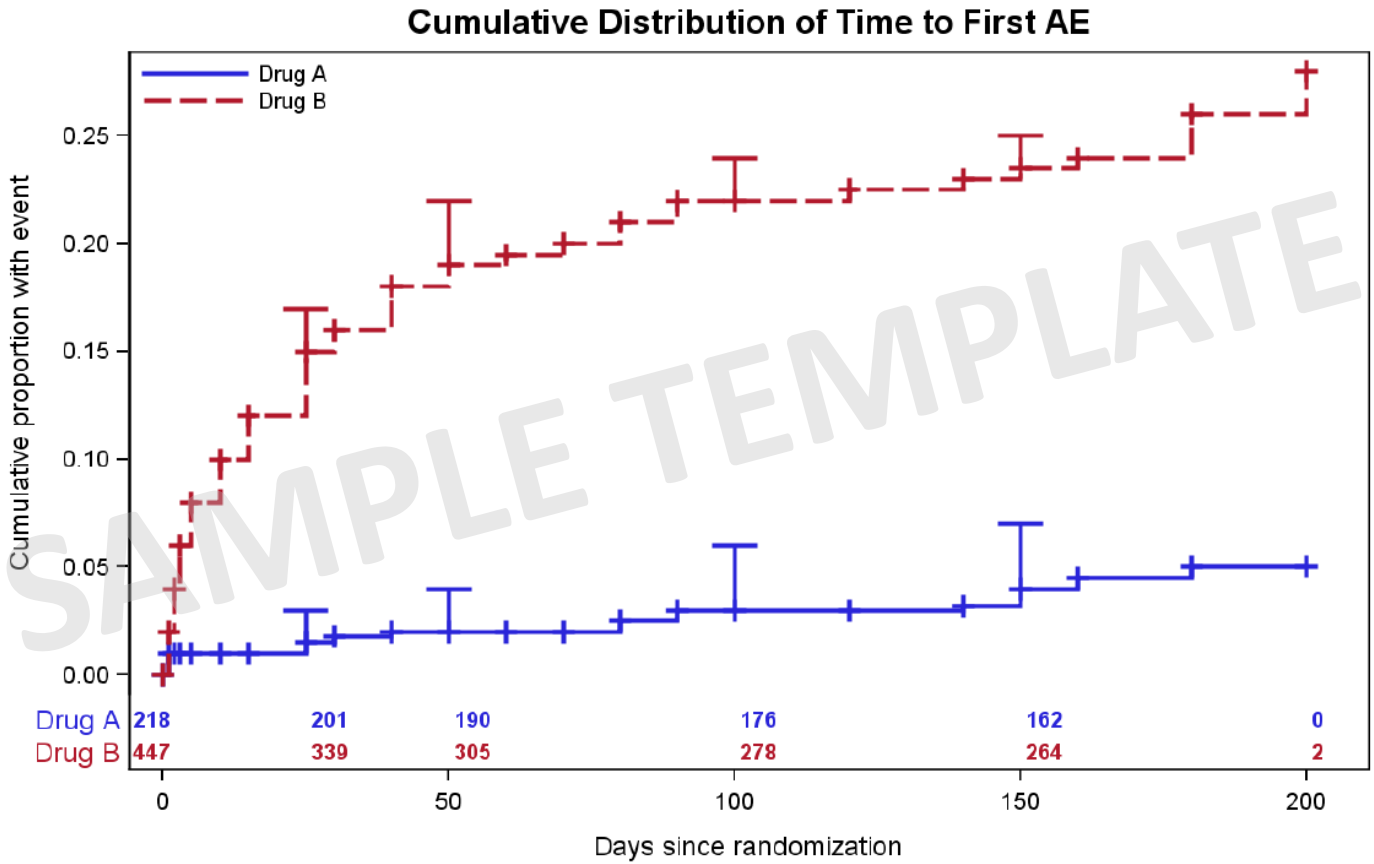
CDASH
DOMAIN
CDASH Variable
Value Level Metadata
Controlled Terminology

SDTM
DOMAIN
SDTM Variable
Value Level Metadata
Controlled Terminology
Computational algorithm

ADaM
DOMAIN
ADaM Variable
ADaM Parameters
Controlled Terminology
Computational algorithm

Figures

- Graphical Approaches to the Analysis of Safety Data from Clinical Trials". Amit, et. al.
- From "Graphical Approaches to the Analysis of Safety Data from Clinical Trials". Amit, et. al.
- Mean Change from Baseline in QTc by time and treatment.
- Distribution of ASAT by time and treatment
- Distribution of maximum LFT values by treatment.
- Panel of LFT shift from baseline to maximum by treatment
- LFT Patient profiles
- Most Frequent On Therapy Adverse Events
- Cumulative distribution (with SEs) of time to first AE of special interest


Cumulative Distribution of Time to First AE
Drug A
Drug B
Cumulative proportion with event
0.25
0.20
0.15
0.10
0.05
0.00
Drug A 218 201 190 176 162 0
Drug B 447 339 305 278 264 2
0
50
100
150
200
Days since randomization
SAMPLE TEMPLATE

CDASH

DOMAIN
CDASH Variable
Value Level Metadata
Controlled Terminology

SDTM

DOMAIN
SDTM Variable
Value Level Metadata
Controlled Terminology
Computational algorithm

ADaM

DOMAIN
ADaM Variable
ADaM Parameters
Controlled Terminology
Computational algorithm

AE Adverse Events | EX Exposure | CM Concomittant Medication | MH Medical History | VS Vital Signs | LB Laboratory Test Results

Log out | Favorites | Previous | Repository | Workspace | Selection

Enter your search here

SELECTION | DATA COLLECTION | ANALYSIS

CDASH
DOMAIN
CDASH Variable
Value Level Metadata
Controlled Terminology

SDTM
DOMAIN
SDTM Variable
Value Level Metadata
Controlled Terminology
Computational algorithm

ADaM
DOMAIN
ADaM Variable
ADaM Parameters
Controlled Terminology
Computational algorithm

Listings

- Listing 2.4 Current Cancer History – All Treated Patients Experiencing Critical Events
- Listing 2.5 Prior and Concomitant Medication – All Treated Patients Experiencing Critical Events
- Listing 2.6 Physical Examination at Screening – All Treated Patients Experiencing Critical Events
- Listing 3.1 Reference Chemotherapy and Concomitant Chemotherapies – All Treated Patients Experiencing ..
- Listing 4.1 Adverse Event Listing. All Pre-Treatment Adverse Events – All Treated Patients Experiencing ...
- Listing 4.2 Adverse Event Listing. Treatment Emergent Adverse Events – All Treated Patients Experiencing ...
- Listing 4.3 Adverse Event Listing. Serious Treatment Emergent Adverse Events – All Treated Patients ..
- Listing 4.4 Adverse Event Listing. Serious Treatment Emergent Adverse Events Related To Study Drug ...
- Listing 4.5 Adverse Event Listing. Serious Treatment Emergent Adverse Events Related To Treatment

x

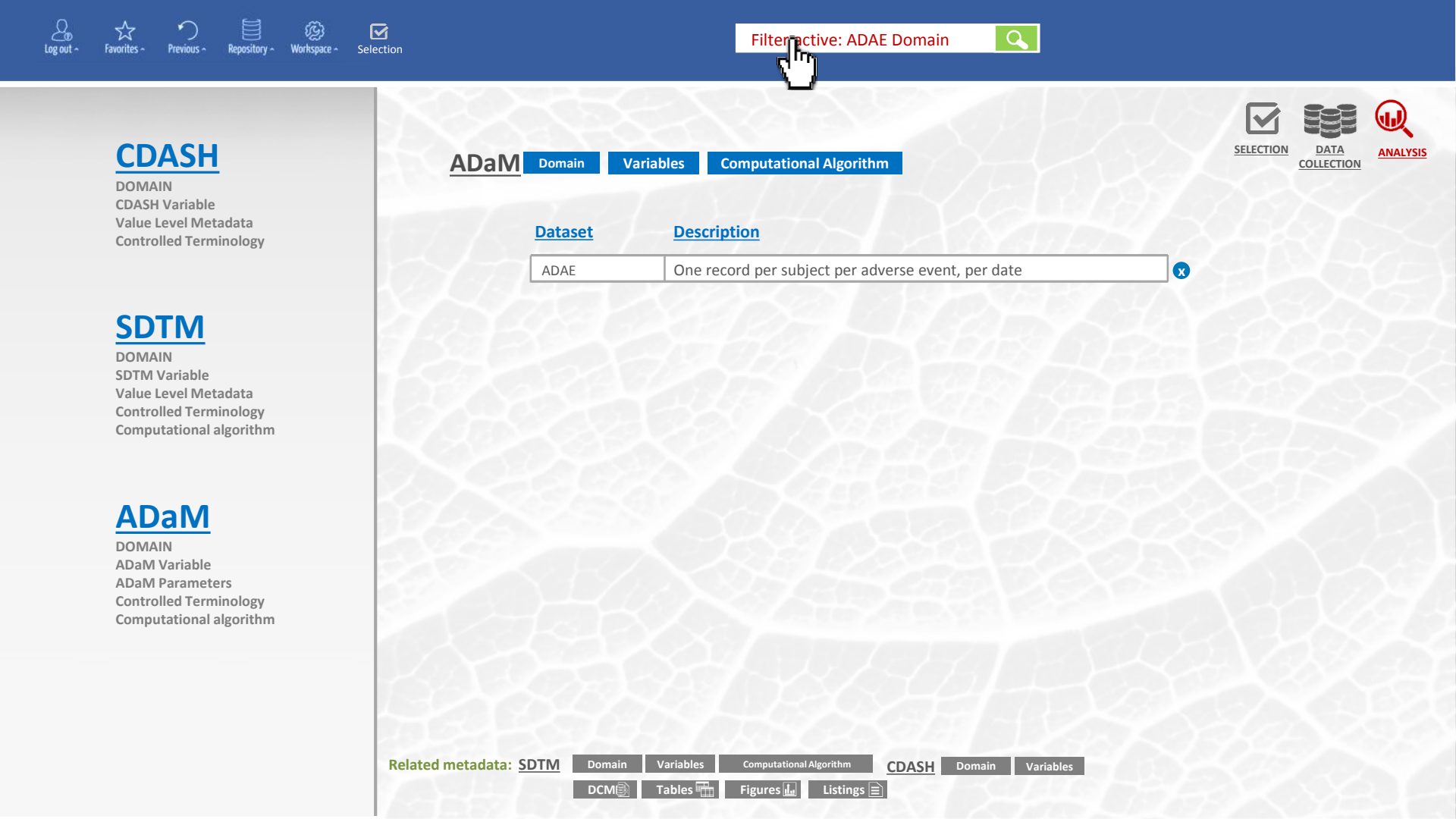
Listing 4.5 Adverse Event Listing. Serious Treatment Emergent Adverse Events Related To "Treatment" – All Treated Patients Experiencing Critical Events

Country	Site/ Patient ID	AE Verbatim Term MedDRA SOC Name MedDRA Preferred Term	Start Date/Time Stop Date/Time Duration (Days/Hours)	Day of onset	Occurrence	Intensity CTC grade	Relationship to Dexamethasone	Action Taken	Outcome
xxxxxxxx	xx/xxx	XXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXX	DDMMMYYYY/HH:MM DDMMMYYYY/HH:MM xxxxx/xxxxx	xx	Intermittent	Grade X	Possibly	None	Resolved

Analysis dataset: ADAE.SAS7BDAT ddmmmyyyy hh:mm
Note: Critical events are defined as: Serious Adverse Events (extracted from the clinical database reconciled with the safety database), Suspected Unexpected Serious Adverse Reactions (extracted from the safety database), wrong study medication used (patients who received a wrong medication kit by mistake in one cycle, resulting in the administration of drug from both treatment groups during the study).
Note: "Treatment" related adverse events are adverse events with a missing relationship to "Treatment" or assessed by the Investigator as definite, probable, possible or unassessable.
Program: <DIRECTORY PATH>\XXXXXX.sas; Date & Time program was run: ddmmmyyyy hh:mm; Date & Time analysis dataset was run: ddmmmyyyy hh:mm

SELECT Close

cdisc


Log out
Favorites
Previous
Repository
Workspace
Selection
Filter active: ADAE Domain
SELECTION
DATA COLLECTION
ANALYSIS
CDASH
DOMAIN
CDASH Variable
Value Level Metadata
Controlled Terminology
SDTM
DOMAIN
SDTM Variable
Value Level Metadata
Controlled Terminology
Computational algorithm
ADaM
DOMAIN
ADaM Variable
ADaM Parameters
Controlled Terminology
Computational algorithm
ADaM
Domain
Variables
Computational Algorithm
Dataset
Description
ADAE
One record per subject per adverse event, per date
Related metadata:
SDTM
Domain
Variables
Computational Algorithm
DCM
Tables
Figures
Listings
CDASH
Domain
Variables

CDASH
DOMAIN
CDASH Variable
Value Level Metadata
Controlled Terminology

SDTM
DOMAIN
SDTM Variable
Value Level Metadata
Controlled Terminology
Computational algorithm

ADaM
DOMAIN
ADaM Variable
ADaM Parameters
Controlled Terminology
Computational algorithm

ADaM Domain | Variables | Computational Algorithm

Domain	Name	Label	Computational Algorithm
ADAE	USUBJID	Unique Subject Identifier	
ADAE	SUBJID	subject identifier for the study	
ADAE	SITEID	Study Site identifier	
ADAE	DOSEAONU	Study Drug Dose at AE Onset Units	ADAE.DOSEAEONU
ADAE	DOSEAEON	Study Drug Dose at AE Onset	ADAE.DOSEAEON
ADAE	COUNTRY	Country	
ADAE	ASTTM	Analysis Start Time	ADAE.ASTTM
ADAE	ASTDT	Analysis Start Time	ADAE.ASTDT
ADAE	AETERM	Reported Term for the Adverse Events	

Related metadata: SDTM Domain | Variables | Computational Algorithm | DCM | Tables | Figures | Listings

CDASH Domain | Variables

CDASH
DOMAIN
CDASH Variable
Value Level Metadata
Controlled Terminology

SDTM
DOMAIN
SDTM Variable
Value Level Metadata
Controlled Terminology
Computational algorithm

ADaM
DOMAIN
ADaM Variable
ADaM Parameters
Controlled Terminology
Computational algorithm

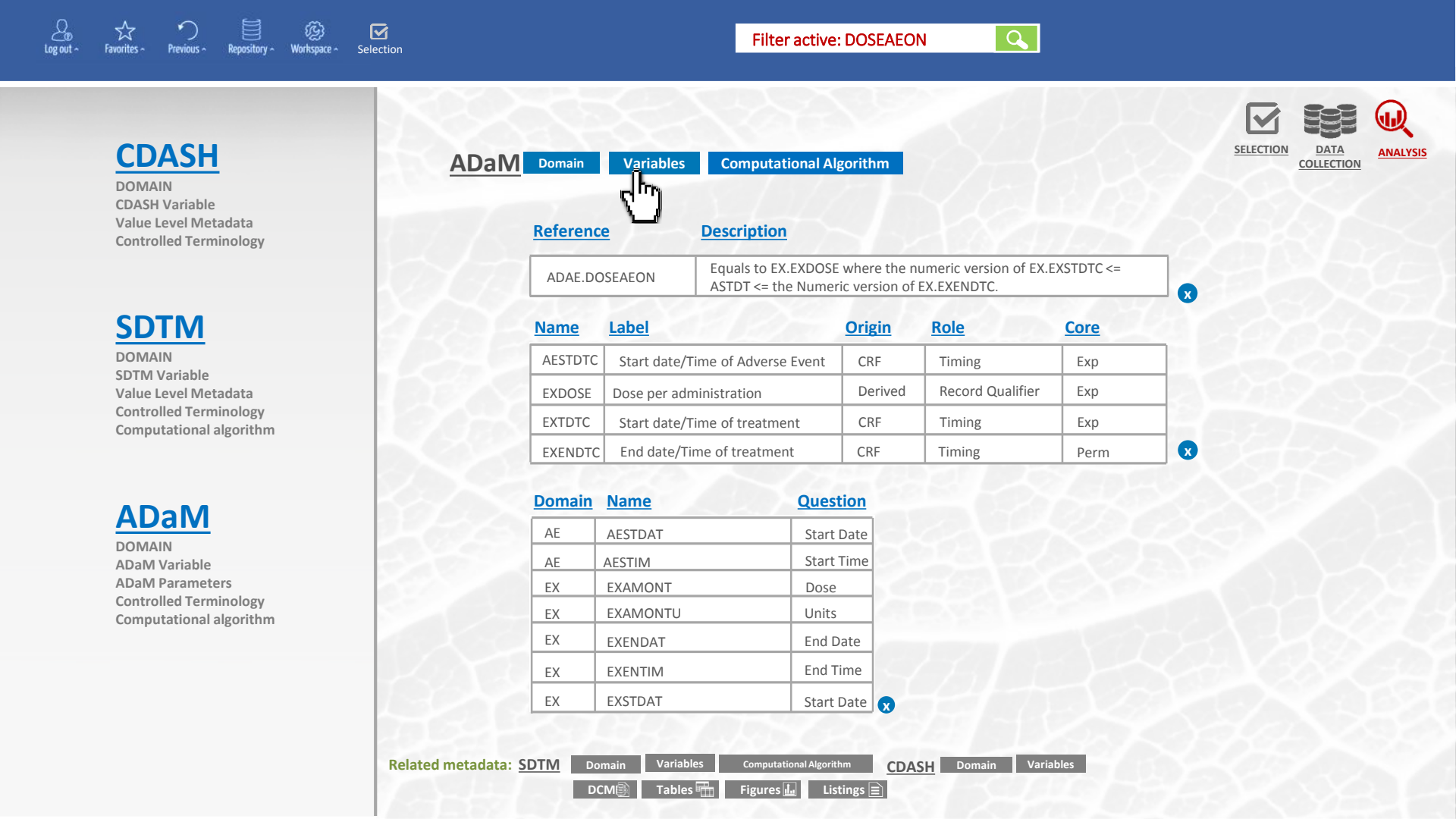
ADaM | Domain | Variables | Computational Algorithm

Reference **Description**

ADAE.AENDT	Equals to % SDTM_DATE_VARIABLE % transformed into % DATE_NUMERIC_FORMAT% when length (%SDTM_DATE_VARIABLE%) > 9
ADAE.ADURN	Equals to ADAE.AENDT – ADAE.ASTDT + 1.
ADAE.DOSEAEON	Equals to EX.EXDOSE where the numeric version of EX.EXSTDTC <= ASTDT <= the Numeric version of EX.EXENDTC.
ADAE.DOSEAEONU	Equals to EX.EXDOSU where the numeric version of EX.EXSTDTC <= ASTDT <= the Numeric version of EX.EXENDTC.
ADAE.DOSEAEON	Equals to ''DAYS''

Related metadata: SDTM | Domain | Variables | Computational Algorithm | CDASH | Domain | Variables

DCM | Tables | Figures | Listings


Log out
Favorites
Previous
Repository
Workspace
Selection
Filter active: DOSEAEON
SELECTION
DATA COLLECTION
ANALYSIS
CDASH
DOMAIN
CDASH Variable
Value Level Metadata
Controlled Terminology
SDTM
DOMAIN
SDTM Variable
Value Level Metadata
Controlled Terminology
Computational algorithm
ADaM
DOMAIN
ADaM Variable
ADaM Parameters
Controlled Terminology
Computational algorithm
ADaM
Domain
Variables
Computational Algorithm
Reference
Description
ADAE.DOSEAEON
Equals to EX.EXDOSE where the numeric version of EX.EXSTDTC <= ASTDT <= the Numeric version of EX.EXENDTC.
Name
Label
Origin
Role
Core
AESTDTC
Start date/Time of Adverse Event
CRF
Timing
Exp
EXDOSE
Dose per administration
Derived
Record Qualifier
Exp
EXTDTC
Start date/Time of treatment
CRF
Timing
Exp
EXENDTC
End date/Time of treatment
CRF
Timing
Perm
Domain
Name
Question
AE
AESTDAT
Start Date
AE
AESTIM
Start Time
EX
EXAMONT
Dose
EX
EXAMONTU
Units
EX
EXENDAT
End Date
EX
EXENTIM
End Time
EX
EXSTDAT
Start Date
Related metadata:
SDTM
Domain
Variables
Computational Algorithm
CDASH
Domain
Variables
DCM
Tables
Figures
Listings

CDASH

DOMAIN
CDASH Variable
Value Level Metadata
Controlled Terminology

SDTM

DOMAIN
SDTM Variable
Value Level Metadata
Controlled Terminology
Computational algorithm

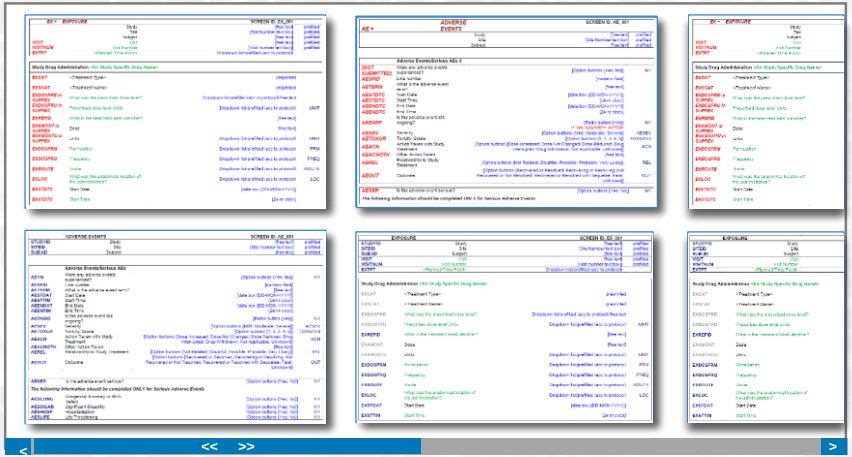
ADaM

DOMAIN
ADaM Variable
ADaM Parameters
Controlled Terminology
Computational algorithm

SELECTION DATA COLLECTION ANALYSIS

ADaM Domain Variables Computational Algorithm

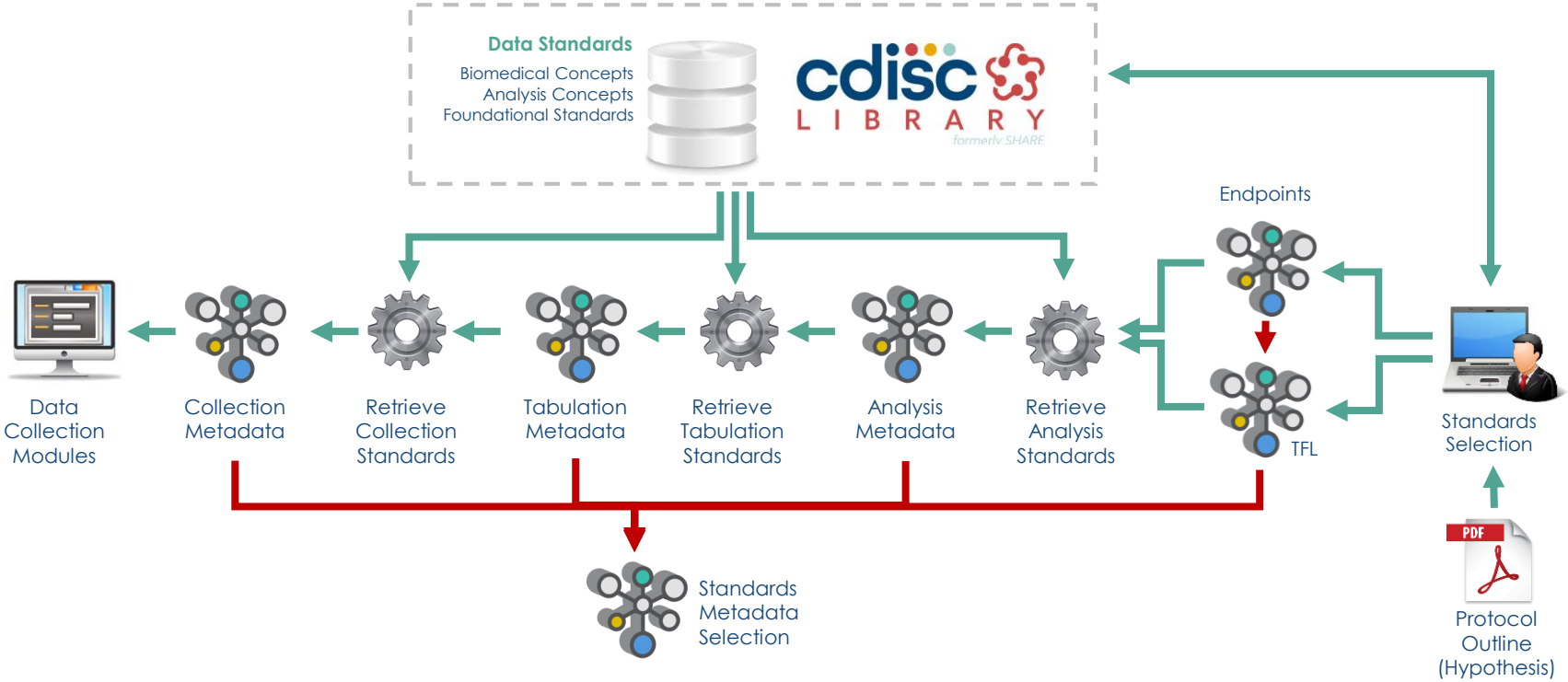
DCM's



Related metadata: SDTM Domain Variables Computational Algorithm CDASH Domain Variables
DCM Tables Figures Listings

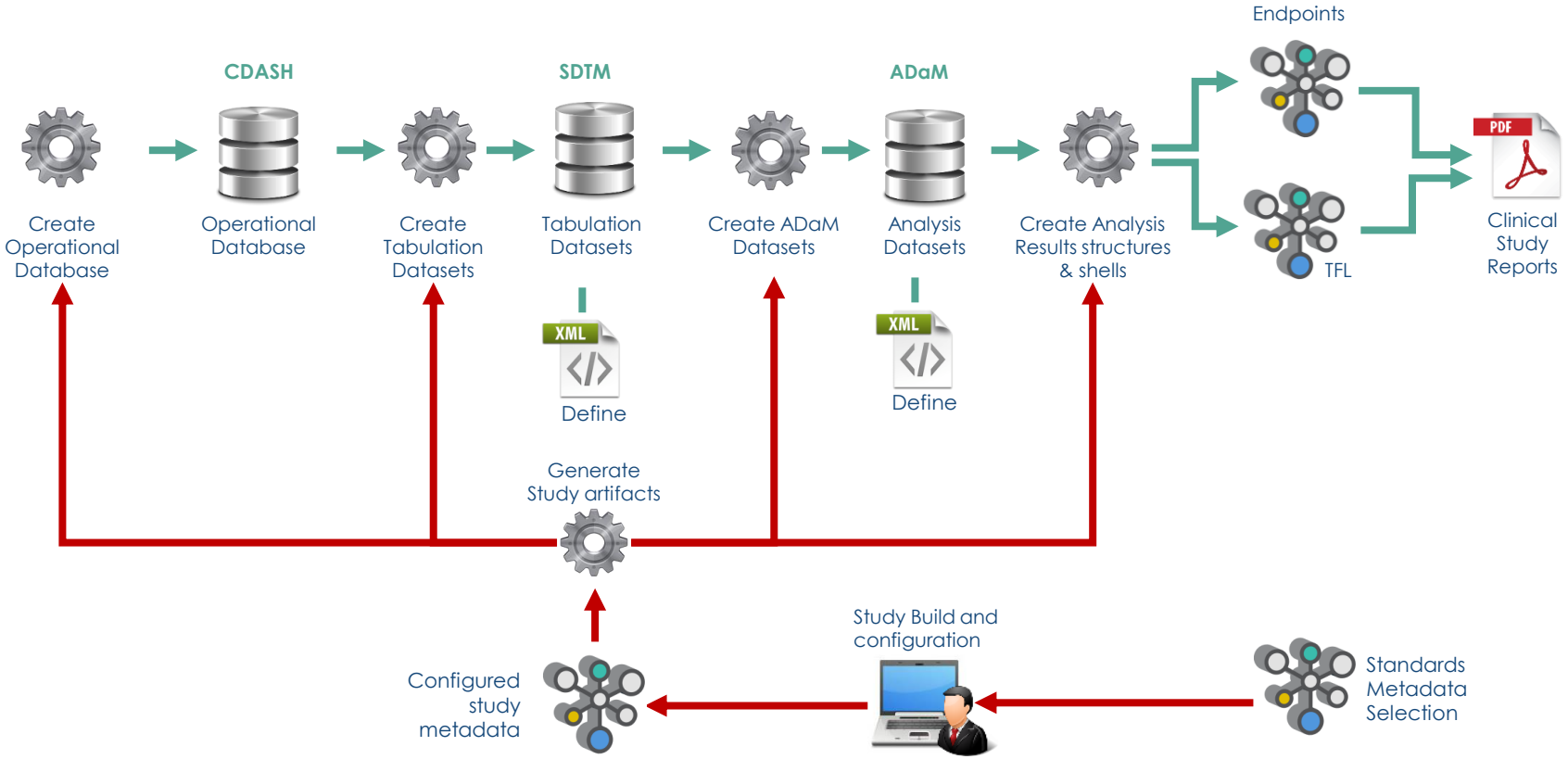
Use Case 1 : End to Start specification

Selecting standards concepts and linked metadata needed for a study

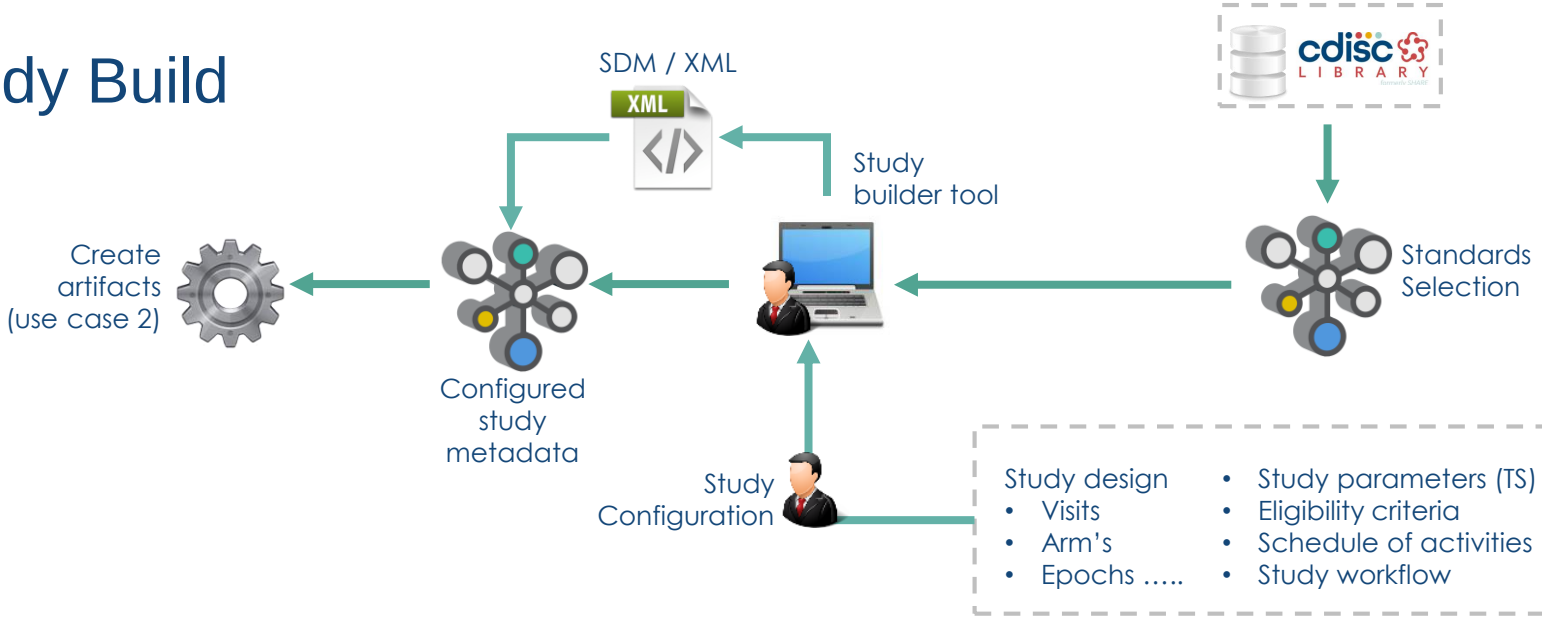


Use Case 2 : Start to End Study Metadata

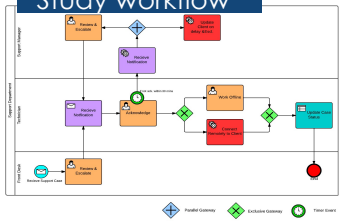
Adding study design, concept configuration & generate artifacts



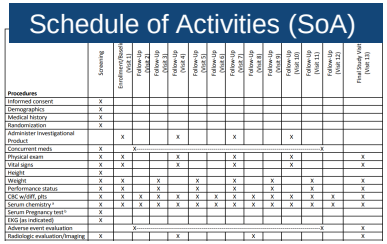
Study Build



Study workflow

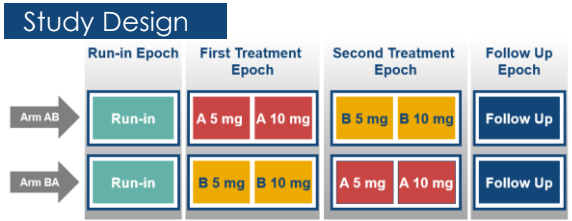


Schedule of Activities (SoA)

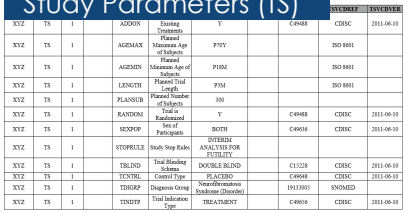


Procedures	Screening	Enrollment/Baseline (Visit 2)	Follow-Up (Week 4)	Follow-Up (Week 8)	Follow-Up (Week 12)	Follow-Up (Week 16)	Follow-Up (Week 20)	Follow-Up (Week 24)	Follow-Up (Week 28)	Follow-Up (Week 32)	Follow-Up (Week 36)	Follow-Up (Week 40)	Follow-Up (Week 44)	Follow-Up (Week 48)	Follow-Up (Week 52)	Final Study Visit (Visit 53)
Informed consent	X															
Demographics	X															
Medical history	X															
Randomization	X															
Administer Investigational Product		X				X				X				X		
Concurrent meds	X			X											X	
Physical exam	X	X				X				X				X		X
Vital signs	X	X				X				X				X		X
Height	X															
Weight	X	X			X		X			X		X			X	X
Performance status	X	X			X		X			X		X			X	X
CBC w/diff, plts	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Serum chemistry*	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Serum Pregnancy test**	X															
EKG (as indicated)	X															
Adverse event evaluation			X												X	X
Radiologic evaluation/Imaging	X					X				X						X

Study Design



Study Parameters (TS)

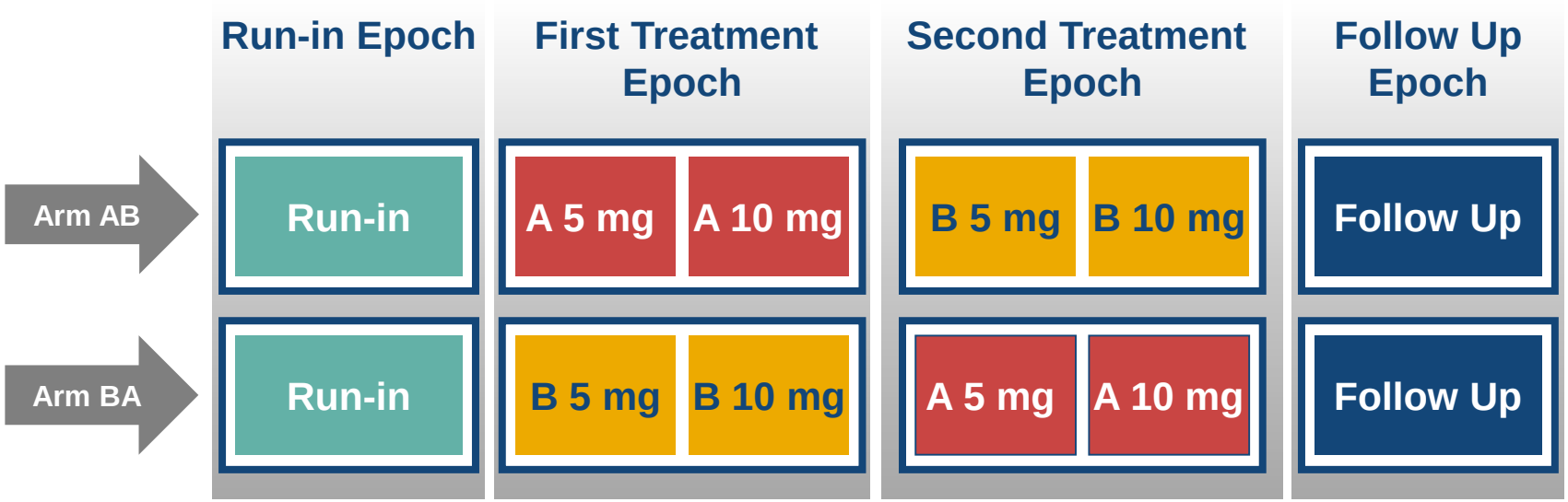


XYZ	TS	1		ADDON	Existing Treatments	Y		C49488	CDISC	2011-06-10
XYZ	TS	1		AGEMAX	Planned Maximum Age of Subjects	P70Y			ISO 8601	
XYZ	TS	1		AGEMIN	Planned Minimum Age of Subjects	P18M			ISO 8601	
XYZ	TS	1		LENGTH	Planned Trial Length	P3M			ISO 8601	
XYZ	TS	1		PLANSUB	Planned Number of Subjects	300				
XYZ	TS	1		RANDOM	Trial is Randomized	Y		C49488	CDISC	2011-06-10
XYZ	TS	1		SEXPOP	Sex of Participants	BOTH		C49636	CDISC	2011-06-10
XYZ	TS	1		STOPRULE	Study Stop Rules	INTERIM ANALYSES FOR FUTILITY				
XYZ	TS	1		TBLIND	Trial Blinding Schema	DOUBLE BLIND		C15228	CDISC	2011-06-10
XYZ	TS	1		TCNTRL	Control Type	PLACEBO		C49648	CDISC	2011-06-10
XYZ	TS	1		TDIGRP	Diagnosis Group	Neuroblastoma Syndrome (Disorder)		19133005	SNOMED	
XYZ	TS	1		TINDTP	Trial Indication Type	TREATMENT		C49656	CDISC	2011-06-10

Study Parameters (TS)

STUDYID	DOMAIN	TSSEQ	TSGRPID	TSPARMCD	TSPARM	TSVAL	TSVALNF	TSVALCD	TSVCDREF	TSVCDVER
XYZ	TS	1		ADDON	Added on to Existing Treatments	Y		C49488	CDISC	2011-06-10
XYZ	TS	1		AGEMAX	Planned Maximum Age of Subjects	P70Y			ISO 8601	
XYZ	TS	1		AGEMIN	Planned Minimum Age of Subjects	P18M			ISO 8601	
XYZ	TS	1		LENGTH	Planned Trial Length	P3M			ISO 8601	
XYZ	TS	1		PLANSUB	Planned Number of Subjects	300				
XYZ	TS	1		RANDOM	Trial is Randomized	Y		C49488	CDISC	2011-06-10
XYZ	TS	1		SEXPOP	Sex of Participants	BOTH		C49636	CDISC	2011-06-10
XYZ	TS	1		STOPRULE	Study Stop Rules	INTERIM ANALYSIS FOR FUTILITY				
XYZ	TS	1		TBLIND	Trial Blinding Schema	DOUBLE BLIND		C15228	CDISC	2011-06-10
XYZ	TS	1		TCNTRL	Control Type	PLACEBO		C49648	CDISC	2011-06-10
XYZ	TS	1		TDIGRP	Diagnosis Group	Neurofibromatosis Syndrome (Disorder)		19133005	SNOMED	
XYZ	TS	1		TINDTP	Trial Indication Type	TREATMENT		C49656	CDISC	2011-06-10

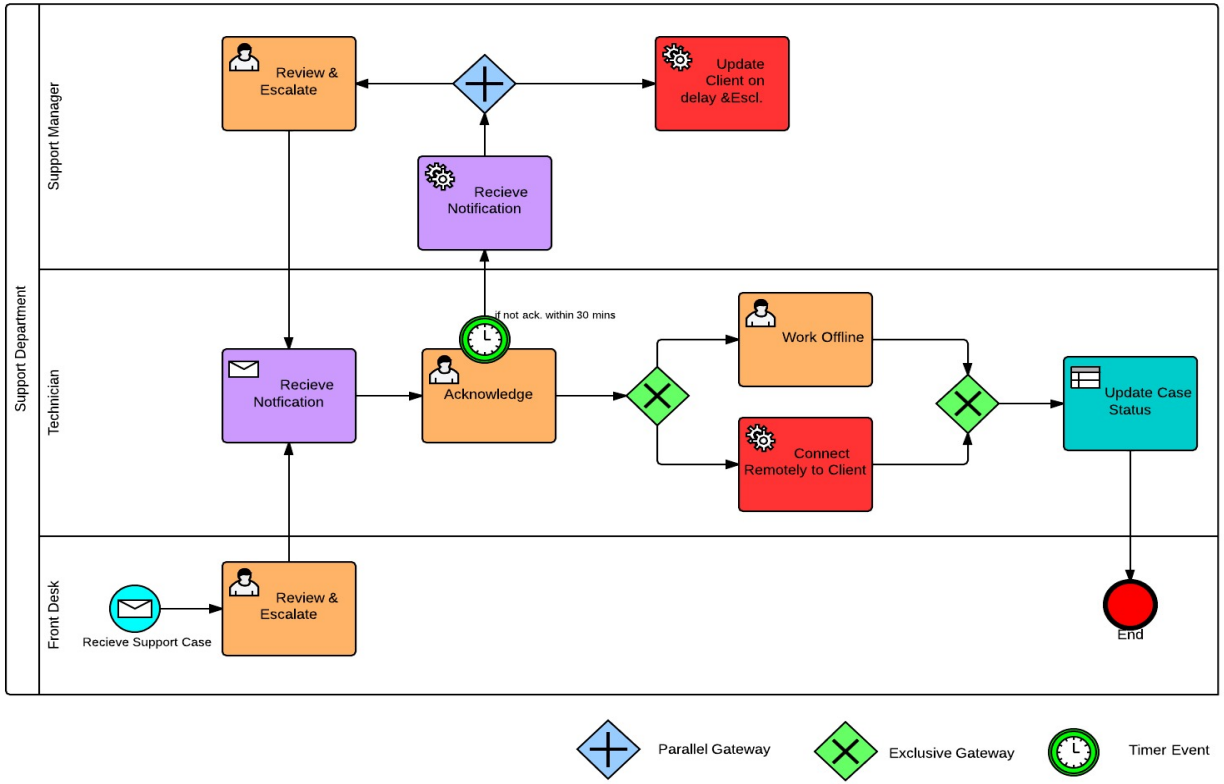
Study Design



Schedule of Activities (SoA)

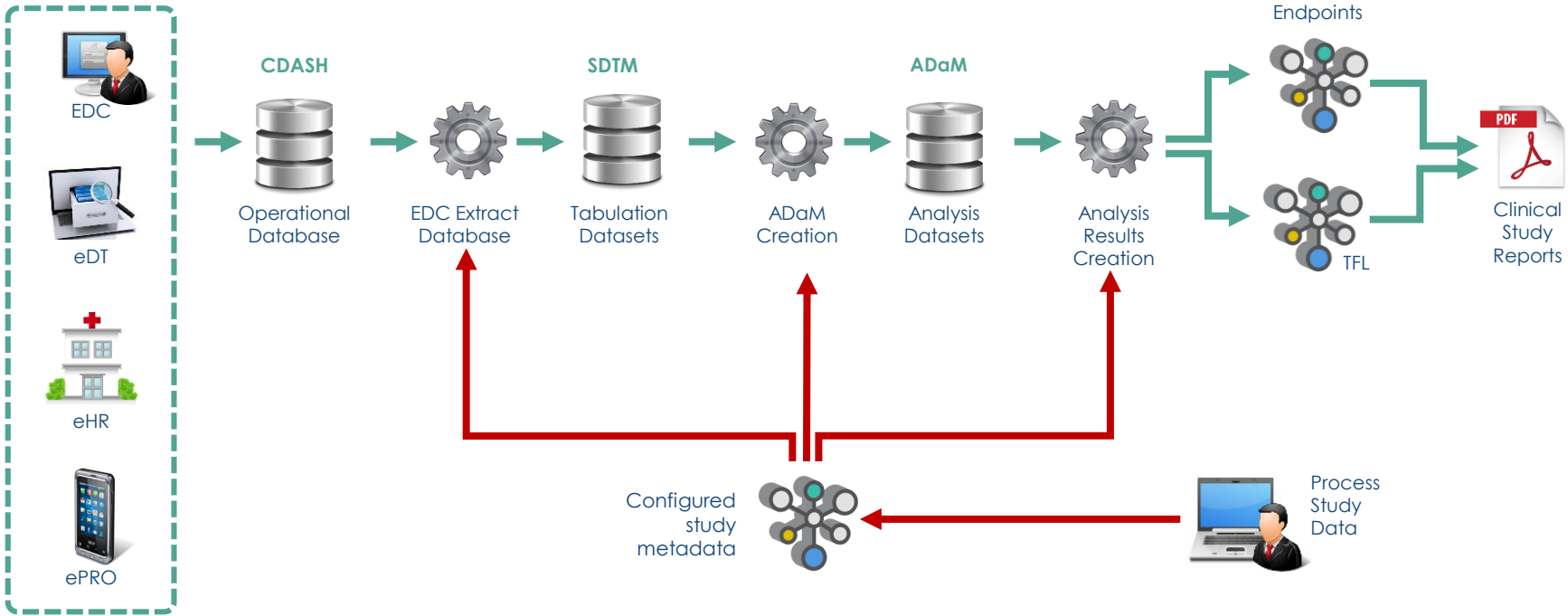
Procedures	Screening	Enrollment/Baseline (Visit 1)	Follow-Up (Visit 2)	Follow-Up (Visit 3)	Follow-Up (Visit 4)	Follow-Up (Visit 5)	Follow-Up (Visit 6)	Follow-Up (Visit 7)	Follow-Up (Visit 8)	Follow-Up (Visit 9)	Follow-Up (Visit 10)	Follow-Up (Visit 11)	Follow-Up (Visit 12)	Final Study Visit (Visit 13)
Informed consent	X													
Demographics	X													
Medical history	X													
Randomization	X													
Administer Investigational Product		X			X			X			X			
Concurrent meds	X	X----------											-X	
Physical exam	X	X			X			X			X			X
Vital signs	X	X			X			X			X			X
Height	X													
Weight	X	X		X		X		X		X		X		X
Performance status	X	X		X		X		X		X		X		X
CBC w/diff, plts	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Serum chemistry [a]	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Serum Pregnancy test [b]	X													
EKG (as indicated)	X													
Adverse event evaluation		X----------											-X	X
Radiologic evaluation/Imaging	X				X				X					X

Study workflow



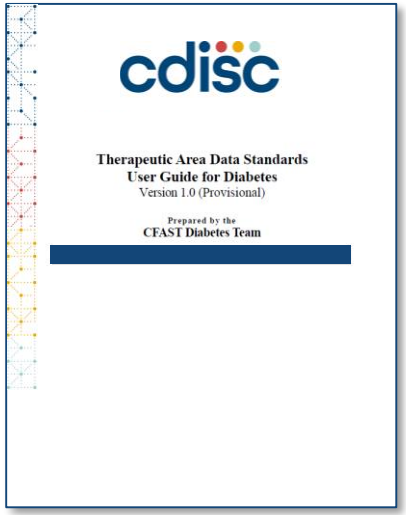
Use Case 3 : Start to End Data Processing

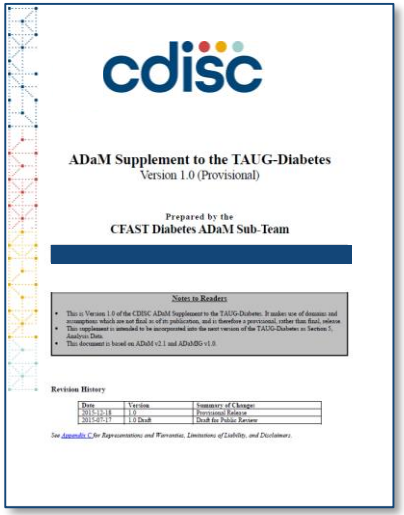
Automatic population of data into artifacts



Project Standards Scope

Diabetes TAUG



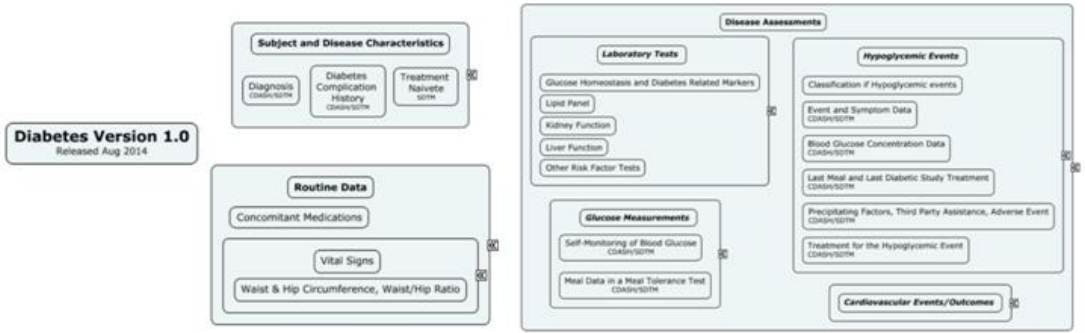


- 1 or 2 statistical endpoints
- 3 to 4 ADaM datasets
- 7 to 8 SDTM datasets
- 15 Data Collection Modules

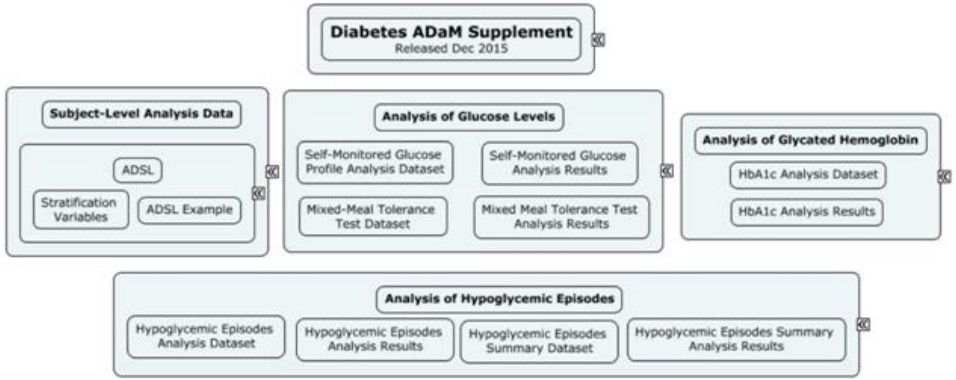
→ ***Reason for this scope:*** *the Diabetes TAUG provides standardized artifacts from analysis outputs to data collection. This allows the project team to focus on innovation and not on establishing a new data standard.*

Diabetes TAUG

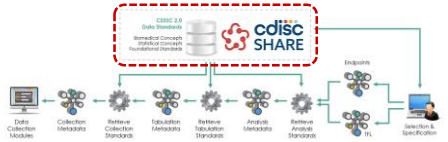
Diabetes V1

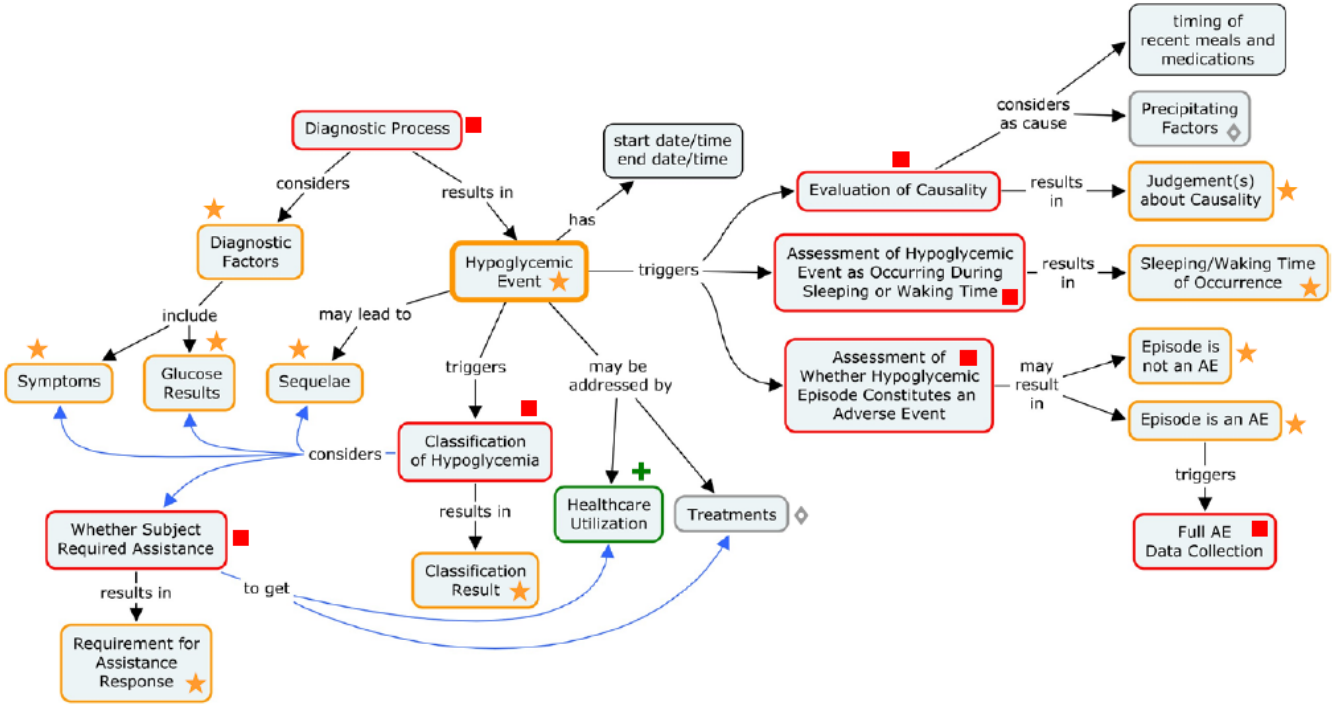


Diabetes ADaM V1



Biomedical Concept Map





cdisc

Analysis Results Shells

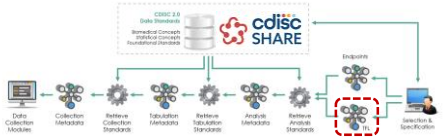


Table 3.2.1: Summary of Post-Meal Hypoglycemic Episodes by Severity – Table Shell

Hypoglycemic episodes within 2 hours since last meal by severity

Summary - Safety Analysis Set

	Drug A			Drug B		
	N	(%)	E	N	(%)	E
Number of subjects	xxx			xx		
Diurnal	xxx	(xx.x)	xxx	xx	(xx.x)	xxx
Documented Symptomatic	xx	(xx.x)	xx	xx	(xx.x)	xx
Pseudo Symptomatic	xx	(xx.x)	xx	xx	(xx.x)	xx
Probable Symptomatic	x	(xx.x)	xx	x	(x.x)	x
Nocturnal	x	(x.x)	x	x	(x.x)	x
Documented Symptomatic	x	(x.x)	x	x		
Probable Symptomatic	x			xx	(x.x)	x

N: Number of subjects; %: Percentage of subjects; E: Number of events

Analysis Dataset Metadata

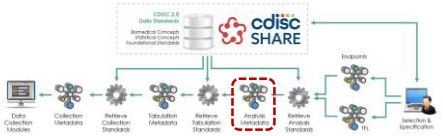


Table 3.3.1: ADHYSUM Analysis Dataset

Row	STUDYID	USUBJID	PARAMCD	PARAM	AVISIT	AVAL	TRTDURD	SEX	AGE	COUNTRY	TRTA
1	XYZ	000008	ASSYMP	Asymptomatic Hypoglycemia (frequency count)	Week 1	3	72	F	35	DZA	Drug B
2	XYZ	000008	ASSYMPC	Asymptomatic Hypoglycemia (cumulative frequency count)	Week 1	3	72	F	35	DZA	Drug B
3	XYZ	000008	ASSYMP	Asymptomatic Hypoglycemia (frequency count)	Week 2	1	72	F	35	DZA	Drug B
4	XYZ	000008	ASSYMPC	Asymptomatic Hypoglycemia (cumulative frequency count)	Week 2	4	72	F	35	DZA	Drug B
5	XYZ	000008	ASSYMP	Asymptomatic Hypoglycemia (frequency count)	Week 3	0	72	F	35	DZA	Drug B
6	XYZ	000008	ASSYMPC	Asymptomatic Hypoglycemia (cumulative frequency count)	Week 3	4	72	F	35	DZA	Drug B
7	XYZ	000008	ASSYMP	Asymptomatic Hypoglycemia (frequency count)	Week 4	1	72	F	35	DZA	Drug B
8	XYZ	000008	ASSYMPC	Asymptomatic Hypoglycemia (cumulative frequency count)	Week 4	5	72	F	35	DZA	Drug B
10	XYZ	000008	ASSYMPC	Asymptomatic Hypoglycemia (cumulative frequency count)	End of Treatment	7	72	F	35	DZA	Drug B
…	…	…	…	…	…	…	…	…	…	…	…
20	XYZ	000008	DOCSEVC	Documented Symptomatic or Severe Hypoglycemia (cumulative frequency count)	End of Treatment	17	72	F	35	DZA	Drug B

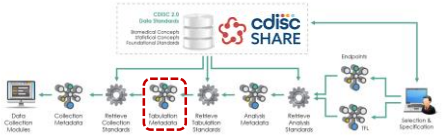
Table 3.3.2: ADHYSUM Dataset Metadata

Dataset Name	Dataset Description	Dataset Location	Dataset Structure	Keys	Class	Documentation
ADHYSUM	Hypoglycemic Episodes Summary Data	ADHYSUM.xpt	One record per subject per analysis visit per parameter	STUDYID, USUBJID, AVISIT, PARAMCD	BDS	ADHYSUM.SAS/SAP

Table 3.3.3: ADHYSUM Variable Metadata

Variable Name	Variable Label	Type	Length/Display Format	Codelist/Controlled Terms	Source/Derivation/Comment
STUDYID	Study Identifier	text	$12		ADSL.STUDYID
USUBJID	Unique Subject Identifier	text	$20		ADSL.USUBJID
PARAMCD	Parameter Code	text	$8		See parameter value metadata. *Note that the tables below do not present all possible values for PARAMCD but only those that correspond to the data display.*
PARAM	Parameter	text	$80		See parameter value metadata. *Note that the tables below do not present all possible values for PARAM but only those that correspond to the data display.*
AVISIT	Analysis Visit	text	$13	Week -1; Week 0; Week 1; Week N; End of Treatment	Refer to Section X.X of the SAP for windowing and imputation algorithms based on ADHYPO.ADY. End-of-treatment is defined as the last week during which the subject is on treatment.
AVAL	Analysis Value	integer	8		See parameter value metadata.
TRTDURD	Total Treatment Duration (Days)	integer	8		ADSL.TRTDURD
SEX	Sex	text	$1		ADSL.SEX
AGE	Age	integer	8		ADSL.AGE
COUNTRY	Country	text	$3		ADSL.COUNTRY
TRTA	Actual Treatment	text	$32		ADSL.TRT01A

Tabulation Metadata



Row	STUDYID	DOMAIN	USUBJID	CESEQ	CECAT	CETERM	CEDECOD	CEPRESP	CEOCCUR	CESTDTC	CESTDY
2	XYZ	CE	XYZ-001-001	2	HYPO SYMPTOMS	SWEATING	Hyperhidrosis	Y	N		
3	XYZ	CE	XYZ-001-001	3	HYPO SYMPTOMS	TREMORS/TREMBLING	Tremor	Y	N		
4	XYZ	CE	XYZ-001-001	4	HYPO SYMPTOMS	DIZZINESS	Dizziness	Y	N		
5	XYZ	CE	XYZ-001-001	5	HYPO SYMPTOMS	COGNITIVE IMPAIRMENT	Cognitive Disorder	Y	Y		
6	XYZ	CE	XYZ-001-001	6	HYPO SYMPTOMS	LOSS OF CONSCIOUSNESS	Loss of Consciousness	Y	Y		
7	XYZ	CE	XYZ-001-001	7	HYPO SYMPTOMS	CONVULSIONS/SEIZURES	Convulsion	Y	N		
8	XYZ	CE	XYZ-001-001	8	HYPO SYMPTOMS	COMA	Coma	Y	N		
9	XYZ	CE	XYZ-001-001	9	HYPO EVENTS	HYPOGLYCEMIC EVENT	Hypoglycaemia			2013-09-24T08:48	50

Row	RELMIDS	MIDS	MIDSDTC
1 (cont)		HYPO 1	
2 (cont)	DURING	HYPO 1	2013-09-01T11:00
3 (cont)	DURING	HYPO 1	2013-09-01T11:00
4 (cont)	DURING	HYPO 1	2013-09-01T11:00
5 (cont)	DURING	HYPO 1	2013-09-01T11:00
6 (cont)	DURING	HYPO 1	2013-09-01T11:00
7 (cont)	DURING	HYPO 1	2013-09-01T11:00
8 (cont)	DURING	HYPO 1	2013-09-01T11:00
9 (cont)		HYPO 2	

suppce.xpt

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL
1	XYZ	CE	XYZ-001-001	CESEQ	1	WHEOCC	When did the hypoglycemic event occur?	BETWEEN BEDTIME AND WAKING
2	XYZ	CE	XYZ-001-001	CESEQ	8	WHENOCC	When did the hypoglycemic event occur?	BETWEEN BEDTIME AND WAKING

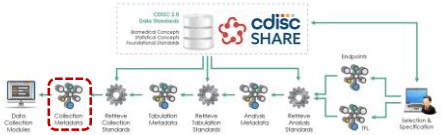
lb.xpt

Row	STUDYID	DOMAIN	USUBJID	SPDEVID	LBSEQ	LBTESTCD	LBTEST	LBORRES	LBORRESU	LBSTRESC	LBSTRESN	LBSTRESU
1	XYZ	LB	XYZ-001-001	GLUCOSE METER	1	GLUC	GLUCOSE	60	mg/dL	3.3	3.3	mmol/l
2	XYZ	LB	XYZ-001-001	GLUCOSE METER	2	GLUC	GLUCOSE	65	mg/dL	3.6	3.6	mmol/l

ml.xpt

Row	STUDYID	DOMAIN	USUBJID	MLSEQ	MLTRT	MLSTDTC	RELMIDS	MIDS	MIDSDTC
1	XYZ	ML	XYZ-001-001	1	MEAL	2013-08-31T20:00	LAST MEAL PRIOR TO	HYPO 1	2013-09-01T11:00
2	XYZ	ML	XYZ-001-001	2	MEAL	2013-09-23T22:30	LAST MEAL PRIOR TO	HYPO 2	2013-09-24T08:48

Collection Metadata



Example CRF 5: Hypoglycemia

CETERM= Hypoglycemic Event
CECAT= HYPO EVENTS

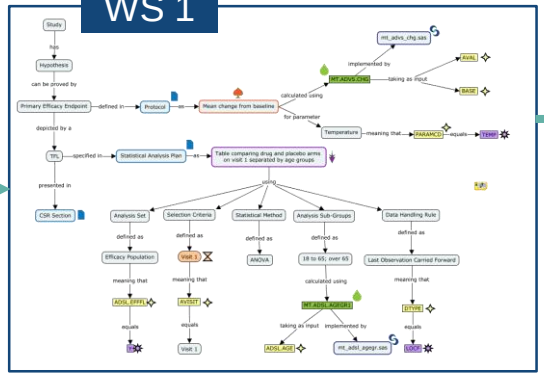
Any Hypoglycemic Events Experienced?	No Yes (If yes complete for each event) CEYN		
Sponsor Defined ID CESPID	001		
Date/Time of Event CESTDTC	-- --- ---- (DD-MMM-YYYY) -:-- (24 hour clock) CESTDAT CESTTIM		
When Did the Hypoglycemic Event Occur?	Between Bedtime and Waking Between Waking and Bedtime QVAL when QNAM= WHENOCC and QLABEL="When Did the Hypoglycemic Event Occur?"		
In the Opinion of the Investigator Was This an Adverse Event?	No Yes WASAEYN FAORRES where FATESTCD= "WASAEYN", FATEST= "Was this an adverse event?" and FAOBJ="HYPOGLYCEMIC EVENT"		
Was a Glucose Measurement Obtained at the Time of the Event? LBSTAT	No Yes (If yes enter result and unit below) LBPERF		
	- - - - Glucose Result LBORRES mg/dL mmol/L LBORRESU		
Last Study Medication Taken	----------Name/Reference EXTRT		
EXCAT= HIGHLIGHTED DOSE EXSTDTC	-- --- ---- (DD-MMM-YYYY) -:-- (24 hour clock) EXSTDAT EXSTTIM		
	- - - dose EXDOSE - - - units EXDOSU EXDSTXT		
Last Concomitant Diabetic Medication Taken	----------Name/Reference CMTRT		
CMCAT= ANTI-HYPERGLYCEMIC MED CMSCAT= HIGHLIGHTED DOSE CMSTDTC	-- --- ---- (DD-MMM-YYYY) -:-- (24 hour clock) CMSTDAT CMSTTIM		
	- - - dose CMDOSE - - - units CMDOSU CMDSTXT		
Date/Time of Last Meal MLSTDTC	-- --- ---- (DD-MMM-YYYY) -:-- (24 hour clock) MLSTDAT MLSTTIM		
Were Signs/Symptoms Present? CECAT= HYPO SYMPTOMS	No Yes (If yes complete following) CEYN		
CETERM= SWEATING	Sweating	No Yes	CEOCCUR with CEPRESP=Y
CETERM= TREMORS/TREMBLING	Tremors/Trembling	No Yes	
CETERM= DIZZINESS	Dizziness	No Yes	
CETERM= COGNITIVE IMPAIRMENT	Cognitive Impairment	No Yes	
CETERM= LOSS OF CONSCIOUSNESS	Loss of Consciousness	No Yes	
CETERM= CONVULSIONS/SEIZURE	Convulsions/Seizure	No Yes	
CETERM= COMA	Coma	No Yes	
	Other (Specify)	No Yes (if yes enter below)	
FACAT= PRECIPITATING FACTORS, FAOBJ= HYPOGLYCEMIC EVENT and:		---------- CETERM	
Were Any Precipitating Factors Reported?	No Yes (If yes complete following) HPFYN		
FATEST= Alcohol Consumption as a Precip Factor	Alcohol Consumption	No Yes	FAORRES
FATEST= Concurrent Illness as a Precip Factor	Concurrent Illness	No Yes	
FATEST= Dosing Deviation as a Precip Factor	Deviation from Dosing Instructions	No Yes	
FATEST= Meal Variance as a Precip Factor	Missed, Delayed or Smaller Meal	No Yes	
FATEST= Physical Activity as a Precip Factor	Physical Activity	No Yes	
	Other (Specify)	No Yes (if yes enter below)	
CMCAT= HYPO TREATMENT		---------- FATEST	
Was Any Treatment Given for the Hypoglycemic Event?	No Yes (If yes complete following) HTGYN		
CMTRT= DRINK	Drink	No Yes	CMOCCUR with CMPRESP= Y
CMTRT= FOOD	Food	No Yes	
CMTRT= GLUCOSE TABLETS	Glucose Tablets	No Yes	
CMTRT= GLUCAGON INJECTION	Glucagon Injection	No Yes	
CMTRT= INTRAVENOUS GLUCOSE	Intravenous Glucose	No Yes	
If Treatment Given Indicate Assistance Needed?	None - Subject Treated Self		FAORRES when FAOBJ= HYPOGLYCEMIC EVENT, FACAT= TREATMENT ADMINISTRATION, FATESTCD= TXASSIST FATEST=Treatment Assistance
	Subject was Capable of Treating Self, but Received Assistance		
	Subject was Not Capable of Treating Self, and Required Assistance		

CRF annotated to show mapping SDTM variables are in **Red**. If CDASH variable differs from SDTM the CDASH variable is in **Blue**.

cdisc

3. Logistics

WS 1



Create concepts in knowledge graphs

WS2

Transform concepts in machine readable form

Load into library

Biomedical Concepts
Analysis Concepts
Foundational Standards

cdisc LIBRARY
formerly SHARE

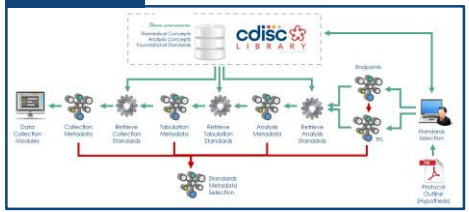
API

Extend API's

WS3

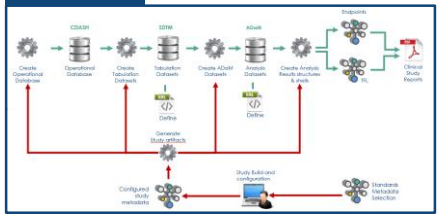
Add transformation semantics

WS 4



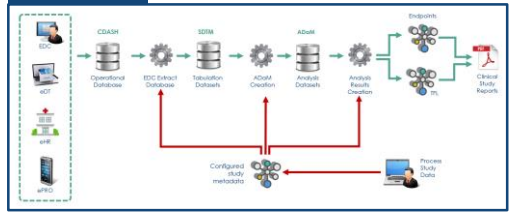
Identify and select standards specification (Use Case 1)

WS 5



Configure study specification and create artifacts (Use Case 2)

WS 6



Automatically process and transform data (Use Case 3)

Workstream 1 & 2

- Workstream 1 - End-to-end concept development
  - Design concept maps
  - Semantic end-to-end expression of concepts
  - Final analysis output to data collection instruments
  - Includes transformation information
  - Combine Biomedical Concepts (BC) with Analysis Concepts (AC)

- Workstream 2 - Machine-readable End-to-end concept development
  - Transform concepts in machine readable form
  - Load in to CDISC Library
  - Extend API's to extract multifunctional metadata

cdisc

Workstream 3 & 4

- Workstream 3 - Standard dataset definition extension to include transformation information
  - Add semantics in the form of transformation information (ETL)

- Workstream 4 - End-to-start standards specification development (Use Case 1)
  - Demonstrate identify and select capability
  - Ensure API output is complete
  - Combine all metadata in specification pool

Workstream 5

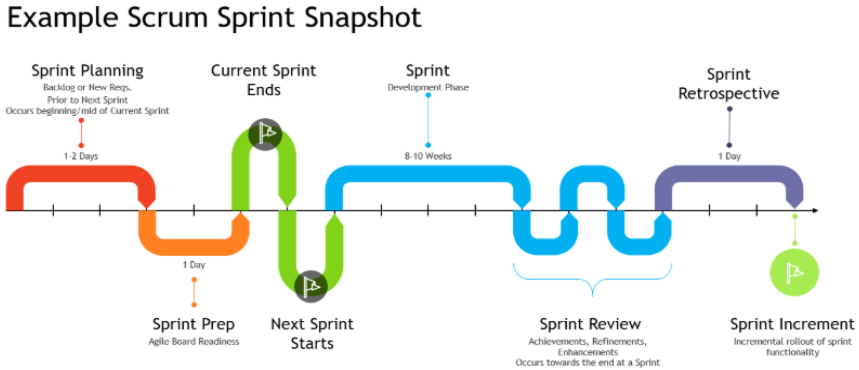
- Workstream 5 - Start-to-end study metadata development (Use Case 2)
  - Study specific configuration of standards metadata
  - Instantiate metadata on a study level
  - Demonstrate study build process (includes trial design information)
  - Create study artifacts
    - Datasets
    - Define xml
    - Analysis shells

Workstream 6

- Workstream 6 - Start-to-end auto process and transform (Use Case 3)
  - Process data from collection to analysis
    - Extract data from collection instruments
    - Create operational database (ODM v2)
    - Map and transform SDTM and ADaM data
    - Auto generate analysis outputs
  - Currently out of scope
    - Creation of study Protocol, SAP, CSR
    - Automated business rules (validation)

Agile Scrum Methodology and Timeframe

- What is agile scrum methodology
  - Continuous flexible development process: workstreams to be nimble, iterative, innovative, incremental, evolutionary, quality driven, adaptive, organized, and collaborative.
- Why use agile scrum methodology
  - Flexible mechanism to handle moments of change; e.g., technical limitations, requirements, or communication.
- Project timeframe: 18 months



www.agilemanifesto.org

CDISC 360 – CDISC Member Participation Process

- Participation Process Calendar:
  - 06Feb2019 – Invitation email sent to CDISC membership
  - 19Feb2019 – Q&A webinar for prospective participants
  - 01Mar2019 – Expressed interest to participate due
  - 02Mar2019 - 07Apr2019 – 360 Leadership Team organizes workstream teams
  - 08Apr2019 – Workstream teams' activities launch

- Resources available to members during process:
  - Webpage with high-level project description
  - Slide deck with mid-level project description
  - Spreadsheet with low-level workstream description and skills/backgrounds sought per workstream

4. Relationship to other Initiatives

Relationship to other initiatives

- Helmsley Transformational Grant
- Blue Ribbon Commission
- TransCelerate Digital Data Flow
- CDISC Data Exchange Standards
  - ODM v2
  - SDM-XML

→ ***CDISC 360:*** *a blueprint for the next generation data standards, aligned with key initiatives*

5. Expected outcome

Expected Outcome

- Learn
  - What works and what doesn't
- Assessment
  - Technology Gap Analysis
  - Standards Gap Analysis
- Building a base for the future
  - Effort calculation
  - Cost / Benefit Analysis
  - Scale up to deliver the standards metadata needed
  - Partnerships with vendors to ensure tools are made available

cdisc

Thank You!

Peter Van Reusel
Sam Hume
Barry Cohen