PhUSE 2016

Paper CD15

How to handle different versions of SDTM & DEFINE generation in a Single Study?

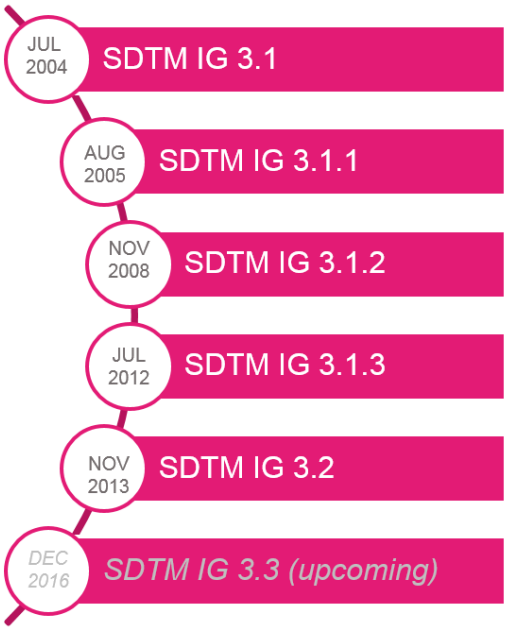
Edwin Ponraj Thangarajan, PRA Health Sciences, Chennai, India
Giri Balasubramanian, PRA Health Sciences, Chennai, India

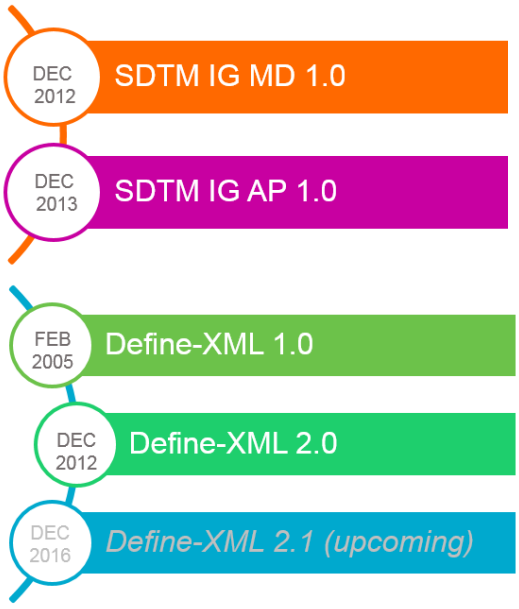
ABSTRACT

Submission Data Standards has been evolving over a period of time and it is still subject to revisions by way inclusion of new therapeutic areas in the form of domains and variables under different classifications. As well, the define.xml which provides a quick access to reviewers to go over the submissions once it is imported into JANUS data warehouse. These revisions and amendments does have impact on ongoing and new studies especially studies which are of longer duration in the case of Oncology studies. Longer duration studies would have started with an SDTM v3.1.3 with Define-XML v1.0 file generation for DSMB analysis and other related submissions. During the course of the study, due to revisions, the same study has to be done under a new version say SDTM v3.2 with Define-XML v2.0 generation. This session would be covering the processes and different steps with demo that needs to be taken care while implementing different versions of SDTM standards and Define-XML generation for a single on-going study.

SDTM/DEFINE-XML RELEASES BACKGROUND - OVERVIEW

SDTM first evolved from the Submission Data Model (SDM) as developed by CDISC Submission Data Standards (SDS) Team around 10 years ago. The SDS Team developed a set of consistent data standards for submitting tabulation data for human clinical trials. The SDS team is also responsible for maintaining the compatibility, consistency and conformity of SDTM data standards and associated SDTM Implementation Guide (SDTMIG) as part of the long running Clinical Data Interchange Standards Consortium (CDISC) strategy for harmonizing the submission of clinical trials data.





The purpose of Define-XML is to support the interchange of dataset metadata for clinical research applications in a machine-readable format. An important use case for Define-XML is to support the submission of clinical trials data in CDISC SDTM, SEND or ADaM format to regulatory authorities. The FDA has also announced the end of support for Define-XML v1.0 for studies would be March, 2017.

Define-XML v2.0 represents a significant update to Define-XML v1.0 in response to implementation experience with v1.0, the evolution of the SDTM, SEND and ADaM standards and best practices by SDTM and ADaM metadata experts. Key enhancements include:

- Support for CDISC Controlled Terminology
- Flexible definition of Value Level Metadata
- Enhanced documentation of data origin or source
- Improved support for ADaM metadata
- Improved handling of comments.

Define-XML v2.0 is the most current version of this standard and its use is recommended.

FDA SUPPORTED VERSION ON DATA STANDARDS:

FDA Data Standards Catalog v4.5 (07-18-2016) - Supported and Required Standards

This table contains a listing of the data exchange, file formats and terminology standards supported at FDA. These standards have gone through all the steps necessary to make this part of the regulatory review process, including posting of regulatory guidance documents and associated implementation guidelines and technical specifications. The submission of standardized data using any standard not listed, or to an FDA Center not listed, should be discussed with the Agency in advance. This catalog is incorporated by reference in the guidance to industry, *Providing Regulatory Submissions in Electronic format-Standardized Study Data (http://www.fda.gov/downloads/Drugs/Guidances/UCM292334.pdf)*.

Use	Data Exchange Standard	Exchange Format	Standards Development Organization (SDO)	Supported Version	Implementation Guide Version	FDA Center(s)	Date Support Begins (MM/DD/YYYY)	Date Support Ends (MM/DD/YYYY)	Date Requirement Begins (MM/DD/YYYY)	Date Requirement Ends	Regulatory Reference and Information Sources
Clinical study datasets	Study Data Tabulation Model (SDTM)	XPT	Clinical Data Interchange Standards Consortium (CDISC)	1.4	3.2	CDER, CBER	08/17/2015		03/15/2018 [1] 03/15/2019 [2]		CDISC.org - SDTM See Technical Conformance Guide
Clinical study datasets	SDTM	XPT	CDISC	1.3	3.1.3	CDER, CBER	12/01/2012		12/17/2016 [1] 12/17/2017 [2]		CDISC.org - SDTM
Clinical study datasets	SDTM	XPT	CDISC	1.2	Version 3.1.2 Amendment 1	CDER, CBER	08/07/2013		12/17/2016 [1] 12/17/2017 [2]		CDISC.org - SDTM
Clinical study datasets	SDTM	XPT	CDISC	1.2	3.1.2	CDER, CBER	10/30/2009		12/17/2016 [1] 12/17/2017 [2]		CDISC.org - SDTM
Clinical study datasets	SDTM	XPT	CDISC	1.1	3.1.1	CDER, CBER	Ongoing	01/28/2015			CDISC.org - SDTM
Clinical study datasets	Analysis Data Model (ADaM)	XPT	CDISC	2.1	1.0	CDER, CBER	Ongoing		12/17/2016 [1] 12/17/2017 [2]		CDISC.org - ADaM
Animal study datasets	Standard for Exchange of Nonclinical Data (SEND)	XPT	CDISC	1.2	3.0	CDER	06/13/2011		12/17/2016 [1] 12/17/2017 [2]		CDISC.org - SEND
Study data definition	Define	XML	CDISC	1.0	N/A	CDER, CBER, CDRH	Ongoing	03/17/2017	12/17/2016 [1] 12/17/2017 [2]		CDISC.org - Define-XML
Study data definition	Define	XML	CDISC	2.0	N/A	CDER, CBER, CDRH	08/07/2013		12/17/2016 [1] 12/17/2017 [2]		CDISC.org - Define-XML
This section is reserved for future therapeutic area data standards											

CHANGES FROM ONE SDS VERSION TO ANOTHER SDS VERSION

Generally CDISC published the following changes and it is classified as major and minor.

1. New domain inclusion
2. New variables included in the existing domain
3. Updates in the existing domain/variables
4. Updates in the assumptions
5. Controlled Terminologies
6. Trial Design
7. Additional more functionality

There are multiple ways to identify changes from one version to another.

A. USING CDISC eSHARE DOWNLOADS

For quick identification of changes from one version to another version can be found in the eSHARE downloads members only area. For example, it contains around 171 changes from SDTM 3.1.3 to 3.2

SDTM 1.4 (IG 3.2)				
2015-02-02	SDTM IG Diff	3.2 - 3.1.3	Metadata	Excel
2014-08-21	SDTM IG	1.4 (3.2)	Metadata	Define-XML v2.0 | Define-XML v1.0 | Excel
2014-08-21	SDTM + CT	1.4+2013-12-20	Bundle	Define-XML v2.0 | Excel
2014-08-21	SDTM IG	1.3 (3.1.3)	Document Bundle	PDF

Downloaded excel from the above table is shown below. The downloaded excel contains changes made in the new version by identifying it as 'Add', 'Update', 'Remove'. Based on the information available, user would be able to identify the changes quickly and list the different type of work to be undertaken in output dataset creation and Define-XML file generation.

1	Change Nbr	Action	Impact	Domain	Variable	Attribute (updated)	Attribute (previous)	Updated SDTM-IG Ver	Previous SDTM-IG Ver
2	0	Add	Domain	APDM				3.2	3.1.3
3	1	Add	Domain	APRELSUB				3.2	3.1.3
4	2	Update	CDISC Notes and Definitions	AE	AEACN	Describes changes to the study treatment as a result of the event. AEACN is specifically for the relationship to study treatment. AEACNOTH is for actions unrelated to dose adjustments of study treatment. Examples of AEACN values include ICH E2B values: DRUG WITHDRAWN, DOSE REDUCED, DOSE INCREASED, DOSE NOT CHANGED, UNKNOWN or NOT APPLICABLE.	Describes changes to the study treatment as a r	3.2	3.1.3
5	3	Add	Controlled Terminology	AE	AEDECOD	MedDRA		3.2	3.1.3
6	4	Update	CDISC Notes and Definitions	AE	AEDECOD	Dictionaryderived text description of AETERM or AEMODIFY. Equivalent to the Preferred Term (PT in MedDRA). The sponsor is expected to provide the dictionary name and version used to map the terms utilizing the define.xml external codelist attributes.	Dictionaryderived text description of AETERM or	3.2	3.1.3
7	5	Update	Controlled Terminology	AE	AEENRTPT	(STENRF)	BEFORE, AFTER, COINCIDENT, ONGOING, U	3.2	3.1.3
8	6	Update	CDISC Notes and Definitions	AE	AELOC	Describes anatomical location relevant for the event (e.g., ARM for skin rash).	Describes anatomical location relevant for the e	3.2	3.1.3
9	7	Update	Role	AE	AEPRESP	Variable Qualifier	Record Qualifier	3.2	3.1.3
10	8	Update	CDISC Notes and Definitions	AE	AEREFID	Internal or external identifier such as a serial number on an SAE reporting form.	Internal or external identifier such as a serial num	3.2	3.1.3
11	9	Update	CDISC Notes and Definitions	AE	AEREL	Records the investigators opinion as to the causality of the event to the treatment. ICH E2A and E2B examples include NOT RELATED, UNLIKELY RELATED, POSSIBLY RELATED, RELATED. Controlled Terminology may be defined in the future. Check with regulatory authority for population of this	Records the investigators opinion as to the caus	3.2	3.1.3
12	10	Update	CDISC Notes and Definitions	AE	AETOXGR	Toxicity grade according to a standard toxicity scale such as Common Terminology Criteria for Adverse Events v3.0 (CTCAE). Sponsor should specify name of the scale and version used in the metadata (see Assumption 6d). If value is from a numeric scale, represent only the number (e.g., 2 and	Toxicity grade according to a standard toxicity s	3.2	3.1.3
13	11	Update	Controlled Terminology	CE	CEENRTPT	(STENRF)	BEFORE, AFTER, COINCIDENT, ONGOING, U	3.2	3.1.3
14	12	Update	Role	CE	CEPRESP	Variable Qualifier	Record Qualifier	3.2	3.1.3
15	13	Update	Controlled Terminology	CE	CESTRTPT	(STENRF)	BEFORE, AFTER, COINCIDENT, U	3.2	3.1.3
16	14	Add	Controlled Terminology	DS	EPOCH	(EPOCH)		3.2	3.1.3
17	15	Update	CDISC Notes and Definitions	DS	EPOCH	EPOCH may be used when DSCAT = DISPOSITION EVENT. Examples: SCREENING, TREATMENT, FOLLOWUP	EPOCH may be used when DSCAT = DISPOSITI	3.2	3.1.3
18	16	Add	Controlled Terminology	DV	EPOCH	(EPOCH)		3.2	3.1.3
19	17	Add	Domain	HO				3.2	3.1.3
20	18	Update	CDISC Notes and Definitions	MH	MHDECOD	Dictionaryderived text description of MHTERM or MHMODIFY. Equivalent to the Preferred Term (PT in MedDRA). The sponsor is expected to provide the dictionary name and version used to map the terms utilizing the define.xml external codelist attributes.	Dictionaryderived text description of MHTERM or	3.2	3.1.3
21	19	Update	Controlled Terminology	MH	MHENRTPT	(ENRTPT)	BEFORE, AFTER, COINCIDENT, ONGOING, U	3.2	3.1.3
22	20	Update	Role	MH	MHPRESP	Variable Qualifier of --TERM	Record Qualifier	3.2	3.1.3
23	21	Update	Variable Label	MH	MHREASND	Reason Medical History Not Done or Not Occurred	Reason Medical History Not Collected	3.2	3.1.3

B. USING SPREADSHEET COMPARE TOOL [OFFICE 2013]

Another way of comparing the metadata published by eSHARE is described below:

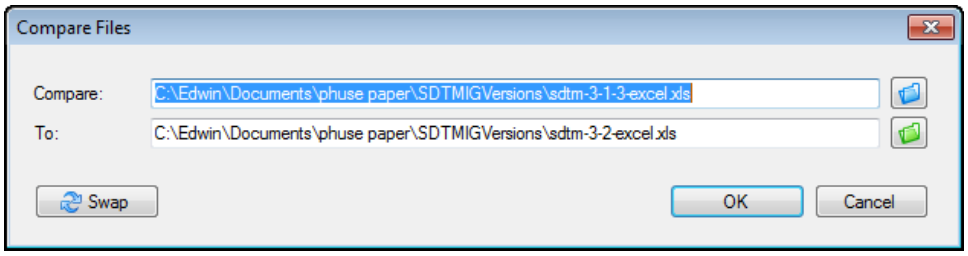
1. Download SDTM IG metadata excel spreadsheet from eSHARE page. By this, two different excel file is downloaded namely, SDTM 3.1.3 and SDTM 3.2 metadata sheets.

SDTM 1.4 (IG 3.2)				
2015-02-02	SDTM IG Diff	3.2 - 3.1.3	Metadata	Excel
2014-08-21	SDTM IG	1.4 (3.2)	Metadata	Define-XML v2.0 | Define-XML v1.0 | Excel
2014-08-21	SDTM + CT	1.4+2013-12-20	Bundle	Define-XML v2.0 | Excel
2014-08-21	SDTM IG	1.3 (3.1.3)	Document Bundle	PDF
SDTM 1.3 (IG 3.1.3)				
2014-08-21	SDTM IG	1.3 (3.1.3)	Metadata	Define-XML v2.0 | Define-XML v1.0 | Excel
2014-08-21	SDTM + CT	1.3+2013-12-20	Bundle	Define-XML v2.0 | Excel
2014-08-21	SDTM IG	1.3 (3.1.3)	Document Bundle	PDF

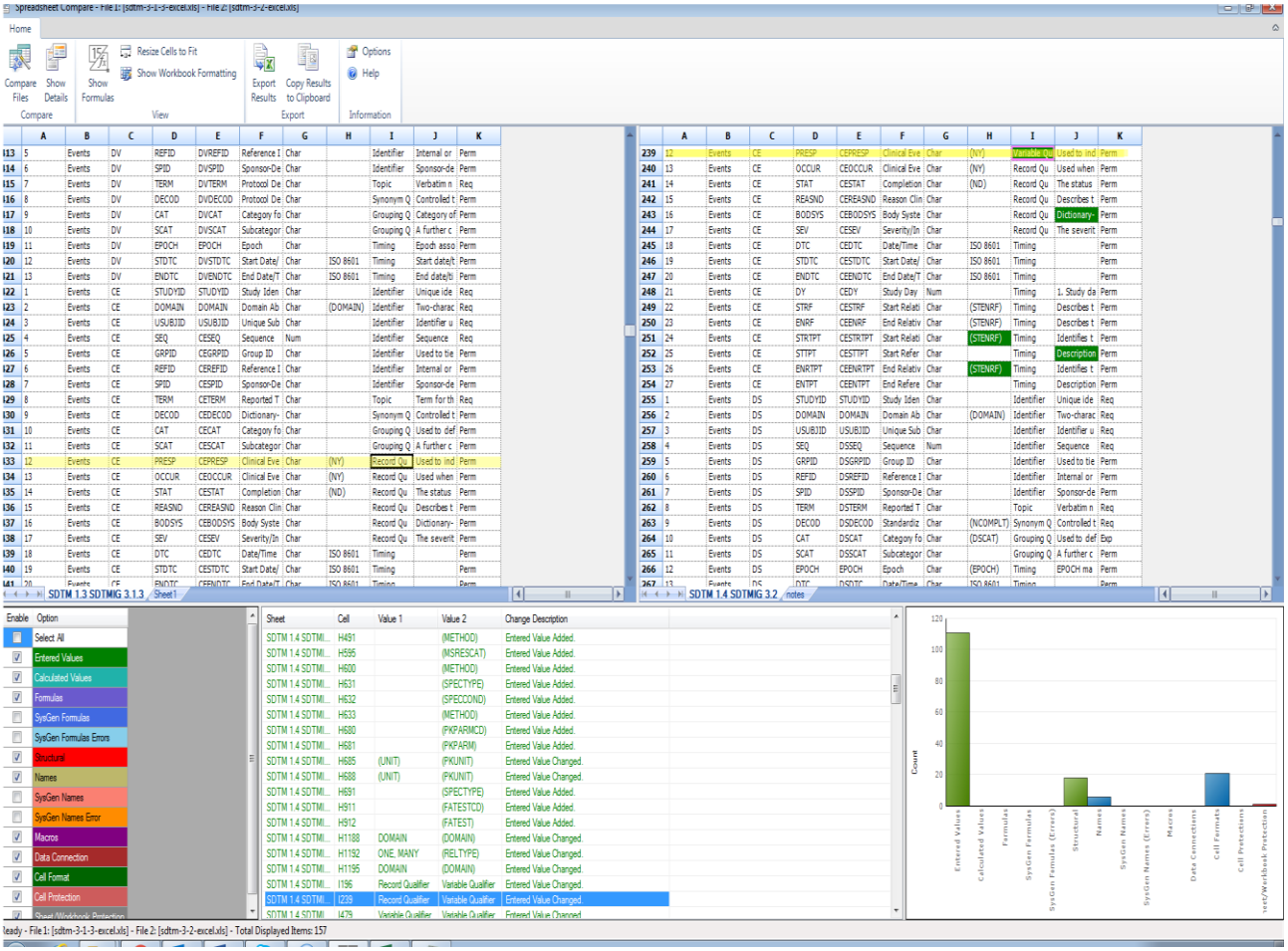
2. Open Spreadsheet Compare → In Start menu, go to Office 2013 Tools, and click Spreadsheet Compare.
3. Click Home → Compare Files.

Compare Files dialog box appears.

4. Click the blue folder icon next to the Compare box to browse the location of older SDTM version. In addition to files saved locally on your computer, it is possible to compare files on a network.



5. Click the green folder icon next to the 'To' box to browse the location of the workbook that you want to compare to the earlier version, and then click OK.
6. In the left pane, choose the compare options you want to see in the results of the workbook comparison by checking or unchecking the options, such as Formulas, Macros, or Cell Format. Or, just Select All.
7. Click Ok to run the comparison.



In the above example shown, results for domain CE.CEPRESP variable role changed from Record Qualifier to Variable Qualifier in the SDTM 3.2 version.

If the compare tool does not provide the accurate result across all domains, then split the excel file domain-wise and compare it.

C. USING SAS PROGRAM OR MACRO

Other option available for a SAS programmer is to develop a macro to compare two versions and display the differences in a customized report form.

1. Display all the newly added domains
2. Display all the variables assigned for the newly added domains
3. Display all the variables names in which added into the existing domains
4. Display all the deleted variables from the existing domains
5. Display all the variables names in which attributes updated

Below is a simple SAS code for listing the added and deleted details from two versions. This example uses the eSHARE metadata excel files as an input.

```
*** Import SDTM IG 3.1.3 version;
libname sdtm "C:\Edwin\Documents\phuse paper\SDTMIGVersions";
proc import datafile="C:\Edwin\Documents\phuse paper\SDTMIGVersions\sdtm-3-1-3-
excel.xls" out=sdtm.sdtm313 dbms=xls replace;
run;
quit;

*** Import SDTM IG 3.2 version;
libname sdtm "C:\Edwin\Documents\phuse paper\SDTMIGVersions\sdtm-3-1-3";
proc import datafile="C:\Edwin\Documents\phuse paper\SDTMIGVersions\sdtm-3-2-
excel.xls" out=sdtm.sdtm32 dbms=xls replace;
```

```
run;
quit;

*** Sorting by domain and variables;
proc sort data=sdtm.sdtm313 out=sdtm313 nodup;
      by Domain_Prefix Variable_Name;
run;

proc sort data=sdtm.sdtm32 out=sdtm32 nodup;
      by Domain_Prefix Variable_Name;
run;

*** To find the list of new variables added [in tdnew] and deleted [in tdold] in the
SDTM IG Versions;
data tdnew tdold ;
      merge sdtm32(in=nw) sdtm313(in=od);
      by Domain_Prefix Variable_Name;
      if nw & not(od) then output tdnew;
      else if od & not(nw) then output tdold;
run;

proc sort data=tdnew out=tdnew;
      by Domain_Prefix Seq__For_Order Variable_Name;
run;

proc sort data=tdold out=tdold;
      by Domain_Prefix Seq__For_Order Variable_Name;
run;

*** Export the output details in to excel;
proc export data=tdnew outfile="C:\Edwin\Documents\phuse
paper\SDTMIGVersions\Added.xls" dbms=xls replace;
run;
quit;

proc export data=tdold outfile="C:\Edwin\Documents\phuse
paper\SDTMIGVersions\Deleted.xls" dbms=xls replace;
run;
quit;
```

PROBLEM STATEMENT

A Study starts with initial prerequisite items filled up under the advice of the Sponsor, which includes the adoption of SDS version for the submission. During the course of the study, there can be changes initiated by FDA or Sponsor themselves would request the study team to make amendments to submission plan by requesting to work on a different SDS version for submission. Such situations would occur in late phase trials where the study duration exceeds a month. At times, it would impact early phase trials which run for few weeks.

Under such circumstances, the study team would have to make amendments for the ongoing study which would have been started, configured and even output created in one version to a different version. The amendment varies from output dataset handling to Define-XML creation.

Going forward, this paper would be taking up a case where the SDS version changes from v3.1.3 to v3.2 and Define-XML changes from v1.0 to v2.0 for an ongoing study.

THE SOLUTION IS…EXACT™

Many solutions were discussed in earlier section to perform a compare of versions and then, proceed manually to implement the changes and rerun the program set. This paper deals with a semi-automated approach with a demonstration outlining the approach & process to make the change with ease and still maintain the quality in less time. Proposed demonstration is done using an enterprise class system: EXACT.

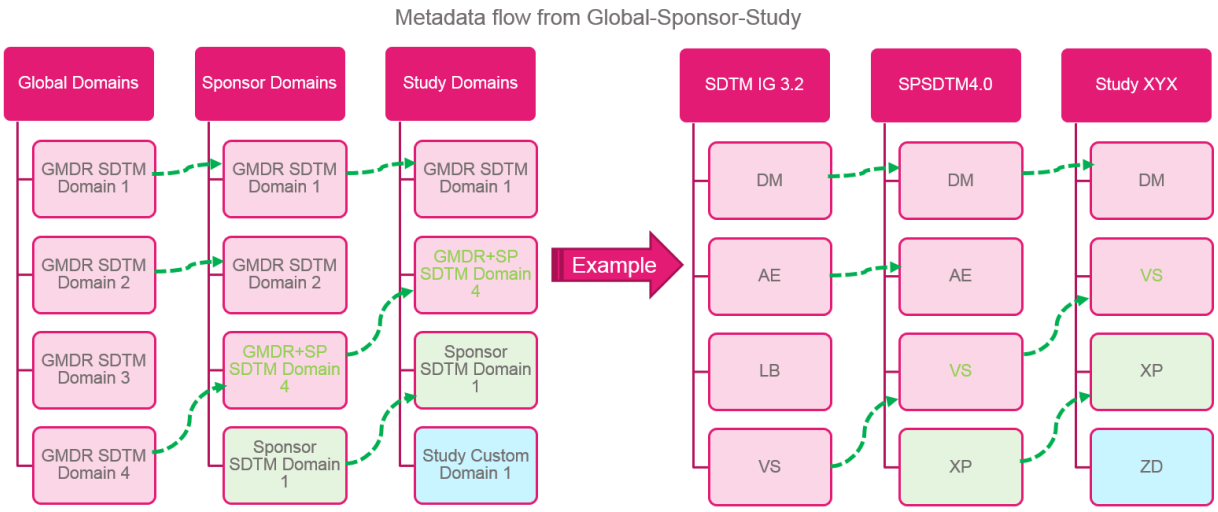
EXtraction, Analysis, Conversion and Tabulation (EXACT) is an outcome of integrating clinical and non-clinical data from multiple sources into a single environment where it can be standardized, analyzed, visualized and reported on by clinical researchers to regulatory agencies in shorter span of time from last observation last visit.

EXACT brings a complete set of process automation using the implementation of Clinical Meta Data Repository model across different stages of clinical development by enabling the user to extract, view, transform, analyze and report data across clinical applications, and across projects or trials in different therapeutics within the same environment and in a regulatory compliant manner. This is achieved in EXACT by extracting data from various sources that are available for a study and creating a fully controlled, traceable, full GCP and 21 CFR Part 11 compliant data for various analyses & within modules that are integrated in the product.

MANAGE CDISC METADATA & SPONSOR METADATA

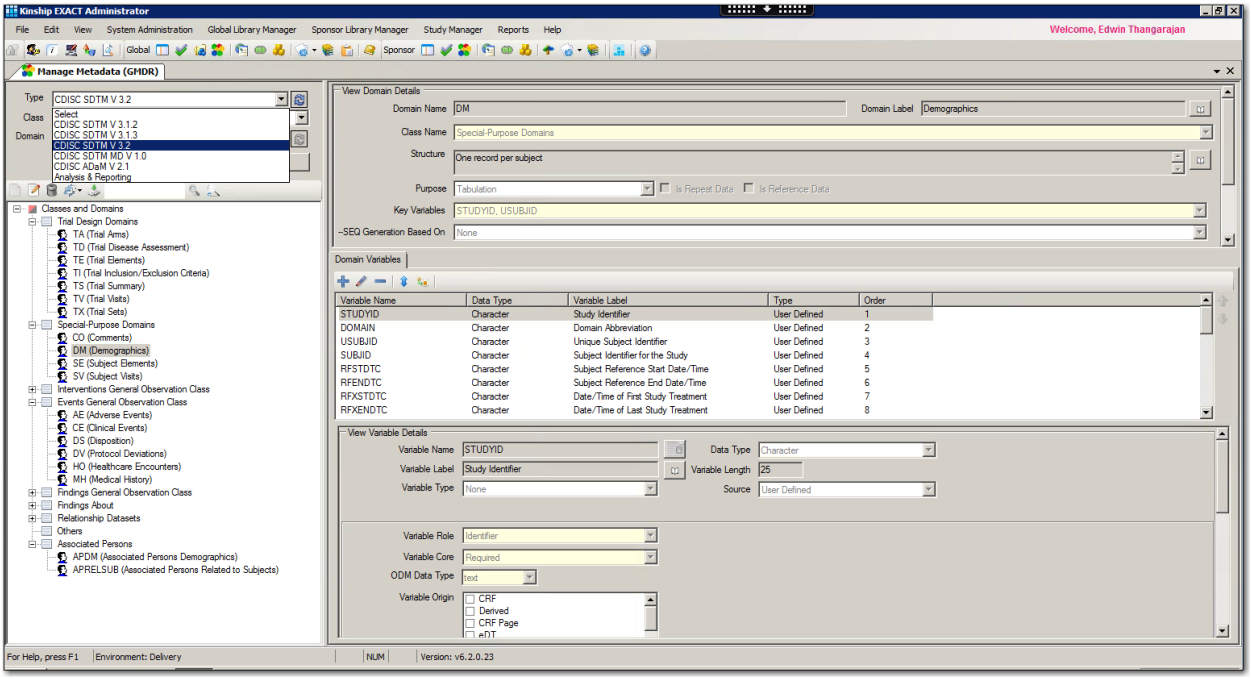
EXACT™ stores and manage all the CDISC published metadata, sponsor metadata with additional attributes which requires for the SDTM dataset and Define-XML generation into its repository. If any new SDTM IG version released by CDISC or Sponsor, it can be easily imported into the system if there is no structural changes. These metadata details will be re-used under each study based on the SDTM IG or Sponsor version. These metadata is stored under EXACT Administrator module and managed by the standards group and the standards librarian within our organization.

Below diagram shows how metadata will be inherited from the global level to below levels namely, sponsor, study.



The following SDTM versions are currently available and loaded into EXACT.

- SDTM IG 3.1.1 (outdated)
- SDTM IG 3.1.2
- SDTM IG 3.1.3
- SDTM IG 3.2
- SDTM IG AP 1.0
- SDTM IG Medical Devices 1.0
- SDTM Terminologies



User Interface 1: Manage Metadata (GMDR)

EXACT facilitates the use of global level version or it permits the use of sponsor version for the study associated with that sponsor. Both the versions are available in library.

HANDLING SDTM ONE VERSION TO ANOTHER IN A STUDY

This section of the paper provides the necessary approach & steps that need to be followed in a semi-automated fashion to change from one SDS version 3.1.3 to v3.2. It is understood that EXACT has both the versions that is required for the ongoing study and one of the version has been used for generating the output dataset. But, with the current change in requirement, it is required to generate the output SDTM dataset in a new version.

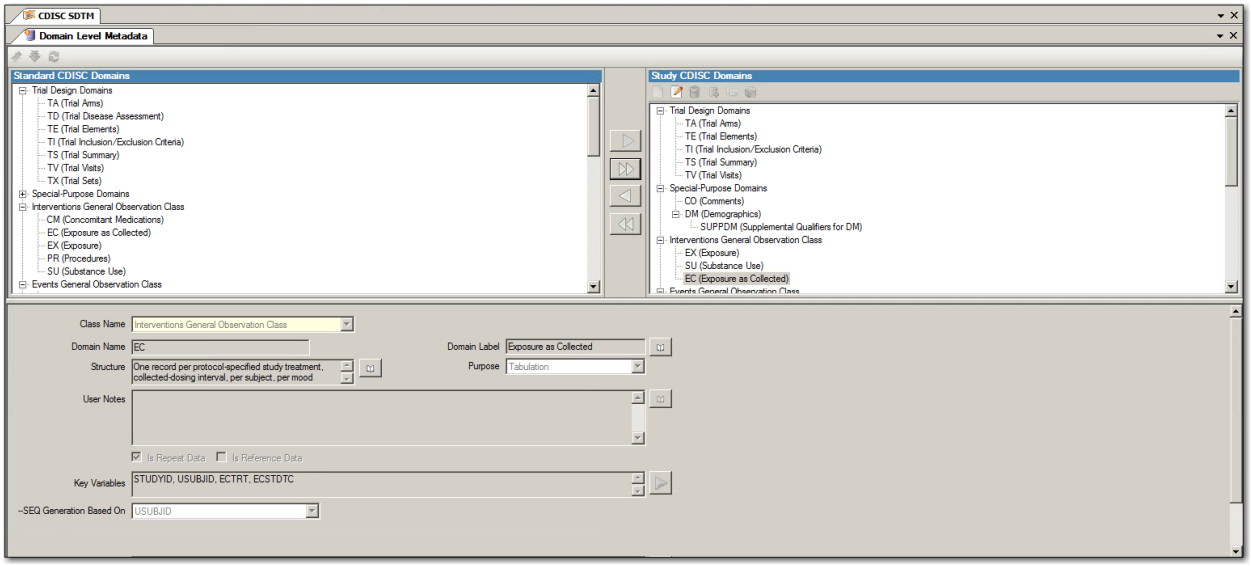
The following steps are to be followed to handle different versions of SDTM in a single study in EXACT.

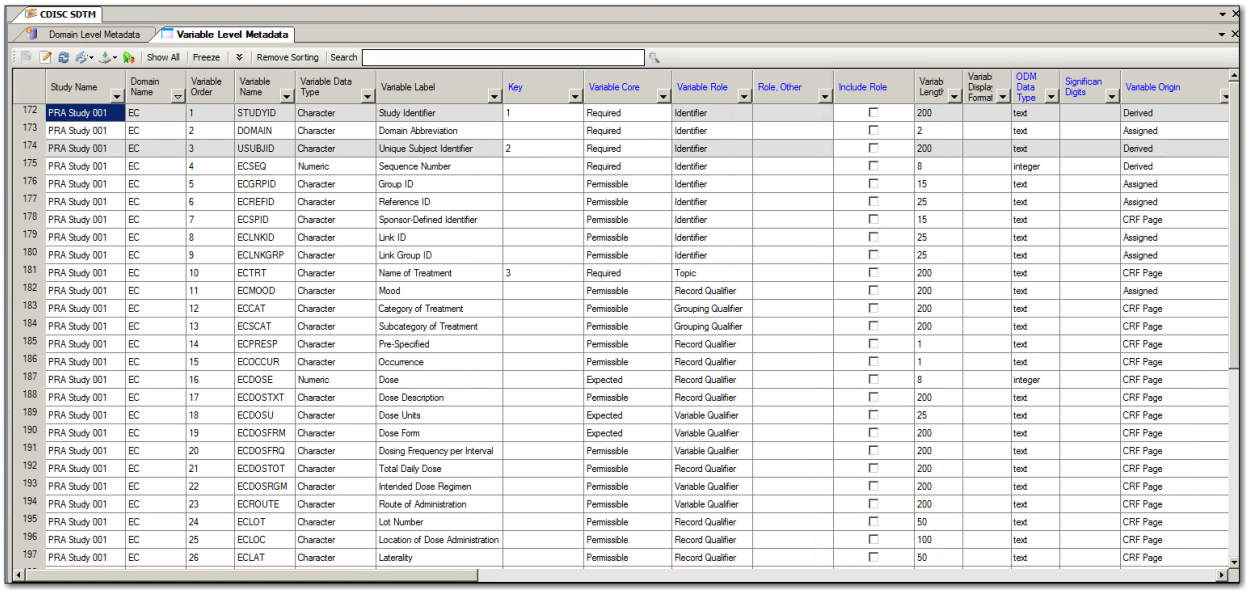
NEW DOMAIN INCLUSION

This scenario is about handling inclusion of new domain into the study when the ongoing study has been configured and output generated using v3.1.3. In this circumstance, the new domain was made available in v3.2 whereas the trial data had information during study capture and was accommodated in a single domain within v3.1.3 structure.

The study team has been instructed to use the new domain available in the new version to move the additional information captured in trial data.

EXACT has the feature in Domain Level Metadata (DLM) UI to move the domains from CDISC/Sponsor repository to study level which in turns makes all the metadata available for the study. Variable Level Metadata (VLM) UI is shown to display all the variables post move from Sponsor to Study using DLM UI.





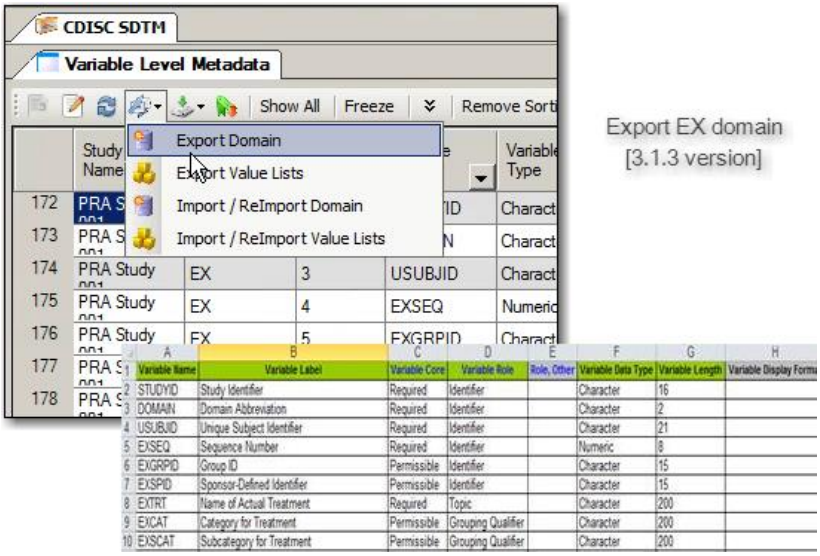
	Study Name	Domain Name	Variable Order	Variable Name	Variable Data Type	Variable Label	Key	Variable Core	Variable Role	Role, Other	Include Role	Variab Length	Variab Display Format	ODM Data Type	Significan Digits	Variable Origin
172	PRA Study 001	EC	1	STUDYID	Character	Study Identifier	1	Required	Identifier			200		text		Derived
173	PRA Study 001	EC	2	DOMAIN	Character	Domain Abbreviation		Required	Identifier			2		text		Assigned
174	PRA Study 001	EC	3	USUBJID	Character	Unique Subject Identifier	2	Required	Identifier			200		text		Derived
175	PRA Study 001	EC	4	ECSEQ	Numeric	Sequence Number		Required	Identifier			8		integer		Derived
176	PRA Study 001	EC	5	ECGRPID	Character	Group ID		Permissible	Identifier			15		text		Assigned
177	PRA Study 001	EC	6	ECREFID	Character	Reference ID		Permissible	Identifier			25		text		Assigned
178	PRA Study 001	EC	7	ECSPID	Character	Sponsor-Defined Identifier		Permissible	Identifier			15		text		CRF Page
179	PRA Study 001	EC	8	ECLNKID	Character	Link ID		Permissible	Identifier			25		text		Assigned
180	PRA Study 001	EC	9	ECLNKGRP	Character	Link Group ID		Permissible	Identifier			25		text		Assigned
181	PRA Study 001	EC	10	ECTRT	Character	Name of Treatment	3	Required	Topic			200		text		CRF Page
182	PRA Study 001	EC	11	ECMOOD	Character	Mood		Permissible	Record Qualifier			200		text		Assigned
183	PRA Study 001	EC	12	ECCAT	Character	Category of Treatment		Permissible	Grouping Qualifier			200		text		CRF Page
184	PRA Study 001	EC	13	ECSCAT	Character	Subcategory of Treatment		Permissible	Grouping Qualifier			200		text		CRF Page
185	PRA Study 001	EC	14	ECPRESP	Character	Pre-Specified		Permissible	Record Qualifier			1		text		CRF Page
186	PRA Study 001	EC	15	ECOCCUR	Character	Occurrence		Permissible	Record Qualifier			1		text		CRF Page
187	PRA Study 001	EC	16	ECDOSE	Numeric	Dose		Expected	Record Qualifier			8		integer		CRF Page
188	PRA Study 001	EC	17	ECDOSTXT	Character	Dose Description		Permissible	Record Qualifier			200		text		CRF Page
189	PRA Study 001	EC	18	ECDOSU	Character	Dose Units		Expected	Variable Qualifier			25		text		CRF Page
190	PRA Study 001	EC	19	ECDOSFRM	Character	Dose Form		Expected	Variable Qualifier			200		text		CRF Page
191	PRA Study 001	EC	20	ECDOSFRQ	Character	Dosing Frequency per Interval		Permissible	Variable Qualifier			200		text		CRF Page
192	PRA Study 001	EC	21	ECDOSTOT	Character	Total Daily Dose		Permissible	Record Qualifier			200		text		CRF Page
193	PRA Study 001	EC	22	ECDOSRGM	Character	Intended Dose Regimen		Permissible	Variable Qualifier			200		text		CRF Page
194	PRA Study 001	EC	23	ECROUTE	Character	Route of Administration		Permissible	Variable Qualifier			200		text		CRF Page
195	PRA Study 001	EC	24	ECLOT	Character	Lot Number		Permissible	Record Qualifier			50		text		CRF Page
196	PRA Study 001	EC	25	ECLOC	Character	Location of Dose Administration		Permissible	Record Qualifier			100		text		CRF Page
197	PRA Study 001	EC	26	ECLAT	Character	Laterality		Permissible	Record Qualifier			50		text		CRF Page

User Interface 2: DLM and VLM

In addition to the above, the system provides a feature to import the domain & variable level information at study level.

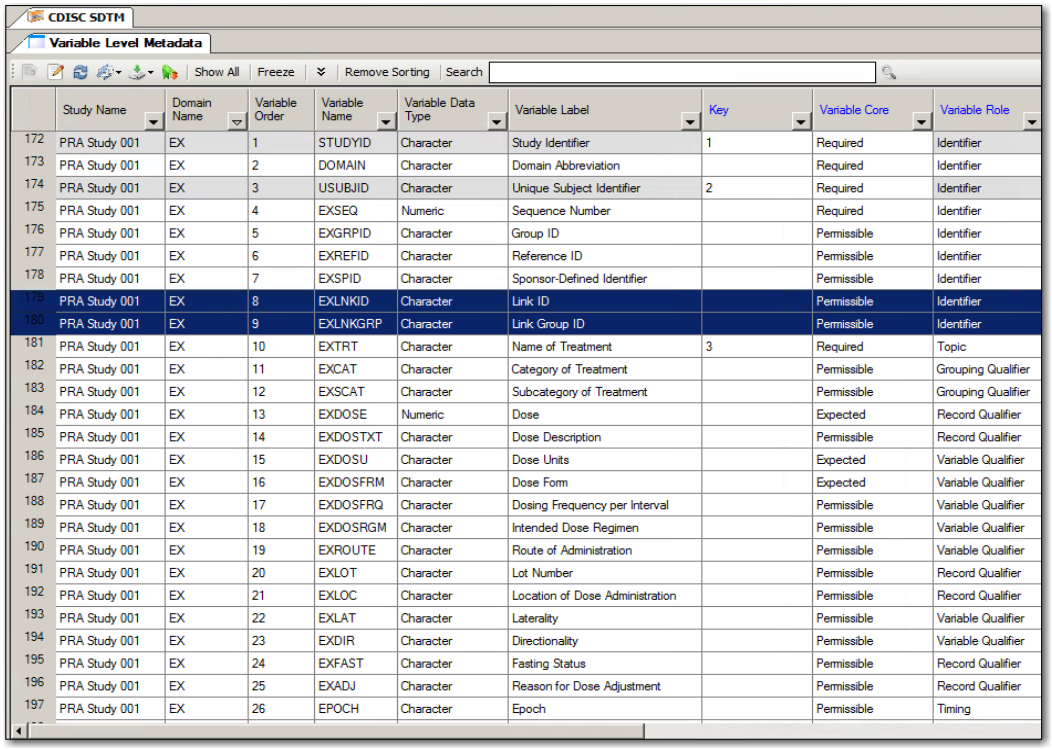
NEW VARIABLES AND UPDATES IN THE EXISTING DOMAIN/VARIABLES

EXACT has the feature to export the existing configured domain metadata and mapping definitions as excel spreadsheet and make necessary modifications and re-import quickly without changing the unchanged metadata and mapping definitions.



User Interface 3: VLM along with Export Sheet

1. Make necessary changes to the exported metadata (add/modify) in the spreadsheet for the SDTM 3.2 usage for a study or use the SDTM 3.2 metadata sheet (Example: v3.2 EX variables 'EXLNKID', 'EXLNKGP' to be added)
2. Select Import / Re Import option to re-import the specific domain.



	Study Name	Domain Name	Variable Order	Variable Name	Variable Data Type	Variable Label	Key	Variable Core	Variable Role
172	PRA Study 001	EX	1	STUDYID	Character	Study Identifier	1	Required	Identifier
173	PRA Study 001	EX	2	DOMAIN	Character	Domain Abbreviation		Required	Identifier
174	PRA Study 001	EX	3	USUBJID	Character	Unique Subject Identifier	2	Required	Identifier
175	PRA Study 001	EX	4	EXSEQ	Numeric	Sequence Number		Required	Identifier
176	PRA Study 001	EX	5	EXGRPID	Character	Group ID		Permissible	Identifier
177	PRA Study 001	EX	6	EXREFID	Character	Reference ID		Permissible	Identifier
178	PRA Study 001	EX	7	EXSPID	Character	Sponsor-Defined Identifier		Permissible	Identifier
179	PRA Study 001	EX	8	EXLNKID	Character	Link ID		Permissible	Identifier
180	PRA Study 001	EX	9	EXLNKGRP	Character	Link Group ID		Permissible	Identifier
181	PRA Study 001	EX	10	EXTRT	Character	Name of Treatment	3	Required	Topic
182	PRA Study 001	EX	11	EXCAT	Character	Category of Treatment		Permissible	Grouping Qualifier
183	PRA Study 001	EX	12	EXSCAT	Character	Subcategory of Treatment		Permissible	Grouping Qualifier
184	PRA Study 001	EX	13	EXDOSE	Numeric	Dose		Expected	Record Qualifier
185	PRA Study 001	EX	14	EXDOSTXT	Character	Dose Description		Permissible	Record Qualifier
186	PRA Study 001	EX	15	EXDOSU	Character	Dose Units		Expected	Variable Qualifier
187	PRA Study 001	EX	16	EXDOSFRM	Character	Dose Form		Expected	Variable Qualifier
188	PRA Study 001	EX	17	EXDOSFRQ	Character	Dosing Frequency per Interval		Permissible	Variable Qualifier
189	PRA Study 001	EX	18	EXDOSRGM	Character	Intended Dose Regimen		Permissible	Variable Qualifier
190	PRA Study 001	EX	19	EXROUTE	Character	Route of Administration		Permissible	Variable Qualifier
191	PRA Study 001	EX	20	EXLOT	Character	Lot Number		Permissible	Record Qualifier
192	PRA Study 001	EX	21	EXLOC	Character	Location of Dose Administration		Permissible	Record Qualifier
193	PRA Study 001	EX	22	EXLAT	Character	Laterality		Permissible	Variable Qualifier
194	PRA Study 001	EX	23	EXDIR	Character	Directionality		Permissible	Variable Qualifier
195	PRA Study 001	EX	24	EXFAST	Character	Fasting Status		Permissible	Record Qualifier
196	PRA Study 001	EX	25	EXADJ	Character	Reason for Dose Adjustment		Permissible	Record Qualifier
197	PRA Study 001	EX	26	EPOCH	Character	Epoch		Permissible	Timing

User Interface 4: VLM

The export feature can be done for a single domain or more than one domain and the same can be reimported for a study. While doing the above steps, it would cover the variable level changes such as new variables or dropping of variables or change in attributes of a variable.

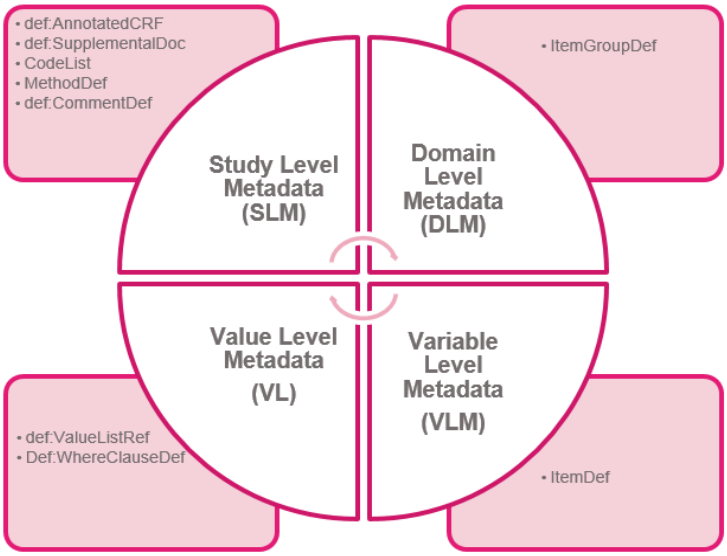
Further to this, changes associated with inclusion of new variable would involve mapping configuration changes. This is done either through the exported sheet which contains the mapping information or post re-import by way of accessing the mapping user interface and make necessary configurations.

HANDLING DEFINE-XML VERSION 1.0 TO DEFINE-XML 2.0 IN A STUDY

It is informed by FDA that the end of support for Define-XML v1.0 for studies is March, 2017. In such case, most of the sponsor studies which were ongoing with Define-XML v1.0 has to plan for migration to Define-XML v2.0 based on the study duration. This situation is easily handled through the system to generate output datasets and define.xml file generation in the new version recommended.

The below data flow diagram represents the user interfaces and how the data flows to define file sections.

EXACT UI structure



Define-XML document structure:

```
<?xml version="1.0" encoding="UTF-8"?>
<ODM xmlns="http://www.cdisc.org/ns/odm/v1.3"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:def="http://www.cdisc.org/ns/def/v2.0"
  FileType="Snapshot" ODMVersion="1.3.2"
  FileOID="Studycdisc01-Define-XML_2.0.0"
  CreationDateTime="2012-06-21T11:07:23-05:00"
  Originator="CDISC XML Technologies Team">
  <Study OID="cdisc01">
    <GlobalVariables>
      <StudyName>CDISC01</StudyName>
      <StudyDescription>CDISC Test Study</StudyDescription>
      <ProtocolName>CDISC 01</ProtocolName>
    </GlobalVariables>
    <MetaDataVersion OID="MDV.CDISC01.SDTMIG.3.1.2.SDTM.1.2"
      Name="Study CDISC01, Data Definitions"
      Description="Study CDISC01, Data Definitions"
      def:DefineVersion="2.0.0"
      def:StandardName="CDISC SDTM"
      def:StandardVersion="3.1.2">
        < Annotated Case Report Forms (def:AnnotatedCRF) >
        < Supplemental Data Definitions (def:SupplementalDoc) >
        < Value Level Metadata (def:ValueListDef) >
        < Where Clause Definitions (def:WhereClauseDef) >
        < Domain Level Metadata (ItemGroupDef) >
        < Variable Level Metadata (ItemDef) >
        < Controlled Terminology Metadata (CodeList) >
        < Computational Algorithms (MethodDef) >
        < Comments (def:CommentDef) >
        < Referenced Documents (def:leaf) >
    </MetaDataVersion>
  </Study>
</ODM>
```

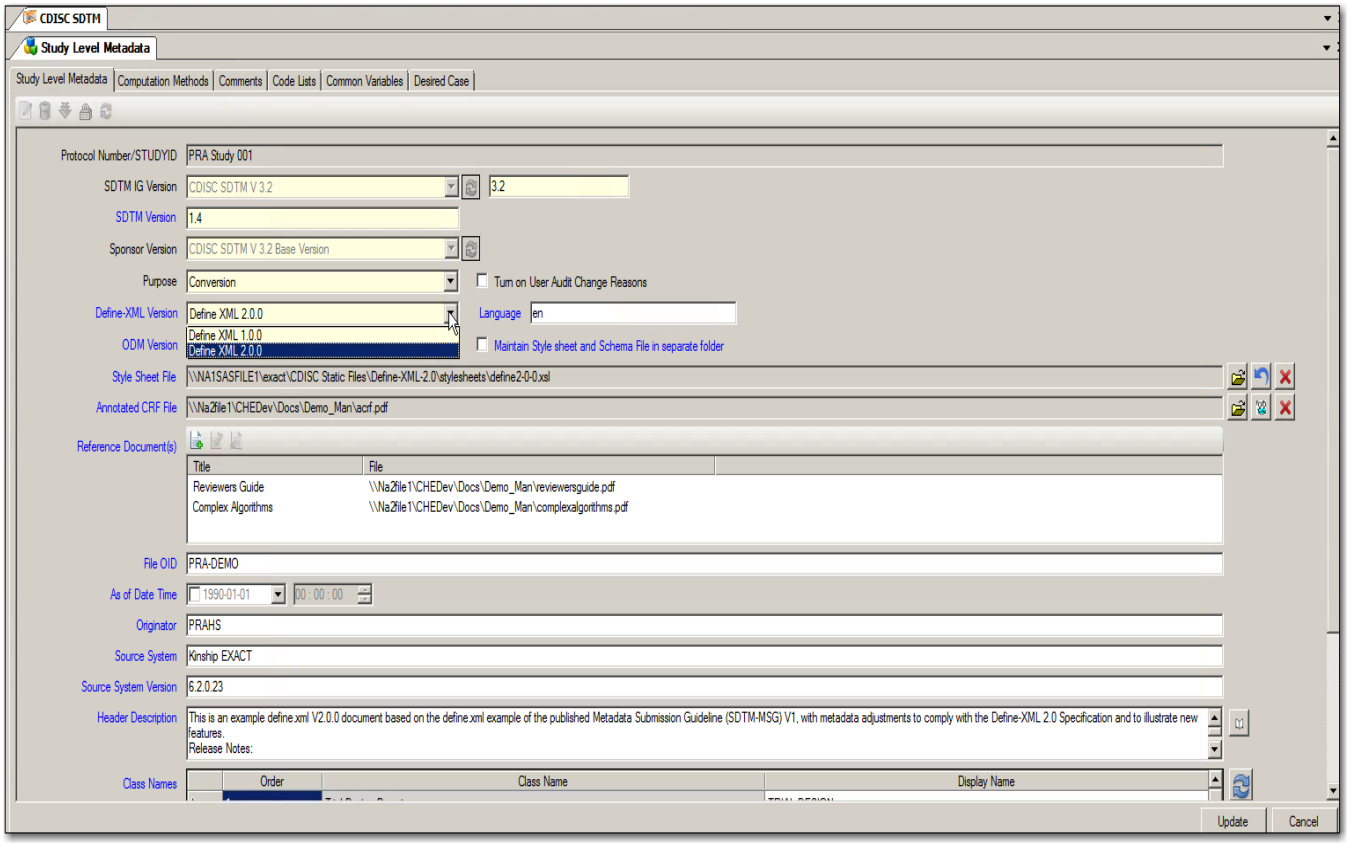
The key metadata components to support submissions are:

- Dataset definitions
- Dataset variable definitions
- Controlled Terminology definitions
- Value list definitions
- Links to supporting documents
- Computational method definitions
- Comments definitions

The below details narrates the configuration steps to be followed by the user in the system to generate the define.xml in the new version. Following sections illustrate the steps to be followed in each of the UIs provided in EXACT to handle the different tags in the new define,xml file.

STUDY LEVEL METADATA

Study Level Metadata (SLM) UI is used to switch the Define-XML 1.0 to Define-XML 2.0. The system automatically initiates the change in metadata information for the specific study once the user switches to a different define version at the time of saving the UI. In addition to change, user needs to configure additional information such as value list, comments for the target define version which may not be available in the previous version.



User Interface 5: SLM

The SLM UI capture the below details,

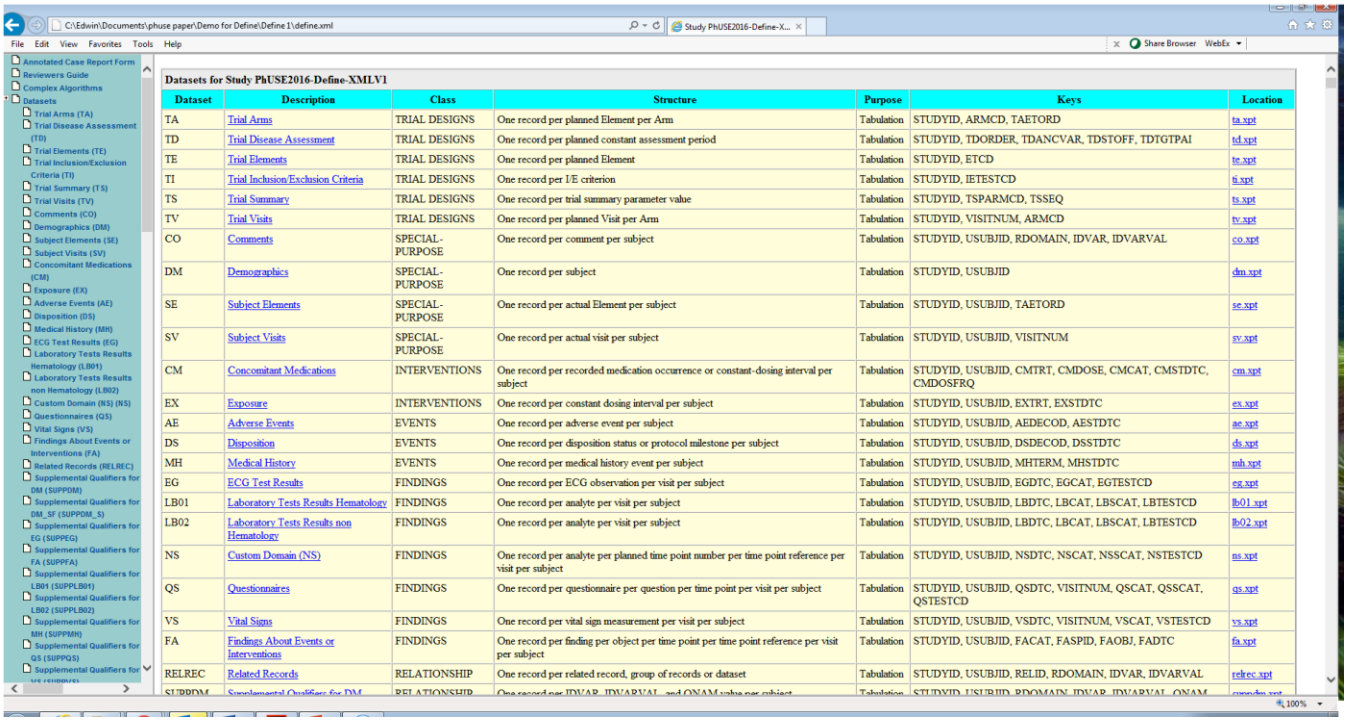
- ODM Attributes such as stylesheet information, File Type, File OID, Creation Date Time and Originator
- Study Global Variables
- Metadata Version Details such as Define-XML version, Standard Name, Standard Version
- Annotated Case Report Forms
- Reference Documents such as reviewer's guide, computational algorithms etc.,
- Display order and names of the Classes

DOMAIN LEVEL METADATA

CDISC SDTM study domains are modeled as tables where the columns represent variables and the rows represent observed or derived values of those variables. Domain level metadata is represented in Define-XML as an ItemGroupDef element for both version 1 and 2. The system permits the user to view with additional information pertaining to domain level changes using the DLM UI. DLM UI with details is shown in the User Interface 2.

Below screen snapshots depicts domain level information of two versions of define.xml file for better clarity.

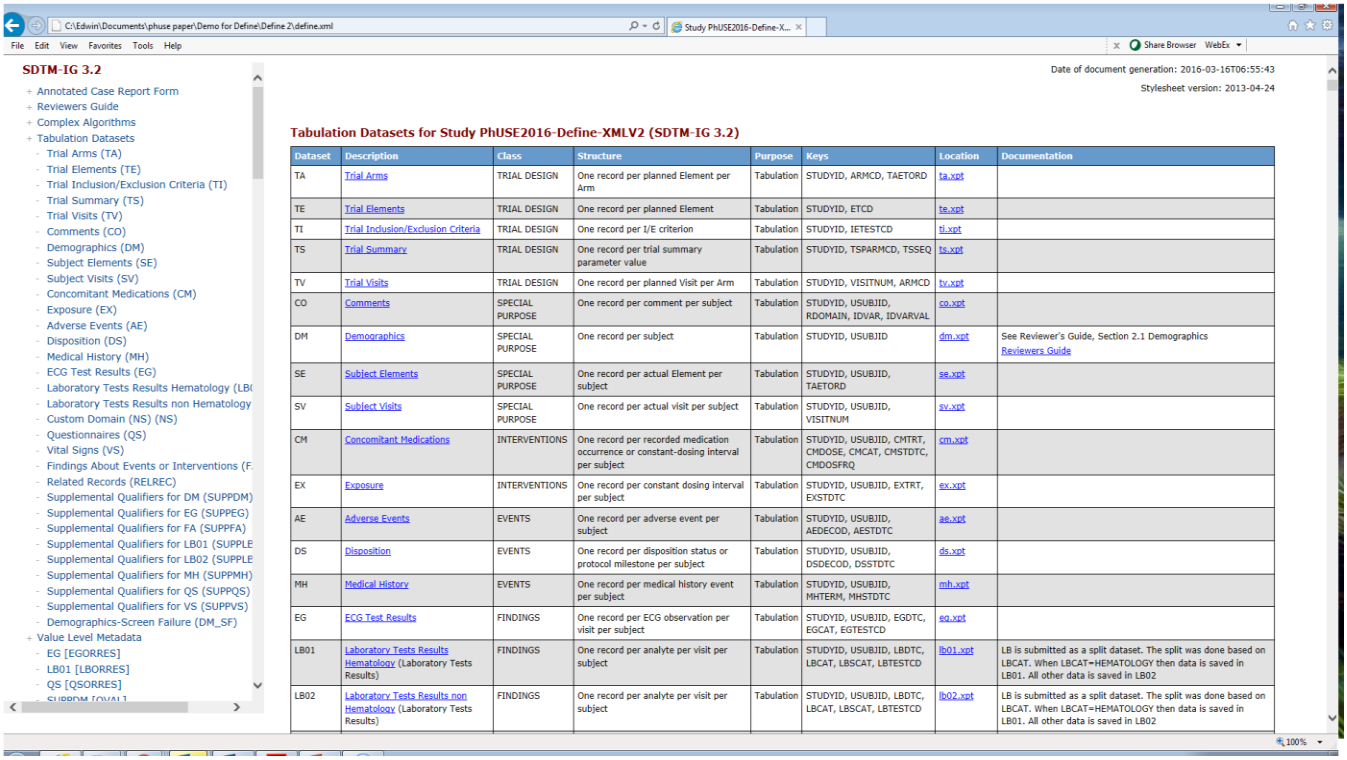
Define-XML Version 1 [Domain Level Metadata]



Datasets for Study PhUSE2016-Define-XMLV1

Dataset	Description	Class	Structure	Purpose	Keys	Location
TA	Trial Arms	TRIAL DESIGNS	One record per planned Element per Arm	Tabulation	STUDYID, ARMCD, TAETORD	ta.xpt
TD	Trial Disease Assessment	TRIAL DESIGNS	One record per planned constant assessment period	Tabulation	STUDYID, TDORDER, TDANCVAR, TDSTOFF, TDTGTPAI	td.xpt
TE	Trial Elements	TRIAL DESIGNS	One record per planned Element	Tabulation	STUDYID, ETCD	te.xpt
TI	Trial Inclusion/Exclusion Criteria	TRIAL DESIGNS	One record per I/E criterion	Tabulation	STUDYID, IETESTCD	ti.xpt
TS	Trial Summary	TRIAL DESIGNS	One record per trial summary parameter value	Tabulation	STUDYID, TSPARMCD, TSSEQ	ts.xpt
TV	Trial Visits	TRIAL DESIGNS	One record per planned Visit per Arm	Tabulation	STUDYID, VISITNUM, ARMCD	tv.xpt
CO	Comments	SPECIAL-PURPOSE	One record per comment per subject	Tabulation	STUDYID, USUBJID, RDOMAIN, IDVAR, IDVARVAL	co.xpt
DM	Demographics	SPECIAL-PURPOSE	One record per subject	Tabulation	STUDYID, USUBJID	dm.xpt
SE	Subject Elements	SPECIAL-PURPOSE	One record per actual Element per subject	Tabulation	STUDYID, USUBJID, TAETORD	se.xpt
SV	Subject Visits	SPECIAL-PURPOSE	One record per actual visit per subject	Tabulation	STUDYID, USUBJID, VISITNUM	sv.xpt
CM	Concomitant Medications	INTERVENTIONS	One record per recorded medication occurrence or constant-dosing interval per subject	Tabulation	STUDYID, USUBJID, CMTRT, CMDOSE, CMCAT, CMSTDTC, CMDOSFRQ	cm.xpt
EX	Exposure	INTERVENTIONS	One record per constant dosing interval per subject	Tabulation	STUDYID, USUBJID, EXTRT, EXSTDTC	ex.xpt
AE	Adverse Events	EVENTS	One record per adverse event per subject	Tabulation	STUDYID, USUBJID, AEDECOD, AESTDTC	ae.xpt
DS	Disposition	EVENTS	One record per disposition status or protocol milestone per subject	Tabulation	STUDYID, USUBJID, DSDECOD, DSSTDTC	ds.xpt
MH	Medical History	EVENTS	One record per medical history event per subject	Tabulation	STUDYID, USUBJID, MHTERM, MHSTDTC	mh.xpt
EG	ECG Test Results	FINDINGS	One record per ECG observation per visit per subject	Tabulation	STUDYID, USUBJID, EGDTC, EGCAT, EGTESTCD	eg.xpt
LB01	Laboratory Tests Results Hematology	FINDINGS	One record per analyte per visit per subject	Tabulation	STUDYID, USUBJID, LBDTC, LBCAT, LBSCAT, LBTESTCD	lb01.xpt
LB02	Laboratory Tests Results non Hematology	FINDINGS	One record per analyte per visit per subject	Tabulation	STUDYID, USUBJID, LBDTC, LBCAT, LBSCAT, LBTESTCD	lb02.xpt
NS	Custom Domain (NS)	FINDINGS	One record per analyte per planned time point number per time point reference per visit per subject	Tabulation	STUDYID, USUBJID, NSDTC, NSCAT, NSSCAT, NSTESTCD	ns.xpt
QS	Questionnaires	FINDINGS	One record per questionnaire per question per time point per visit per subject	Tabulation	STUDYID, USUBJID, QSDTC, VISITNUM, QSCAT, QSSCAT, QSTESTCD	qs.xpt
VS	Vital Signs	FINDINGS	One record per vital sign measurement per visit per subject	Tabulation	STUDYID, USUBJID, VSDTC, VISITNUM, VSCAT, VSTESTCD	vs.xpt
FA	Findings About Events or Interventions	FINDINGS	One record per finding per object per time point per time point reference per visit per subject	Tabulation	STUDYID, USUBJID, FACAT, FASPID, FAOBJ, FADTC	fa.xpt
RELREC	Related Records	RELATIONSHIP	One record per related record, group of records or dataset	Tabulation	STUDYID, USUBJID, RELID, RDOMAIN, IDVAR, IDVARVAL	relrec.xpt

Define-XML Version 2 [Domain Level Metadata]



SDTM-IG 3.2

Date of document generation: 2016-03-16T06:55:43

Stylesheet version: 2013-04-24

Tabulation Datasets for Study PhUSE2016-Define-XMLV2 (SDTM-IG 3.2)

Dataset	Description	Class	Structure	Purpose	Keys	Location	Documentation
TA	Trial Arms	TRIAL DESIGN	One record per planned Element per Arm	Tabulation	STUDYID, ARMCD, TAETORD	ta.xpt	
TE	Trial Elements	TRIAL DESIGN	One record per planned Element	Tabulation	STUDYID, ETCD	te.xpt	
TI	Trial Inclusion/Exclusion Criteria	TRIAL DESIGN	One record per I/E criterion	Tabulation	STUDYID, IETESTCD	ti.xpt	
TS	Trial Summary	TRIAL DESIGN	One record per trial summary parameter value	Tabulation	STUDYID, TSPARMCD, TSSEQ	ts.xpt	
TV	Trial Visits	TRIAL DESIGN	One record per planned Visit per Arm	Tabulation	STUDYID, VISITNUM, ARMCD	tv.xpt	
CO	Comments	SPECIAL PURPOSE	One record per comment per subject	Tabulation	STUDYID, USUBJID, RDOMAIN, IDVAR, IDVARVAL	co.xpt	
DM	Demographics	SPECIAL PURPOSE	One record per subject	Tabulation	STUDYID, USUBJID	dm.xpt	See Reviewer's Guide, Section 2.1 Demographics Reviewers Guide
SE	Subject Elements	SPECIAL PURPOSE	One record per actual Element per subject	Tabulation	STUDYID, USUBJID, TAETORD	se.xpt	
SV	Subject Visits	SPECIAL PURPOSE	One record per actual visit per subject	Tabulation	STUDYID, USUBJID, VISITNUM	sv.xpt	
CM	Concomitant Medications	INTERVENTIONS	One record per recorded medication occurrence or constant-dosing interval per subject	Tabulation	STUDYID, USUBJID, CMTRT, CMDOSE, CMCAT, CMSTDTC, CMDOSFRQ	cm.xpt	
EX	Exposure	INTERVENTIONS	One record per constant dosing interval per subject	Tabulation	STUDYID, USUBJID, EXTRT, EXSTDTC	ex.xpt	
AE	Adverse Events	EVENTS	One record per adverse event per subject	Tabulation	STUDYID, USUBJID, AEDECOD, AESTDTC	ae.xpt	
DS	Disposition	EVENTS	One record per disposition status or protocol milestone per subject	Tabulation	STUDYID, USUBJID, DSDECOD, DSSTDTC	ds.xpt	
MH	Medical History	EVENTS	One record per medical history event per subject	Tabulation	STUDYID, USUBJID, MHTERM, MHSTDTC	mh.xpt	
EG	ECG Test Results	FINDINGS	One record per ECG observation per visit per subject	Tabulation	STUDYID, USUBJID, EGDTC, EGCAT, EGTESTCD	eg.xpt	
LB01	Laboratory Tests Results Hematology (Laboratory Tests Results)	FINDINGS	One record per analyte per visit per subject	Tabulation	STUDYID, USUBJID, LBDTC, LBCAT, LBSCAT, LBTESTCD	lb01.xpt	LB is submitted as a split dataset. The split was done based on LBCAT. When LBCAT=HEMATOLOGY then data is saved in LB01. All other data is saved in LB02
LB02	Laboratory Tests Results non Hematology (Laboratory Tests Results)	FINDINGS	One record per analyte per visit per subject	Tabulation	STUDYID, USUBJID, LBDTC, LBCAT, LBSCAT, LBTESTCD	lb02.xpt	LB is submitted as a split dataset. The split was done based on LBCAT. When LBCAT=HEMATOLOGY then data is saved in LB01. All other data is saved in LB02

VARIABLE LEVEL METADATA

Define-XML version represents variable metadata using an ItemDef element in both the versions and it's associated with,

- Controlled Terminology
- Value Level Metadata
- Computational Method
- Comments
- Origin

The way of representing the XML elements differs on Define-XML version 1 and 2.

The VLM UI shows the details of variable level metadata for the new version and it permits the user to export, make changes to comments, value list & any other additional details, followed by re-import in order to reflect all the changes at the variable level. VLM UI with details is shown in the User Interface 4.

Below screen snapshots depicts variable level information of two versions of define.xml file for better clarity.

Define-XML Version 1 [Variable Level Metadata]

Demographics Dataset (DM)						**dm.xpt**
Variable	**Label**	**Type**	**Controlled Terms or Format**	**Origin**	**Role**	**Comment**
STUDYID	Study Identifier	text		Protocol	IDENTIFIER	
DOMAIN	Domain Abbreviation	text		Assigned	IDENTIFIER	Two-character abbreviation for the domain most relevant to the observation. The Domain abbreviation is also used as a prefix for variables to ensure uniqueness when datasets are merged.
USUBJID	Unique Subject Identifier	text		Derived	IDENTIFIER	See Computational Method: COMPMETHOD.USUBJID
SUBJID	Subject Identifier for the Study	text		CRF Page 1	TOPIC	
RFSTDTC	Subject Reference Start Date/Time	datetime	ISO8601	Derived	RECORD QUALIFIER	See Computational Method: COMPMETHOD.RFSTDTC
RFENDTC	Subject Reference End Date/Time	datetime	ISO8601	Derived	RECORD QUALIFIER	See Computational Method: COMPMETHOD.RFENDTC
RFXSTDTC	Date/Time of First Study Treatment	datetime	ISO8601	Derived	RECORD QUALIFIER	See Computational Method: COMPMETHOD.RFXSTDTC
RFXENDTC	Date/Time of Last Study Treatment	datetime	ISO8601	Derived	RECORD QUALIFIER	See Computational Method: COMPMETHOD.RFXENDTC
RFICDTC	Date/Time of Informed Consent	date	ISO8601	CRF Page 1	RECORD QUALIFIER	
RFPENDTC	Date/Time of End of Participation	date	ISO8601	Derived	RECORD QUALIFIER	See Computational Method: COMPMETHOD.RFPENDTC
DTHDTC	Date/Time of Death	date	ISO8601	CRF Page 907	RECORD QUALIFIER	From AE page; when outcome is Death then Y then date/time, else Null

Define-XML Version 2 [Variable Level Metadata]

Demographics (DM) [Location: dm.xpt]

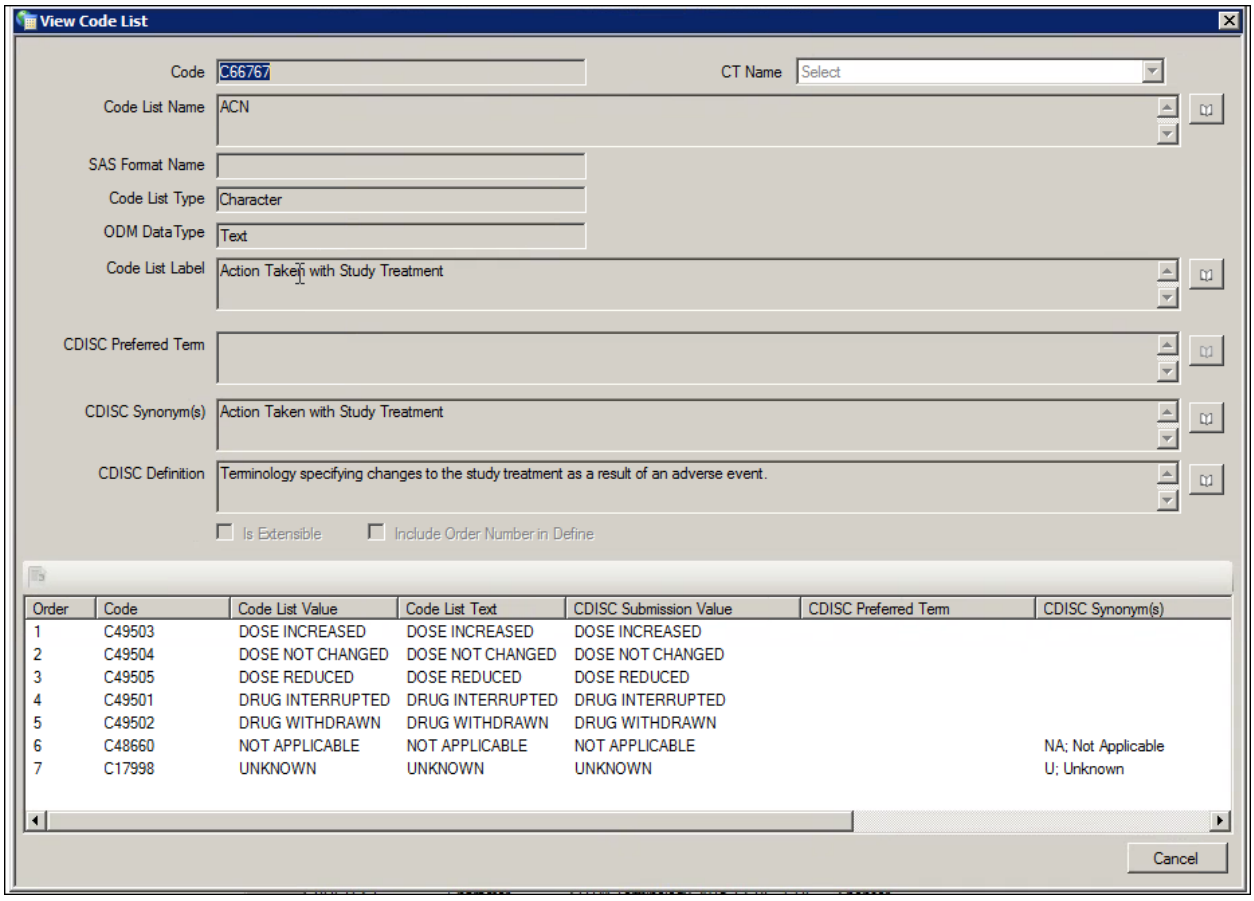
Variable	Label	Key	Type	Length	Controlled Terms or Format	Origin	Derivation/Comment
STUDYID	Study Identifier	1	text	40		Protocol	
DOMAIN	Domain Abbreviation		text	2		Assigned	Two-character abbreviation for the domain most relevant to the observation. The Domain abbreviation is also used as a prefix for variables to ensure uniqueness when datasets are merged.
USUBJID	Unique Subject Identifier	2	text	20		Derived	Concatenation of STUDYID and SUBJID
SUBJID	Subject Identifier for the Study		text	3		CRF Page 1	
RFSTDTC	Subject Reference Start Date/Time		datetime		ISO8601	Derived	RFSTDTC = First date/time of study drug treatment
RFENDTC	Subject Reference End Date/Time		datetime		ISO8601	Derived	RFENDTC = Date of Completion or date of Withdrawal. Null for screen failures
RFXSTDTC	Date/Time of First Study Treatment		datetime		ISO8601	Derived	RFXSTDTC = Date/time of first exposure to protocol specified treatment from CRF
RFXENDTC	Date/Time of Last Study Treatment		datetime		ISO8601	Derived	RFXENDTC = Date/time of last exposure to protocol specified treatment from CRF
RFICDTC	Date/Time of Informed Consent		date		ISO8601	CRF Page 1	
RFPENDTC	Date/Time of End of Participation		date		ISO8601	Derived	RFPSTDTC = Date/time when subject ended participatient or follow-up in a trial (last date/time of contact)
DTHDTC	Date/Time of Death		date		ISO8601	CRF Page 907	From AE page; when outcome is Death then Y then date/time, else Null
DTHFL	Subject Death Flag		text	1	["Y" = "Yes"] <No Yes Response Subset (Y)>	CRF Page 907	From AE page; when outcome is Death then Y, else Null

CODE LISTS

The term "Controlled Terminology" in the context of a study refers to the set of all allowable values across all variables that have finite sets of allowable values in the study. A "Codelist" is a unique subset of the controlled terminology to which one or more variables are subject.

The CodeList element can define either an internal or external code list. Internal code lists include a list of allowable codes and, if applicable, their corresponding decodes. In most cases where controlled terminology is just an enumeration of allowed values, the EnumeratedItem element can be used to define the list of values. In cases where it is useful to provide decodes for the coded values, the CodeListItem element can be used.

EXACT can import the SDTM Terminology spreadsheet published in CDISC eSHARE downloads (SDTM Terminology Section) and based on the Define-XML version, its populates the C-code information, permitted values etc.



User Interface 6: Code Lists

In Define-XML version 1 there was only one way to define internal codelists in the define.xml file. You always had to specify both coded values and decode values. Define-XML version 2 adds extended support for Controlled Terminology. The CodeList element can define either an internal or external CodeList. Internal CodeLists include a list of allowable codes and, if applicable, their corresponding decodes. In most cases where controlled terminology is just an enumeration of allowed values, the EnumeratedItem element can be used to define the list of values. In cases where it is useful to provide decodes for the coded values, the CodeListItem element can be used. Both of these CodeLists items now can have an attribute that defines whether the item is an extension to an extensible CodeList.

For CodeList elements that define the contents of a CDISC Controlled Terminology an identification of the C-codes has been implemented that is used to identify Controlled Terminology Codelists and Coded Terms within the National Cancer Institute's Enterprise Vocabulary System.

Below screen snapshots depicts code list information of two versions of define.xml file for better clarity.

Define-XML Version 1 [Codelists]

Controlled Terminology (Code Lists)	
ACN, Reference Name (ACN)	
Code Value	**Code Text**
DOSE NOT CHANGED	DOSE NOT CHANGED
DRUG WITHDRAWN	DRUG WITHDRAWN
NOT APPLICABLE	NOT APPLICABLE
AEENRF, Reference Name (AEENRF)	
Code Value	**Code Text**
AFTER	AFTER
DURING/AFTER	DURING/AFTER
AEREL, Reference Name (AEREL)	
Code Value	**Code Text**
DEFINITE	DEFINITE
POSSIBLY	POSSIBLY
PROBABLE	PROBABLE
UNLIKELY	UNLIKELY
UNRELATED	UNRELATED
AESEV, Reference Name (AESEV)	
Code Value	**Code Text**
MILD	MILD
MODERATE	MODERATE
SEVERE	SEVERE
ARM, Reference Name (ARM)	
Code Value	**Code Text**
2 mg Filgrastim	2 mg Filgrastim
ARMCD, Reference Name (ARMCD)	
Code Value	**Code Text**
A	2 mg Filgrastim

Define-XML Version 2 [Codelists]

Relation to Reference Period (AEENRF) [CL.AEENRF, *C66728*]

Permitted Value (Code)
AFTER [*C38008*]
DURING/AFTER [*C49640*]

Causality [CL.AEREL]

Permitted Value (Code)
UNRELATED
UNLIKELY
POSSIBLY
PROBABLE
DEFINITE

Severity/Intensity Scale for Adverse Events [CL.AESEV, *C66769*]

Permitted Value (Code)
MILD [*C41338*]
MODERATE [*C41339*]
SEVERE [*C41340*]

Description of Planned Arm [CL.ARM]

Permitted Value (Code)
2 mg Filgrastim

Planned Arm Code [CL.ARMCD]

Permitted Value (Code)	Display Value (Decode)
A	2 mg Filgrastim

Units for Vital Signs Results (CM only) [CL.CM, *C66770*]

Permitted Value (Code)	Display Value (Decode)
cm [*C49668*]	Centimeter

VALUE LEVEL METADATA

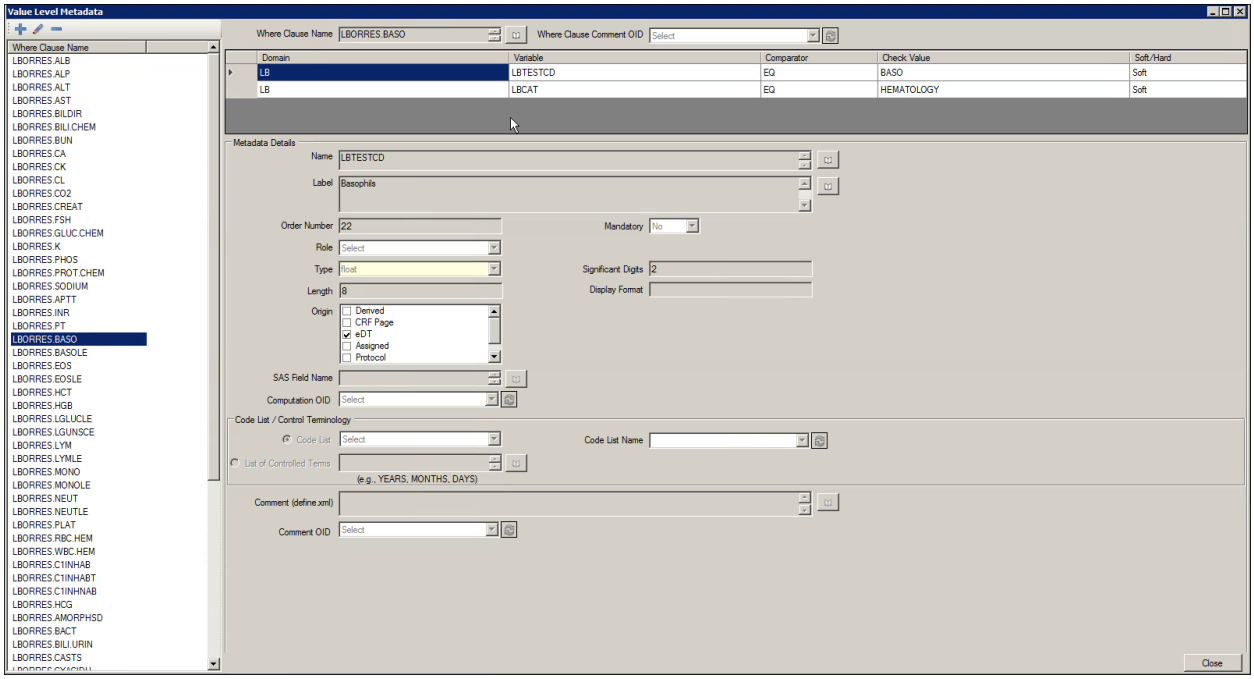
The normalized data structure used by datasets based on the SDTM, SEND and ADaM models (generally one record per subject per test code per visit or observation) provides an efficient method for transmitting information. However, there are cases where the dataset variable metadata does not provide sufficient detail to support data review and analysis.

In these cases Value Level Metadata should be provided in the Define-XML document. Value Level Metadata enables the specification of the metadata of a variable under conditions involving one or more other dataset variables. The definition of a variable for a specific condition is known as Value Level Metadata.

Value Level Metadata should be provided when there is a need to describe differing metadata attributes for subsets of cells within a column. It is most often used on SDTM Findings domains to provide definitions for Variables such as --ORRES, --ORRESU, --STRES, --STRESU that are specific to each test code (value of --TESTCD). It is not required for Findings domains where the results have the same characteristics in all records, such as IE domains.

Although support for Value Level Metadata has always been part of the Define-XML standard, there were shortcomings. In Define-XML version 1 Value Level Metadata was typically attached to –TESTCD variables or QNAM variables. It was defined by convention which data set variable was being described by the Value Level Metadata (e.g. –ORRES or QVAL).

In EXACT, the Value Level Metadata can be configured either through user interface or an excel spreadsheet using the import functionality.

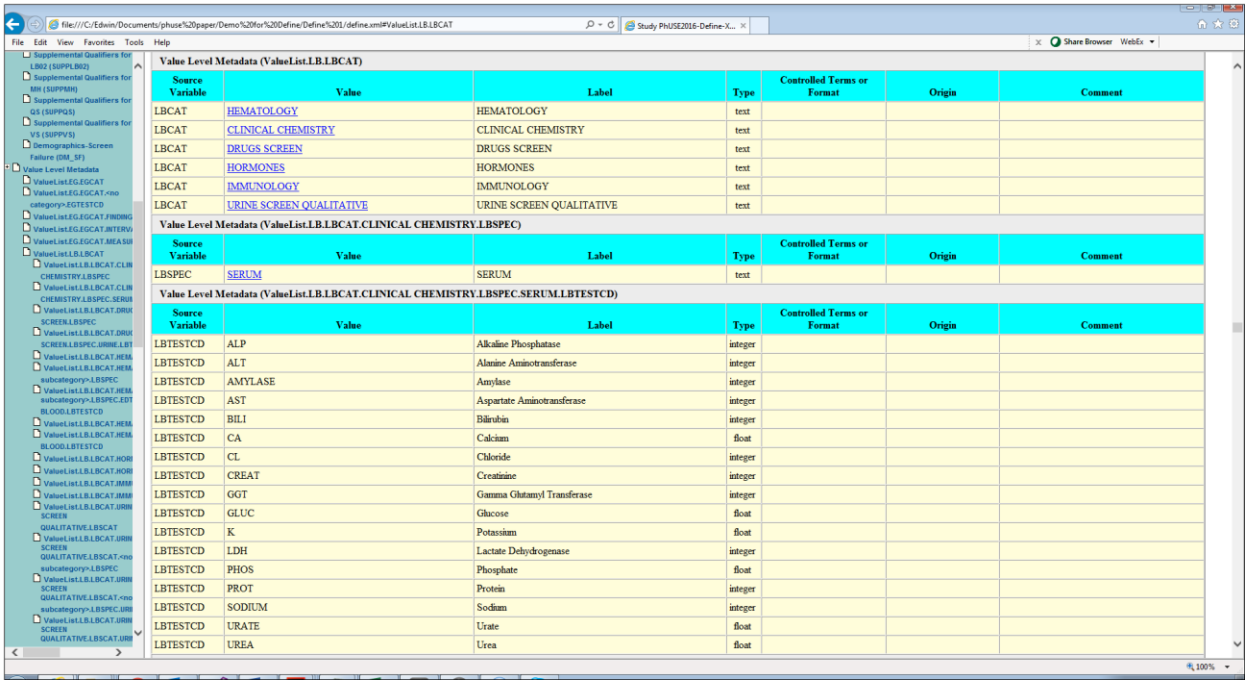


User Interface 7: Value Level Metadata

In Define-XML version 2 where clauses are being used to explicitly describe – in a machine-readable form – under which condition the Value Level definitions apply. Value Level Metadata can be provided whenever there is a need to describe differing metadata attributes for values or subsets of values of a variable. In Define-XML version 2 valuelists should be attached to the variable being defined so that any number of variables can be defined in detail.

Below screen snapshots depicts value level information of two versions of define.xml file for better clarity.

Define-XML Version 1 [Value Level Metadata]



Value Level Metadata (ValueList.LB.LBCAT)

Source Variable	Value	Label	Type	Controlled Terms or Format	Origin	Comment
LBCAT	HEMATOLOGY	HEMATOLOGY	text			
LBCAT	CLINICAL CHEMISTRY	CLINICAL CHEMISTRY	text			
LBCAT	DRUGS SCREEN	DRUGS SCREEN	text			
LBCAT	HORMONES	HORMONES	text			
LBCAT	IMMUNOLOGY	IMMUNOLOGY	text			
LBCAT	URINE SCREEN QUALITATIVE	URINE SCREEN QUALITATIVE	text			

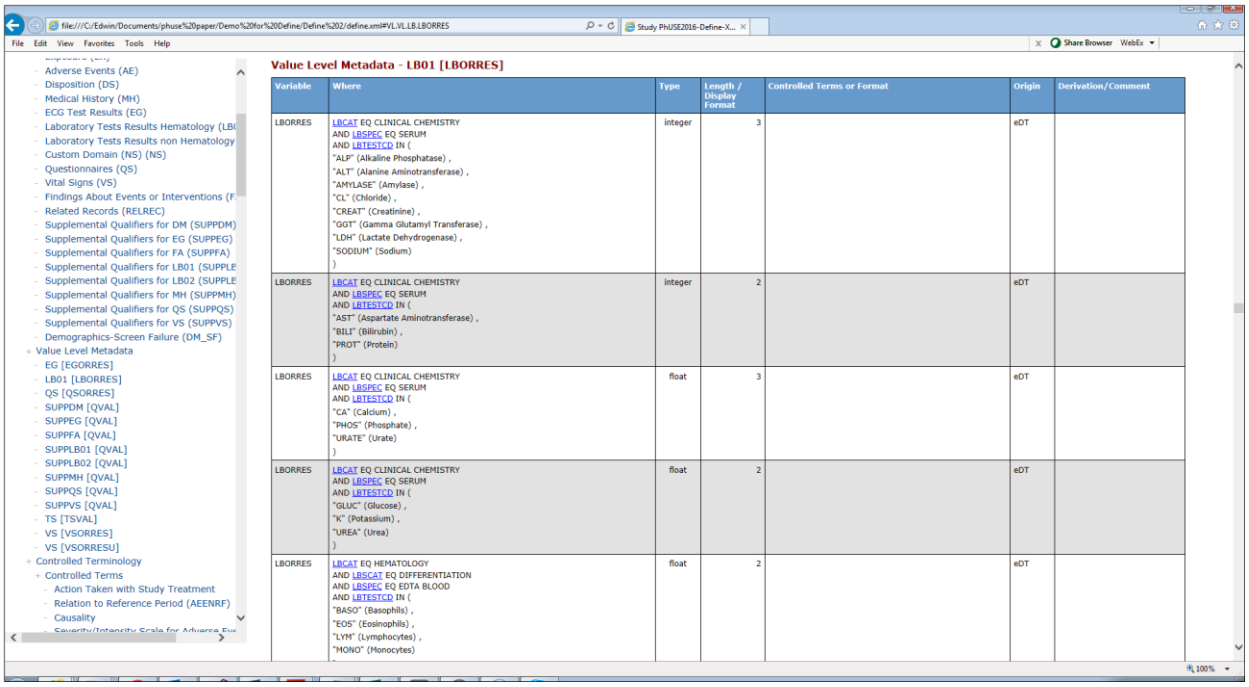
Value Level Metadata (ValueList.LB.LBCAT.CLINICAL CHEMISTRY.LBSPEC)

Source Variable	Value	Label	Type	Controlled Terms or Format	Origin	Comment
LBSPEC	SERUM	SERUM	text			

Value Level Metadata (ValueList.LB.LBCAT.CLINICAL CHEMISTRY.LBSPEC.SERUM.LBTESTCD)

Source Variable	Value	Label	Type	Controlled Terms or Format	Origin	Comment
LBTESTCD	ALP	Alkaline Phosphatase	integer			
LBTESTCD	ALT	Alanine Aminotransferase	integer			
LBTESTCD	AMYLASE	Amylase	integer			
LBTESTCD	AST	Aspartate Aminotransferase	integer			
LBTESTCD	BILI	Bilirubin	integer			
LBTESTCD	CA	Calcium	float			
LBTESTCD	CL	Chloride	integer			
LBTESTCD	CREAT	Creatinine	integer			
LBTESTCD	GGT	Gamma Glutamyl Transferase	integer			
LBTESTCD	GLUC	Glucose	float			
LBTESTCD	K	Potassium	float			
LBTESTCD	LDH	Lactate Dehydrogenase	integer			
LBTESTCD	PHOS	Phosphate	float			
LBTESTCD	PROT	Protein	integer			
LBTESTCD	SODIUM	Sodium	integer			
LBTESTCD	URATE	Urate	float			
LBTESTCD	UREA	Urea	float			

Define-XML Version 2 [Value Level Metadata]



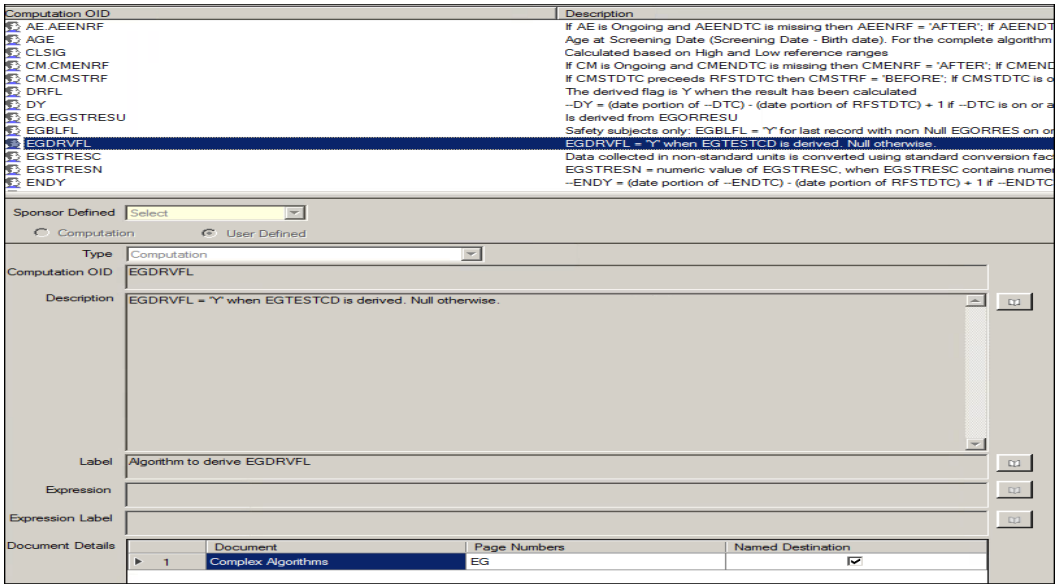
Value Level Metadata - LB01 [LBORRES]

Variable	Where	Type	Length / Display Format	Controlled Terms or Format	Origin	Derivation/Comment
LBORRES	LBCAT EQ CLINICAL CHEMISTRY AND LBSPEC EQ SERUM AND LBTESTCD IN ("ALP" (Alkaline Phosphatase) , "ALT" (Alanine Aminotransferase) , "AMYLASE" (Amylase) , "CL" (Chloride) , "CREAT" (Creatinine) , "GGT" (Gamma Glutamyl Transferase) , "LDH" (Lactate Dehydrogenase) , "SODIUM" (Sodium))	integer	3		eDT	
LBORRES	LBCAT EQ CLINICAL CHEMISTRY AND LBSPEC EQ SERUM AND LBTESTCD IN ("AST" (Aspartate Aminotransferase) , "BILI" (Bilirubin) , "PROT" (Protein))	integer	2		eDT	
LBORRES	LBCAT EQ CLINICAL CHEMISTRY AND LBSPEC EQ SERUM AND LBTESTCD IN ("CA" (Calcium) , "PHOS" (Phosphate) , "URATE" (Urate))	float	3		eDT	
LBORRES	LBCAT EQ CLINICAL CHEMISTRY AND LBSPEC EQ SERUM AND LBTESTCD IN ("GLUC" (Glucose) , "K" (Potassium) , "UREA" (Urea))	float	2		eDT	
LBORRES	LBCAT EQ HEMATOLOGY AND LBSCAT EQ DIFFERENTIATION AND LBSPEC EQ EDTA BLOOD AND LBTESTCD IN ("BASO" (Basophils) , "EOS" (Eosinophils) , "LYM" (Lymphocytes) , "MONO" (Monocytes)	float	2		eDT	

COMPUTATIONAL ALGORITHMS

The MethodDef element describes the algorithms used to generate values for variables defined as derived. For cases where the algorithm description is longer than a few lines the method can link to a section in a computational algorithms document containing the additional details. A formal expression can be provided that contains a machine-readable expression that implements the algorithm.

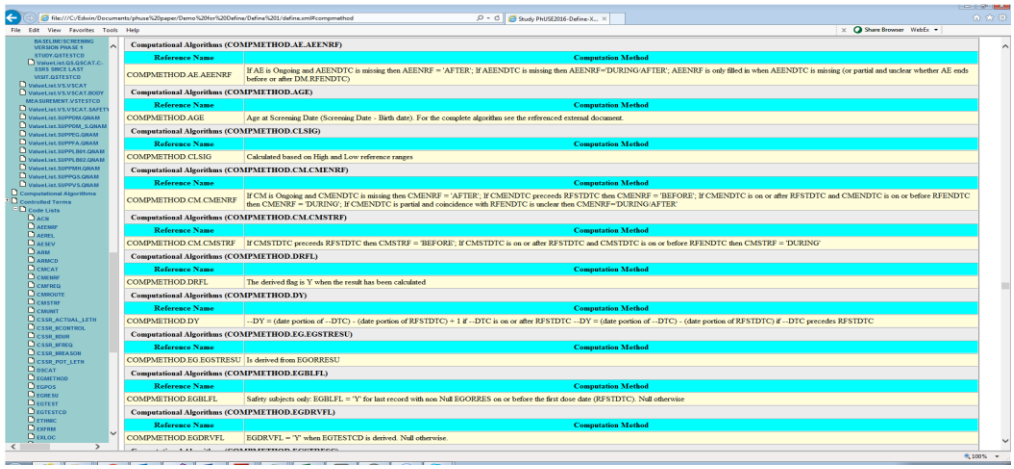
In EXACT, the computational algorithm can be configured/re-used from sponsor library either through user interface or an excel spreadsheet using the import functionality.



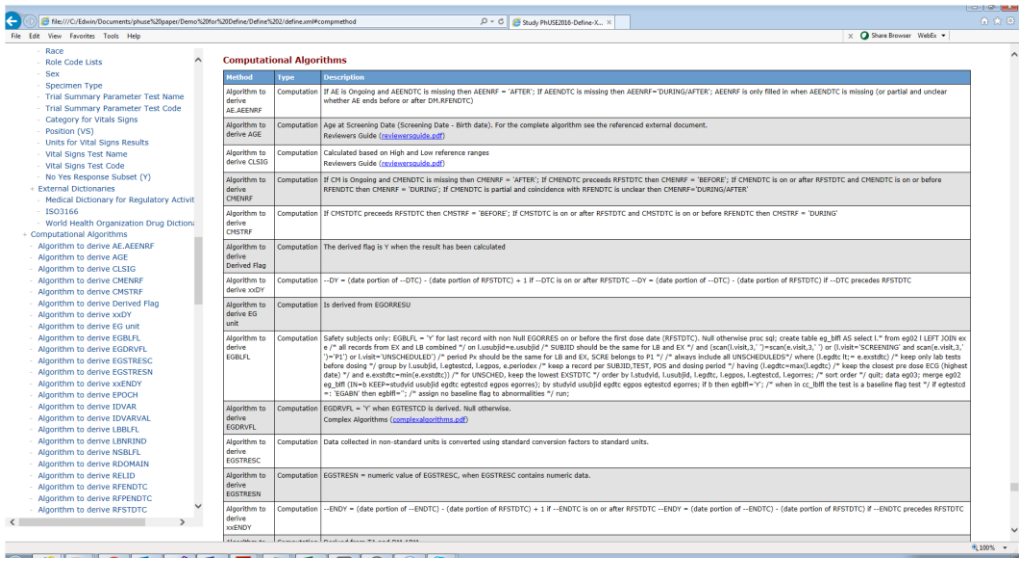
User Interface 8: Computation Methods

Below screen snapshots depicts computation method details of two versions of define.xml file for better clarity.

Define-XML Version 1 [Computational Algorithms]



Define-XML Version 2 [Computational Algorithms]



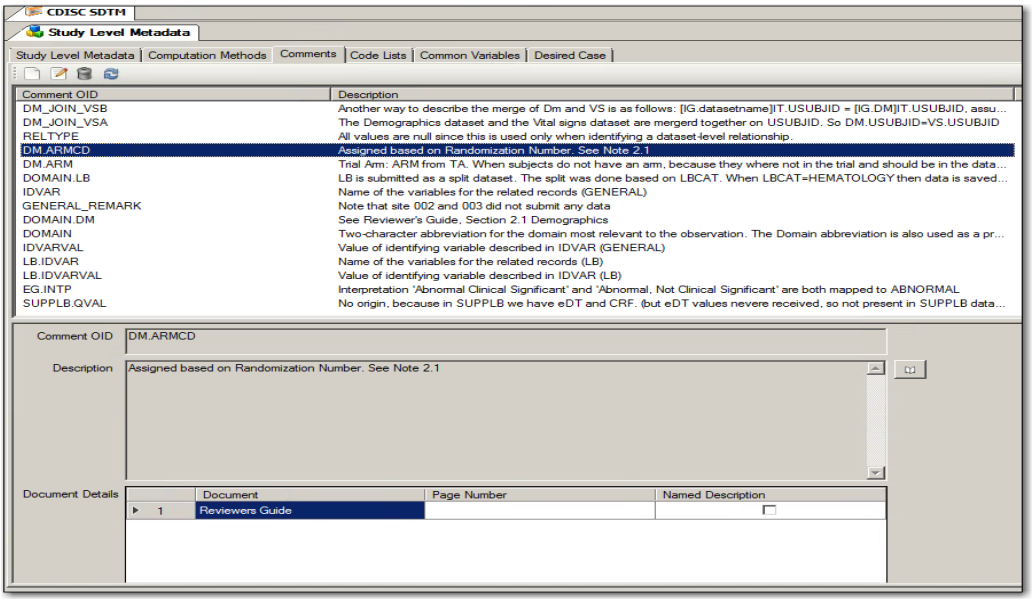
COMMENTS

Define-XML v2.0 allows the definition of comments at dataset, variable and value levels.

The mechanism allows referencing short comments self-contained in the Define-XML document or long comments referenced in external documents. For comments in external documents, the reference can include specific pages within the document.

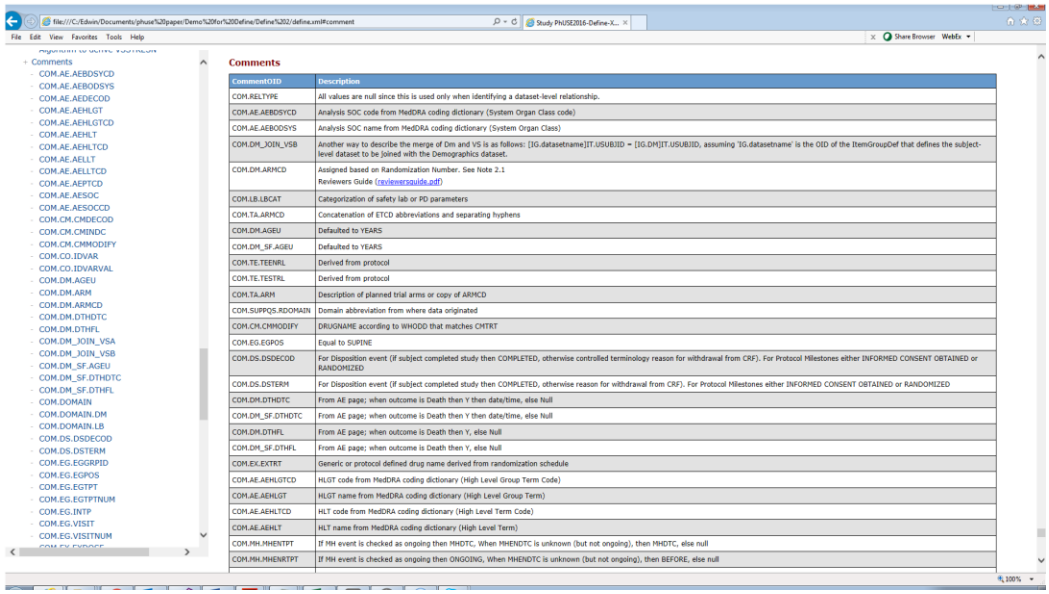
Comments are not intended to replace a properly defined computational algorithm, which is expected for derived variables. Define-XML version 1 had very limited possibilities for providing comments related to the various metadata objects in a clinical study.

Below screen snapshots depicts comments information of Define-XML v2.0 for better clarity whereas the comments are part of the VLM UI.



User Interface 9: Comments

Define-XML Version 2 [Comments]



Pinnacle21 is used within the system to validate the output dataset and define.xml based on the version selected by the user.

CONCLUSION

The paper was focused towards handling the different version of SDTM and Define-XML for a single study and went onto explain the details on how this can be done using a system in a more efficient, time save and consistent quality while dealing with such situations for many studies by a Sponsor or CRO. There were many tools available and even the version comparison were provided by CDISC, which would involve manual effort by the users. The demonstration outlined with steps in this paper would facilitate ease of use in making such change for a study by a user and bring in a process framework to be followed by a CRO or a Sponsor who has global operations dealing with such situations across studies.

REFERENCES

1. Study Data Tabulation Model (SDTM) (http://www.cdisc.org/standards/foundational/sdtm)
2. Define-XML (http://www.cdisc.org/standards/foundational/define-xml)
3. CDISC eSHARE Downloads (http://www.cdisc.org/members-only/eshare-downloads)
4. FDA Study Data Standards Catalog (Version 4.5; Effective 2016-07-18) (http://www.fda.gov/downloads/ForIndustry/DataStandards/StudyDataStandards/UCM340684.xlsx)

ACKNOWLEDGMENTS

The author(s) would like to thank the members of the EXACT Team from PRA Health Sciences.

RECOMMENDED READING

CDISC: http://www.cdisc.org
FDA's Study Data Standards Resources: http://www.fda.gov/ForIndustry/DataStandards/StudyDataStandards/

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Edwin Ponraj Thangarajan, Giri Balasubramanian
PRA Health Sciences
40, II Main Road, R.A. Puram
Chennai - 600 028, TamilNadu, India
Email: ThangarajanEdwin@prahs.com, BalasubramanianGiri@prahs.com
Web: www.prahs.com