CTSU

Cancer Trials Support Unit

CDISC-Compliant Rave GLIB ALS v1.0 Release Notes

Document Version 1.0

July 31, 2019

Document Information

Revision Information for the CDISC-Compliant Rave GLIB ALS v1.0 Release Notes

Revision History			
#	**Date**	**By**	**Description**
1.0	07/31/2019	Quan Duong	Initial Release.
Last Saved By on 7/31/2019 3:10:00 PM			

This document was prepared by:

WESTAT, Cancer Trials Support Unit
1600 Research Boulevard
Rockville, Maryland 20850

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 Overview

The use of Clinical Data Interchange Standards Consortium (CDISC) standards is required for data submissions to the US Food and Drug Administration (FDA). CDISC submissions to the FDA require the use of controlled terminology and datasets formatted according to Study Data Tabulation Model (SDTM). Collecting clinical data using Clinical Data Acquisition Standards Harmonization (CDASH) facilitates this submission process.

The following provides an overview of the National Cancer Institute's (NCI's) CDISC Implementation, as per the *NCI CDISC Implementation Wiki*:

> "The FDA has mandated sponsors whose studies start after Dec 17, 2016, must submit their clinical study data sets in the Study Data Tabulation Model (SDTM) standard format. For INDs, the requirement applies for studies that start after Dec. 17, 2017. SDTM provides a standard for organizing and formatting data to streamline the process in collection, management, analysis and reporting. The Clinical Data Interchange Standards Consortium (CDISC) is a global nonprofit standards development organization with a worldwide team of staff and volunteer experts across the medical community. CDISC provides data standards to streamline clinical research, one of which being SDTM. CDISC is also developing Clinical Data Acquisition Standards Harmonization which establishes a standard way to collect data in a similar way across studies and sponsors so that data formats and structures provide clear traceability of submission into the SDTM. To support the FDA mandate of submitting clinical study stat sets to the FDA in the SDTM format, the NCI is transitioning their current Network Rave Data Standards (NRDS) Initiative, led by the Cancer Therapy Evaluation Program (CTEP) into the CDISC implementation." (***https://wiki.nci.nih.gov/display/CDISC/NCI+CDISC+Implementation+Home***)
>
> "The NCI/CTEP is transitioning the existing NRDS initiative to the CDISC Implementation initiative to meet the FDA mandate of submitting clinical study data sets in the SDTM format. The NCI is working in collaboration with CDISC to collect data in the CDASH format for the Oncology Patient Enrollment Network System (OPEN), Clinical Therapy Evaluation Program Adverse Event Reporting System (CTEP-AERS) and the Clinical Data Update System (CDUS). According to the FDA Study Data Technical Conformance Guide, section 4.1.2 (SDTM General Considerations), it is recommended that sponsors implement the collection of data in a format that is harmonized with SDTM such as CDASH." (***https://wiki.nci.nih.gov/display/CDISC/CDISC+Implementation+Project+Overview***)

Per the NCI, all Investigational New Drug (IND) studies activated on or after 01/01/2020 must be CDISC-compliant. The Cancer Trials Support Unit (CTSU) is supporting the NCI's CDISC implementation by developing a CDISC-Compliant Rave Global Library (GLIB) Architect Loader Specification (ALS). The CDISC-compliant Rave GLIB ALS will specify CDASH and SDTM variables; the Lead Protocol Organizations (LPOs) will use this ALS for building their own Rave CDISC-compliant GLIBs.

1.2 Acronyms and Definitions

Table 1: Acronyms and Definitions

Acronym	Definition
AE	Adverse Events
ALS	Architect Loader Specification
AP	Associated Persons
ARM	In SDTM, an ARM is a complete planned path that a study subject will have from the beginning to the end of a study.
caDSR	Cancer Data Standards Registry and Repository
CBER	Center for Biologics Evaluation and Research
CDASH	Clinical Data Acquisition Standards Harmonization. Basic standards for the collection of clinical trial data and how to implement the standard for specific case report forms (CRFs). Optimized for data capture, investigator site activities and data cleaning. The CDASH standard includes the CDASHIG (including the metadata) and the CDASH Model.
CDASH Metadata Table	Includes variables commonly implemented by a significant number of the organizations/companies (e.g., Medical History, Adverse Events).
CDASH Model	Provides a general framework and root metadata for creating fields to collect information on forms. Root metadata includes root variables and root questions. The root CDASH Model variables are intended to facilitate mapping to the SDTMIG variables while addressing specific data collection needs.
CDASHIG	CDASH Implementation Guide provides information on the implementation of CDASH standards for specific topics of data. Each topic is represented by a CDASH domain. CDASH domains, variables and controlled terminology are aligned with SDTM. Each CDASHIG domain contains a description of the data topic, a specification table, including standard metadata for data collection, general assumptions/rules, and example forms.
CDER	Center for Drug Evaluation and Research
CDUS	Clinical Data Update System
Controlled Terminology	Set of CDISC-developed or CDISC-adopted standard expressions (values) used with data items within CDISC-defined datasets (e.g., a CDISC specific list of values). Controlled terminology should be used throughout the clinical research lifecycle, from data collection through analysis and submission.
CTEP	Cancer Therapy Evaluation Program
CTEP-AERS	Clinical Therapy Evaluation Program Adverse Event Reporting System
CTSU	Cancer Trials Support Unit
EDC	Rave Electronic Data Capture
ELEMENT	In SDTM an ELEMENT is a discrete state of being, or building block for ARMs. ELEMENTS are defined with start and end rules, and then placed in a specific order to create a study ARM.
FDA	US Food and Drug Administration

Acronym	Definition
GLIB	Global Library
GLV	Global Library Volume
IND	Investigational New Drug
ISO	International Organization for Standardization
LPO	Lead Protocol Organization
NCI	National Cancer Institute
NIH	National Institutes of Health
NRDS	Network Rave Data Standards
OCI	Object Cart Importer
OPEN	Oncology Patient Enrollment Network System
QRS	Questionnaires, Ratings, and Scales
SDTM	Study Data Tabulation Model provides a standard for organizing and formatting data to streamline processes in collection, management, analysis and reporting. SDTM describes the general conceptual model for organizations/companies involved in the collection, preparation and analysis of clinical study data that is submitted to regulatory authorities. The model also contains general information that can be used across domains.
SDTM Metadata Table	Includes the standard metadata found in the applicable SDTM model and/or Implementation Guide.
SDTMIG	SDTM Implementation Guide contains detailed information for each domain, associated variables, assumptions, business rules, and examples for preparing standard tabulation datasets that are based on the SDTM.
TA	Therapeutic Area
TAC	Treatment Assignment Code
TAD	Treatment Assignment Description
WG	Working Group

1.3 Scope

This document outlines the details of the CDISC-Compliant Rave GLIB ALS created by the CTSU. It also contains information to assist LPOs in configuring CDISC-compliant, LPO-specific forms directly in Rave, using the CDISC-compliant GLIB ALS. This document does not include instruction for the review, approval, and curation of variables not included in this ALS.

1.4 Audience

This document was developed by CTSU for release to the LPO staff responsible for study build activities in Medidata Rave.

2. CDISC Versions

The CDISC-compliant Rave GLIB ALS Version 1.0 includes forms that were created based on versions of the CDISC components listed in Table 2: *CDISC Version and Links*. As subsequent CDISC versions are released, an assessment will be made to determine the need for updates to the GLIB ALS.

Instructions: To access the links, first log in to the CDISC website (https://www.cdisc.org) using your National Institutes of Health (NIH) email address. These links only work for NIH staff members or LPOs that have obtained their own NIH account access.

Table 2: CDISC Version and Links

CDISC Version	Links
CDASH Model v1.0	https://www.cdisc.org/standards/foundational/cdash/cdash-model-10
CDASH Implementation Guide (CDASHIG) v2.0	https://www.cdisc.org/standards/foundational/cdash/cdash-20
CDASH and SDTM Controlled Terminology packages P37 released through March 29, 2019 *Note: Controlled Terminology packages are released quarterly The older versions can be accessed via the CDISC Library Archive.*	https://evs.nci.nih.gov/ftp1/CDISC/SDTM
CDASHIG v2.0 Metadata Table	https://www.cdisc.org/cdisc-library https://www.cdisc.org/members-only/cdisc-library-archives
SDTM Model v1.7	https://www.cdisc.org/standards/foundational/sdtm/sdtm-v1-7
SDTM Implementation Guide (SDTMIG) v3.3	https://www.cdisc.org/standards/foundational/sdtmig/sdtmig-v3-3
SDTMIG v3.3 Metadata Table	https://www.cdisc.org/cdisc-library https://www.cdisc.org/members-only/cdisc-library-archives
Oncology CDISC Therapeutic Area (TA) User Guides	https://www.cdisc.org/standards/therapeutic-areas/disease-area/oncology

3. CDISC Standard Forms

The CDISC-compliant Rave GLIB ALS Version 1.0 is the initial version released for use by the LPOs; it contains the standard CDASH and SDTM form details to assist the LPOs in building their CDISC-compliant Rave GLIB and their subsequent CDISC-compliant studies. The following CDISC-compliant GLIB components are included in the ALS:

- CDASH and SDTM Domains or Forms
- CDASH Variables
- SDTM Variables (no direct mapping to CDASH)
- Controlled Terminologies Assigned to Variables

3.1 CDASH and SDTM Domains

3.1.1 List of Domains Included in this Release

Table 3: *List of Domains Included in this Release* provides a listing of domains that are included in the GLIB ALS. There are 25 CDASH and SDTM domains (e.g., AE, CE, CM) containing both CDASH and SDTM variables, and 18 SDTM domains (AG, CV, FT) containing only SDTM variables. Controlled terminologies assigned to variables are included in the DataDictionaries and DataDictionaryEntries tabs in the ALS.

Table 3: List of Domains Included in this Release

Seq. #	Domain/Form	Form Name	CDASH and SDTM or SDTM?
1.	AE	Adverse Events	CDASH and SDTM
2.	AG	Procedure Agents	SDTM
3.	CE	Clinical Events	CDASH and SDTM
4.	CM	Concomitant/Prior Medications	CDASH and SDTM
5.	CO	Comments	CDASH and SDTM
6.	CV	Cardiovascular System Findings	SDTM
7.	DA	Drug Accountability	CDASH and SDTM
8.	DD	Death Details	CDASH and SDTM
9.	DM	Demographics	CDASH and SDTM
10.	DS	Disposition	CDASH and SDTM
11.	DV	Protocol Deviations	CDASH and SDTM
12.	EC	Exposure as Collected	CDASH and SDTM
13.	EG	ECG Test Results	CDASH and SDTM
14.	FA	Findings About Events or Interventions	CDASH and SDTM
15.	FT	Functional Tests	SDTM
16.	HO	Healthcare Encounters	CDASH and SDTM
17.	IE	Inclusion/Exclusion Criteria Not Met	CDASH and SDTM

Seq. #	Domain/Form	Form Name	CDASH and SDTM or SDTM?
18.	IS	Immunogenicity Specimen Assessments	SDTM
19.	LB	Laboratory Test Results	CDASH and SDTM
20.	MB	Microbiology Specimen	SDTM
21.	MH	Medical History	CDASH and SDTM
22.	MI	Microscopic Findings	CDASH and SDTM
23.	MK	Musculoskeletal System Findings	SDTM
24.	ML	Meal Data	SDTM
25.	MS	Microbiology Susceptibility	SDTM
26.	NV	Nervous System Findings	SDTM
27.	OE	Ophthalmic Examinations	SDTM
28.	PC	Pharmacokinetics Concentrations	CDASH and SDTM
29.	PE	Physical Examination	CDASH and SDTM
30.	PP	Pharmacokinetics Parameters	SDTM
31.	PR	Procedures	CDASH and SDTM
32.	QS	Questionnaires	SDTM
33.	RE	Respiratory System Findings	SDTM
34.	RP	Reproductive System Findings	CDASH and SDTM
35.	RS	Disease Response and Clin Classification	SDTM
36.	SC	Subject Characteristics	CDASH and SDTM
37.	SR	Skin Response	CDASH and SDTM
38.	SS	Subject Status	SDTM
39.	SU	Substance Use	CDASH and SDTM
40.	TR	Tumor/Lesion Results	SDTM
41.	TU	Tumor/Lesion Identification	SDTM
42.	UR	Urinary System Findings	SDTM
43.	VS	Vital Signs	CDASH and SDTM

3.1.2 List of Domains Not Included in this Release

Table 4: *List of Domains Not Included in this Release* provides a list of domains that are not included in the GLIB ALS.

Table 4: List of Domains Not Included in this Release

Domain	Domain Name	Comments
EX	Exposure	EX has very specific rules for what can and cannot go into it, and it is not designed to be used for data collection. Instead, use EC as it is designed for data collection.
MO	Morphology	This domain has been replaced by Body System-based Morphology/Physiology domains (i.e., CV, MK, UR).
OI	Non-host Organism Identifiers	This is not collected data. It is a study-level reference table that will be set up once per study.
SE	Subject Elements	Due to decisions made by NCI during the Treatment Assignment Code (TAC)/Treatment Assignment Description (TAD)/ARM discussions, avoid collecting data that has to do with Subject Elements until further guidance from NCI is available.
SM	Subject Disease Milestones	This is an SDTM-only domain. Data for this domain is derived from variables collected in other domains.
SV	Subject Visits	This is not collected data in Rave.
TA	Trial Arms	These domains will not be developed in the database as they represent study level information and do not contain subject data.
TD	Trial Disease Assessments	
TE	Trial Elements	
TI	Trial Inclusion/Exclusion Criteria	
TM	Trial Disease Milestones	
TS	Trial Summary Information	
TV	Trial Visits	

3.1.3 List of Variables Not Included in this Release

Table 5: *List of Variables Not Included in this Release* provides a list of variables that are not included in the GLIB ALS.

Table 5: List of Variables Not Included in this Release

Variable Name	Variable Label/ Description	Comments/Instructions
DOMAIN	Domain Abbreviation	Variable is not used for data collection.
RDOMAIN	Related Domain Abbreviation	Variable is not used for data collection.
USUBJID	Unique Subject Identifier	Variable is not used for data collection.
STUDYID, SITEID, SUBJID	Study Identifier Study Site Identifier Subject Identifier for the Study	Concepts that are *built in* to the Rave application do not require to repeat these variables in each form. These variables should be defaulted during the study setup and should not be an enterable field.
--SEQ, --DY, --STDY, --ENDY, VISITDY, TAETORD, --STRF, --ENRF, --BODSYS, --BDSYCD, --STRESC, --STRESN, --STRESU, --DRVFL	SDTM Variables	These SDTM variables should be derived based on specific rules for population that cannot be done through data collection.
--LOBXFL, --BLFL	SDTM Variables	The rule for –LOBXFL is: Operationally-derived indicator used to identify the last non-missing value prior to RFXSTDTC. Should be *Y* or null. The baseline flag (--BLFL) can be used to define a different baseline flag rule, if needed. --BLFL can only be defined one time per domain. Usually it is not reliable to collect a value for either of these variables. They should be populated consistently based on the rules defined for them.
TAETORD	Planned Order of Element within Arm	Due to decisions made by NCI during the TAC/TAD/ARM discussions, avoid collecting anything that has to do with Subject Elements in the ALS until the internal implementation is aligned with SDTM.
--DTC	Date and time variables in International Organization for Standardization (ISO) 8601 character format	Not included for the CDASH and SDTM domains as the CDASH date and time variables do not map directly to an SDTMIG variable. All collected CDASH DATE and TIME components are concatenated and populate the SDTMIG variable --DTC in ISO 8601 format.

Variable Name	Variable Label/ Description	Comments/Instructions
--DUR, --STRTPT, --STTPT, --ENRTPT, --ENTPT	Using CDASH Concept instead of SDTM Concept	Avoid having duplicate concepts in the ALS (e.g., CDASH, SDTM). If there is a choice, include the CDASH concept instead of the SDTM concept because the CDASH concept was created to address the needs of data collection and data management. Best practice is to collect any numeric value that has an associated unit in two separate fields. Using --CDUR and --CDURU is more user friendly and supports data management processes. --DUR (from SDTM) was removed and added --CDUR and --CDURU (from CDASH) in the ALS. In addition, SDTM variables –STRTPT, --STTPT, --ENRTPT and --ENTPT were removed and instead added the CDASH variables --PRIOR, --ONGO and --RFTDAT in the ALS.
VISITNUM and --TPTNUM	SDTM Variables	Not recommended as a collection variable. If added to the domain, do not display on the form and use these variables as defined in the SDTMIG and in the Trial Visits (TV) study-level domain for the study.
CODAT, COTIM	Comment Date and Time	CDASH recommends to **only** collect comments (COVAL) in conjunction with other data instead of on a separate Comment CRF. Comments collected with other data inherit the timing (e.g., VISIT, --DAT/--TIM) of the data with which they are collected. SDTM rules do not allow populating CODTC or other timing fields when the comment is linked to another record.
--XFN	External File Path	--XFN is intended to store the name and path of an external file. It is valid to use in Demographics and all Findings Class domains in SDTM to point to a file that is external to the SDTM data but is supporting information (usually a document). If LPOs want to use something similar in data collection, recommend to use a collected version --CXFN. SDTM variable –XFN was removed and replaced with the CDASH variable --CXFN in the ALS.

Variable Name	Variable Label/ Description	Comments/Instructions
--TESTCD (removing all --TESTCD from ALS except IETESTCD)	SDTM Variables	--TESTCD (test short name) would not typically be used in data collection or displayed on the form, with the exception of IETESTCD (display both IETESTCD - 8 character code that represents the individual IE criteria and IETEST- up to 200 characters to describe the IE criteria) on the IE eCRF. --TESTCD is primarily used in CDASH to construct variable names (Field OIDs) for Findings data in a de-normalized structure. When implementing CDASH in a de-normalized structure, create variable names for the Findings using available CDISC Controlled Terminology for –TESTCD as part of the naming convention for the Field OID (e.g., VS_VSORRES_HEIGHT) and then maps the data to the normalized SDTM structure. The associated --TEST value is what is displayed on the Rave Electronic Data Capture (EDC) screen.
--DOSTXT	SDTM Variable	--DOSTXT is for capturing dosing amounts or a range of dosing information collected in text form (e.g., 200-400). There is a CDASH free text field --DSTXT that allows entering the dose in any format. --DSTXT can then be mapped into either the --DOSE or --DOSTXT variables in SDTM depending on whether the values are numeric or non-numeric. Typically, only a single field is used on any given form - either --DOSE (forcing the site to enter a numeric value) or --DSTXT (allowing the site to enter anything) or --DOSTXT (assuming the site will enter a non-numeric value). CDASH recommended to use CDASH variable DSTXT instead of DOSTXT. SDTM variable –DOSTXT was removed and replaced with the CDASH variable --DSTXT in the ALS.
RFENTIM, RFXENTIM, RFPENTIM, RFENDAT, RFXENDAT, RFPENDAT, RFXSTDAT, RFSTDAT, RFSTTIM, RFXSTTIM	SDTM Variables	These are SDTM submission-only variables and they are typically derived from other collected dates. Example: RFXSTDAT is derived from ECSTDAT for the earliest exposure record.

Variable Name	Variable Label/ Description	Comments/Instructions
--MODIFY	Synonym Qualifier Variable	This variable is for internal coding purposes only. If the collected/reported terms in an Interventions or Event domain is modified (e.g., splitting terms, changing spelling) to facilitate coding, the modified verbatim is stored in –MODIFY. This variable has been removed since coding staff are not allowed to modify the collected/reported terms.
EVLINT	SDTM Variable	--EVLINT (Evaluation Interval) requires ISO 8601 Period format when it is a specific period of time. CDASH recommended to use CDASH variable –CEVINT (Collected Evaluation Interval) for data collection. SDTM variable --EVLINT was removed and replaced with the CDASH variable –CEVINT in the ALS.
EVINTX	SDTM Variable	--EVINTX (Evaluation Interval) is a text description of the evaluation interval. It is determined by the Protocol and is usually a pre-populated value on the EDC to remind the site what it is. LPOs can provide protocol-specific text. CDASH recommended to use CDASH variable –CEVINT (Collected Evaluation Interval) for data collection. SDTM variable --EVINTX was removed and replaced with the CDASH variable –CEVINT in the ALS.
MSANMETH	Analysis Method	MS domain is a findings domain that represents drug susceptibility testing results only. This includes phenotypic testing (where drug is added directly to a culture of organisms) and genotypic tests that provide results in terms of susceptible or resistant. Drug susceptibility testing may occur on a wide variety of non-host organisms, including bacteria, viruses, fungi, protozoa, and parasites. Analysis Method is described in the SDTMIG as "Analysis method applied to obtain a summarized result. It describes the method of secondary processing applied to a complex observation result (e.g. an image or a genetic sequence)." This variable was removed due to lack of use case by LPOs.

Variable Name	Variable Label/ Description	Comments/Instructions
MSSPCUFL	Specimen Usability for the Test	This variable should be used with caution. There are specific rules around how this is populated in SDTM. In the SDTMIG, the value will be *N* if the specimen is not usable, and null if the specimen is usable. If LPOs want to allow collection using both Y and N, consider using a collected version by adding *C* as a prefix to the variable fragment (e.g., MS_MSCSPCUFL). This variable was removed due to lack of use case by LPOs.
RSONGO MSONGO	Findings Variables	Since Findings do not typically have end dates (they are usually Point In Time measurements), these variables are not useful in data collection. These tests would end during the study at some point and an end date would be available. If it were used, it would have to mean that an interval collection (e.g., 24 HR Urine Collection) or measurement (e.g., 48 HR Holter Monitoring) has not ended. This variable was removed due to lack of use case by LPOs.
MSPRIOR	Findings Variable	Most of the time if the Finding measurement/result is done before the study, the date of the result would be collected and would not need the --PRIOR field. PRIOR is only used when there is no date available for the Findings result. This variable was removed due to lack of use case by LPOs.

3.2 Standard Conventions

3.2.1 CDISC-compliant Rave GLIB Forms

The CDISC-compliant Rave GLIB forms were set up in the ALS based on the standard conventions indicated in Table 6: *Standard Conventions for CDISC-Compliant GLIB*.

Table 6: Standard Conventions for CDISC-compliant GLIB

Form Attributes	CDISC-Compliant GLIB CDASH Domain	CDISC-Compliant GLIB SDTM Domain	LPO Allowed to Modify During *Study Build*
Form Name	Domain Name *(Example: Adverse Events)*	Domain Name *(Example: Tumor/Lesion Results)*	Yes
Form OID	2 Letter Domain *(Example: AE)*	2 Letter Domain *(Example: TR)*	Yes
Variable OID	Domain + CDASH variable from Cancer Data Standards Registry and Repository (caDSR) (CDE Short Name)	Domain + SDTM variable from Cancer Data Standards Registry and Repository (caDSR) (CDE Short Name)	Yes
Field Name	Variable Label + PID (caDSR) + Version	Variable Label + PID (caDSR) + Version	No
Field OID	Domain + CDASH variable from caDSR (CDE Short Name)	Domain + SDTM variable from caDSR (CDE Short Name)	No
Field Label (PreText)	Question Text or Prompt from CDASH Metadata table or CDASH Model (For Field with no labels, the Field Label will be the CDASHIG Variable Label from CDASH Metadata table)	Question Text or Prompt from CDASH Model or Variable Label from SDTM Metadata table	Yes
Format - Date	dd MMM yyyy	dd MMM yyyy	Yes
Format - Time	HH:nn:ss (24 hr)	HH:nn:ss (24 hr)	Yes
Format- Char (W Dictionary)	Same as caDSR	Same as caDSR	Yes
Format- Char (W/O Dictionary)	Same as caDSR	Same as caDSR	Yes
Format - Numeric	Same as caDSR	Same as caDSR	Yes
Control Type	Text = Character or numeric variable with format <= 50 Long Text = Non-enumerated variables with format > $50 DATE/TIME = Date or time variable Dropdown = Variables having dictionary <=10 entries SearchList = Variables having dictionary > 10 entries	Text = Character or numeric variable with format <= 50 Long Text = Non-enumerated variables with format > $50 DATE/TIME = Date or time variable Dropdown = Variables having dictionary <=10 entries SearchList = Variables having dictionary > 10 entries	Yes

Form Attributes	CDISC-Compliant GLIB CDASH Domain	CDISC-Compliant GLIB SDTM Domain	LPO Allowed to Modify During *Study Build*
SAS Label	Variable Label from SDTM if variable directly mapping to SDTM. If variable not directly mapping to SDTM or does not exist in SDTM Model, use CDASH variable label	Variable Label from SDTM Metadata table	No
Auto-Query for Future Date	Checked for all dates	Checked for all dates	Yes
Data Dictionary Name	Value Domain Long Name (caDSR) + PID + version	Value Domain Long Name (caDSR) + PID + version	No
Data Dictionary Values:			
User Data String	PV Meaning from caDSR	PV Meaning from caDSR	No
Coded Data	PV from caDSR	PV from caDSR	No

3.2.2 Pattern-based Variable Names

SDTM programming is on the critical path and timeline for FDA Center for Drug Evaluation and Research (CDER) and Center for Biologics Evaluation and Research (CBER) submissions. Significant, unexpected delays in submission timelines and increased costs can result when SDTM is not planned for up front. Increasing efficiency in this part of the process is crucial to the success of a submission.

SDTM variable names and dataset names are standardized and predictable, and the naming convention of most SDTMIG variables includes the domain code as the first two characters followed by a root variable that is consistent across domains within a particular SDTM Observation Class. The domain code is also used in naming the SDTM datasets. SDTM programmers still often struggle when creating SDTM datasets because the data coming to them is often not communicated in advance. Some things that make SDTM dataset creation unpredictable and complex include:

1) Data collection can use SDTM variable names for any data that is collected exactly the same way as it will be represented in SDTM, but some data have to be collected in a different way, such as dates and times that are typically collected in two separate fields (and are then concatenated and formatted in ISO 8601 for the SDTM variable), so different variable names have to be used for date and time data collection.

2) Sometimes data collection requires mixing questions from multiple SDTM domains in the same EDC form, making the target domain in SDTM difficult to determine from the EDC form name. While most SDTMIG variables start with the two-character SDTM domain code as a prefix, not all do, so the SDTM programmer cannot write one program to extract the domain code from all variables. For example:

- Most Demographics (DM) domain variables do not contain the DM prefix (e.g., AGE) because they are only used in the Demographics dataset.
- Associated Persons (AP) variables never contain the four-character AP domain code as a variable prefix, so these domains are not reliable for extracting the domain code/dataset name.
- Non-standard (custom) variables use the Supplemental Qualifier structure in SDTM. Supplemental Qualifier variables (e.g., QNAM, QLABEL, QVAL) are normalized and do not contain domain-specific prefixes. The values for QNAM may start with the parent domain code, but are not required to, so they are not reliable for identifying the parent domain for the Supplemental Qualifier record.

3) Typically, the SDTM programmers have to create a study-specific mapping document for every variable in every study, and code and validate study-specific SDTM programming. If anything changes during the study (adding values to a data dictionary, adding or removing vital signs tests from a visit, etc.) they have to update their mapping and update and revalidate their conversion programs.

Use of a consistent, **pattern-based** naming convention can increase predictability of variable names and increase reusability of SDTM conversion programs by allowing SDTM programmers to write pattern-based programming that can automatically:

- Find or create the appropriate domain or dataset;
- Find or create the target variable; and
- Work without having to know in advance what data will come to them from a particular study.

CDASH includes some example pattern-based variables for some Findings-class domains. However, CDASH pattern-based variables are far from exhaustive, and patterns can be useful for all variable names, not just for Findings-class variables. Creating pattern-based variables for most or all data collection fields supports common data collection needs for flexibility while supporting efficiency in SDTM programming. Data collection flexibility needs may include:

- Mixing questions from different domains on a single CRF (e.g., adding height and weight questions to a Demographics form);
- Asking a different set of questions per study or per visit (e.g., Vital Signs with height and weight at screening, and without them at subsequent visits);
- Arranging the forms in different configurations (e.g., multiple timepoint-based ECG measurements within a single visit vs. one ECG at a visit); and
- Collecting Informed Consent for adults and Assent for pediatric participants in the same study.

Rave GLIB forms were set up in the ALS following a pattern-based variable naming convention:

1) Pattern-based variable naming convention:

   [targetDatasetAbbreviation_ targetVariable_optionalPrespecifiedValue]

   Example 1: For collecting age in the Demographics domain:

   - The target dataset abbreviation is the two-character domain code for Demographics=DM.

- The target variable is AGE.
- There is no pre-specified value, so the resulting pattern-based variable is: DM_AGE.

Example 2: For collecting a temperature result in the Vital Signs domain:

- The target dataset abbreviation is the two-character domain code for Vital Signs=VS.
- The target variable is VSORRES.
- The pre-specified value for this record is the VSTESTCD representation of temperature=TEMP.
- The resulting pattern-based variable is: VS_VSORRES_TEMP.
- VS_VSORRESU_TEMP would be the associated variable for the temperature unit.

Example 3: For a non-standard variable collecting whether the subject discontinued due to an adverse event:

- The target dataset abbreviation is SUPPAE (Supplemental Qualifiers for Adverse Events).
- The target variable is QVAL.
- The pre-specified value comes from the question we asked and is represented by its QNAM value=AEDIS.
- The resulting pattern-based variable is SUPPAE_QVAL_AEDIS.

One limitation of this pattern is that if the prespecifiedValue has multiple words (e.g., Informed Consent), this requires an abbreviation to be created for the variable name. For example, Informed Consent can be abbreviated to IC (e.g., DS_DSSTDAT_IC) for the variable name. A standard mapping should be created for each abbreviation needed (e.g., IC for Informed Consent, T2DM for Type 2 Diabetes Mellitus) to support conversion to the full value during the SDTM creation process. These abbreviations and their associated mapping should be standardized as part of the GLIB implementation.

Example 4: For collecting pre-specified family medical history of Type 2 Diabetes Mellitus:

- The target dataset abbreviation is APMH (Associated Persons Medical History).
- The target variable is MHOCCUR (answered by Y or N).
- The pre-specified medical history value which will populate MHTERM is Type 2 Diabetes Mellitus and is represented by a standard abbreviation=T2DM.
- The resulting pattern-based variable is: APMH_MHOCCUR_T2DM.

Note: Applying the re-use rule, multiple version of the variable APMH_MHOCCUR_T2DM may need to be created for different biological relatives (e.g., APMH_MHOCCUR_T2DM_01, APMH_MHOCCUR_T2DM_02).

2) Instructions for pattern-based variable creation:

- Start each variable name with the two- to eight-character dataset name (usually this will be the two-character domain code) followed by an underscore;
- Add the target variable name;

- If needed to make the variable name unique, add another underscore followed by the prespecifiedValue (e.g., pre-specified Topic value like a value from --TESTCD for Findings records, or the QNAM value for SUPP-- records); and
- An additional, optional two-digit value may be added if the variable will be used more than once on that form, for example if Respiratory Rate will be collected at multiple timepoints within the same visit: VS_VSORRES_RESP_01, VS_VSORRES_RESP_02.

3) Instructions for pattern-based SDTM programming:

- Scan the characters up to the first underscore, find or create that dataset (domain);
- Scan the characters up to the end (or to the next underscore), find or create that variable in the target dataset and populate the target variable with the value that was collected in this variable; and
- If there is a second underscore, use the value after that to populate the appropriate value in the record, such as:
    - The Topic variable for domain-based data; or
    - The QNAM for SUPP-- data.

3.3 Additional Standards

Table 7: *List of Items with Special Instructions* provides a list of variables/items and recommendations/ instructions for their use in study build while following the CDISC standards.

Table 7: List of Items with Special Instructions

Variable/Item	Description	Recommendations/Instructions
--DECOD, --PTCD, --LLT, --LLTCD, --HLT, --HLTCD, --HLGT, --HLGTCD, --SOC, --SOCCD	Coding Variables	These coding variables are specifically designed for use with the MedDRA structure, but may be useful with other dictionaries. These variables should only be set up to support the internal coding process (i.e., derive from the MedDRA coding dictionary for Adverse Events) and should **not** be available for use by the site. For –DECOD (which can be used to select values from a provided list of values in the form), CTSU created a dummy coding dictionary (e.g., LPO_TBD) and attached it to the --DECOD coding variable. For AEDECOD LPOs are required to change the control type and attached the MedDRA defined dictionary in their study build. There are separate codelists used for DSDECOD. Codelist NCOMPLT is used for disposition events and codelist PROTMLST is used for protocol milestones. The disposition codelist (CDISC_SDTM_REAS_PID6352263_V1_0F data dictionary) was attached to the DSDECOD variable in the ALS. If LPOs need to use the protocol milestones codelist, work with the NCI CDISC Harmonization Working Group (WG) to define the appropriate codelist values.

Variable/Item	Description	Recommendations/Instructions
--CLAS, --CLASCD, --ATC1, --ATC1CD, --ATC2, --ATC2CD, --ATC3, --ATC3CD, --ATC4, --ATC4CD, --ATC5, --ATC5CD	Coding Variables	These coding variables are specifically designed for use with the WHO-DD structure, but may be useful with other dictionaries. These variables should only be set up to support the internal coding process (i.e., derive from the WHO-DD coding dictionary for Concomitant Medication) and should **not** be available for use by the site.
Partial Dates	Date Format	If date fields must allow for the entry of partial dates, Rave provides different date formats that allow partial dates entry.
EPOCH, VISIT, --SPID, --TEST, --CAT, --SCAT, --TPT, --GRPID, --LNKID, --LNKGRP	SDTMIG Variables Used for Display Only	These fields should not be used for data entry, but should be used for displaying a value on the form.
--TEST, --CAT, --SCAT, --TPT *Note: These are some example of variables that are intended to use controlled terminology. This list may not be complete.*	Variables Required Control Terminology	Certain CDASHIG and SDTMIG variables are intended to use controlled terminology even though a codelist is not specified for them in the Implementation Guides. These variables require sponsor-defined terminology. CTSU created a dummy data dictionary (e.g., LPO_TBD) and attached it to these variables. LPOs need to attach the required sponsor-defined terminology in study build, except AEREL, which will use the CTSU Standard form dictionary. In addition, if LPOs add a variable from SDTM that requires terminology but is not listed here and does not have a specific codelist defined, the Implementation Notes in the CDASH Metadata table could be useful to help LPOs understand when controlled terminology is needed for the variables.
SDTMIG Variables	Creating PreText	If LPOs want to add in variables from SDTMIG that are not already existing in that domain in the CDASHIG and SDTM-only domains, use the root metadata Question Text from the CDASH Model to create PreText.
--AENO, --MHNO	Identifier Variables	For the --AENO and --MHNO variables, consider writing a custom function that will build search lists for all of the AEs and all of the MHs that are entered, and make the search list available on the forms that use those variables (e.g., CM, PR).

<table>
<tr><th>Variable/Item</th><th>Description</th><th>Recommendations/Instructions</th></tr>
<tr><td>CDASH Concept vs. SDTM Concept</td><td></td><td>Avoid having duplicate concepts in the ALS (e.g., CDASH, SDTM), and if there is a choice, include the CDASH concept instead of the SDTM concept because the CDASH concept was created to address the needs of data collection and data management. An example is the use of SDTM --DUR vs. the use of CDASH --CDUR and --CDURU. Best practice is to collect any numeric value that has an associated unit using two separate fields to allow for edit checks, etc. --DUR has a required format using the ISO 8601 Period standard, which would require the site to enter data using that format (e.g., P2W). Using --CDUR and --CDURU is more user friendly and supports data management processes better.

Other examples of preferring the CDASH concept over the SDTM concept are:
<table>
<tr><th>CDASH Concept</th><th>SDTM Concept</th></tr>
<tr><td>--DAT and --TIM</td><td>--DTC</td></tr>
<tr><td>--PRIOR, --ONGO, --RFTDAT</td><td>--STRTPT, --STTPT, --ENRTPT, --ENTPT</td></tr>
</table></td></tr>
<tr><td>--DOSFRM, --DOSFRQ, --ROUTE</td><td>CDASH Subset Codelists</td><td>For CDASHIG variables that have both a complete SDTM codelist and a subset CDASH codelist, the subset codelists were attached to the CDASH variables that come from caDSR. LPOs can modify the subset codelists to include other values from the complete codelists at the study level, and create subsets using values from the complete SDTM codelists.

Following are examples of CDASH codelists that have been subsetted from the complete SDTM codelists:
<table>
<tr><th>CDASH
Subset Codelists</th><th>SDTM
Complete Codelists</th></tr>
<tr><td>--DOSFRM</td><td>FRM</td></tr>
<tr><td>--DOSFRQ</td><td>FREQ</td></tr>
<tr><td>--ROUTE</td><td>ROUTE</td></tr>
</table></td></tr>
<tr><td>Questionnaires (QS), Functional Tests (FT),
Disease Response and Clin Classifications (RS)</td><td>Domains</td><td>Questionnaires, Ratings, and Scales (QRS) Instruments: These domains are governed by both the SDTMIG domain sections AND the relevant QRS Supplements (https://www.cdisc.org/foundational/qrs). For all published QRS Supplements, there are related controlled terminology codelists published for QSTESTCD, QSTEST, QSCAT, FTTESTCD, FTTEST, FTCAT, RSTESTCD, RSTEST, RSCAT, and some other variables. Reference the individual Supplements for more information on which fields require Controlled Terminology.</td></tr>
</table>

Variable/Item	Description	Recommendations/Instructions
Rave Field Names	Field Naming Convention	The default naming convention is the *caDSR CDE Long Name + CDE PID + version#*. LPOs can use the *caDSR CDE Long Name + CDE PID + version#* or choose to use the *caDSR CDE Short Name + CDE PID + version#* as the Rave field name. If the LPO chooses to use *caDSR CDE Short Name + CDE PID + version#* for the Rave field name, they are responsible for this update in Rave.

Variable/Item	Description	Recommendations/Instructions
Question Text vs. Prompt		As an alternative to using the complete Question Text, LPOs can use the CDASH Prompts for PreText. In some cases, the Prompt provides a more clear explanation of what should appear on the form (e.g., Category values should be displayed rather than asking *What is the category?*). Some text of the Question Text or Prompt is presented using brackets [ ], parentheses ( ) and/or incorporating forward slashes. These different formats are used to indicate how the Question Text or Prompt should be modified by the LPOs. • The text inside the brackets provides an option on the tense of the question, or text that can be replaced with protocol-specific verbiage. • The text inside the parentheses provides options (e.g., singular/plural) or text that may be eliminated. • Text separated with a forward slash provides optional words that the LPOs may choose. Example: Add a question to a CRF that asks whether a lab specimen was collected using a yes or no response. • Question Text (recommended text): [Were any/Was the] [--TEST/ topic] [measurement(s)/test(s)/examination(s)/specimen(s)/sample(s)] [performed/collected]? • Question Text (modified text): Was the laboratory specimen collected? In the first set of brackets, the text option *Was the* is selected as the study required only one lab test to be performed. *[Were any/**Was the**]* In the second set of brackets, the text used is *laboratory* which is the topic of interest. *[--TEST/**Topic (laboratory)**]* In the third set of brackets, the text option *specimen* without the optional *s* is selected. *[measurement(s)/test(s)/examination(s)/**specimen**(s)/sample(s)]* In the fourth set of brackets, the text option *collected* is selected. *[performed/**collected**]*

Variable/Item	Description	Recommendations/Instructions
ALS including domains based on the Findings Class	Multiple use cases of the same field	The ALS includes some domains that are based on the SDTM/CDASH Findings Observation Class. These domains **may** contain multiple copies of the same field because CDASH presents multiple use cases (e.g., core lab vs. local lab) and some CDASH domains include de-normalized examples. The intention is that the LPOs will build a relevant set of fields for each of these domains based on **tests** that are required for the studies.
RACE vs. CRACE		Only use RACE or CRACE, but not **both** for any given study. Use RACE when the five designations for race used by the FDA are collected (American Indian or Alaska Native; Asian; Black or African American; Native Hawaiian or Other Pacific Islander; White). Use CRACE when more detailed race categorizations are collected (e.g., more than the five minimum designations for race used by the FDA).
ETHNIC vs. CETHNIC		Only use ETHNIC or CETHNIC, but not **both** for any given study. Use ETHNIC when values are being collected using the exact non-extensible ETHNIC codelist (C66790) values. Use CETHNIC when values are collected using the NCI Thesaurus codelist for Ethnicity As Collected (C128690), the extended HL7 hierarchy of codelist values, or other Regulatory Agency specific controlled terminology for Ethnic Group. LPOs may append a suffix to denote multiple collected ethnicities (e.g., CETHNIC1, CETHNIC2).
CO	Comments Domain	CDASH does **not** recommend to create a separate Comments form. The COVAL variable may be added to any forms when requires to collect comment.
--PERF vs. --STAT		Either or both of these variables may be used on the form. Use --PERF when a *Yes/No* question is used. Use --STAT when LPOs can collect the value as *Not Done* --PERF is useful at the form or form section level, (i.e., general prompt question to be used as a data management tool to verify that missing results are confirmed missing). --STAT may be used on the --TEST level, (i.e., use to indicate that data are not available by having the site recording the value as *Not Done)*.

Variable/Item	Description	Recommendations/Instructions
Reusing/Subsetting a Data Dictionary		If a CDASH/SDTM variable from the CDISC-compliant GLIB ALS with a data dictionary is used in one form, the same variable cannot be reused in another form without a data dictionary in Rave. Rave does not require reused data dictionaries to have a unique dictionary name/appended integer at any level, study, or form. Rave does require subsetted data dictionaries to have a unique dictionary name/appended integer at the study/form level. • If subsetting a dictionary directly in Rave, use the standard dictionary naming convention and append an integer in front of the *F*. • If subsetting a dictionary in the caDSR FormBuilder and importing the form into Rave via the Object Cart Importer (OCI), the OCI will append an integer to the dictionary name in front of the *F*.
Reusing a Variable		A variable can be reused multiple times within the same form in Rave by appending an integer to the Rave Field OID for each occurrence of reuse (e.g., MHLLT_01, MHLLT_02, MHLLT_03). • If building a form directly in Rave, append an integer to the end of the Rave Field OID for each occurrence of reuse. • If building a form in the caDSR FormBuilder and importing the form into Rave via the OCI, the OCI will append an integer to the end of the Rave Field OID for each occurrence of reuse.
--PRESP		--PRESP is not a collected value. It is used as a hidden field on a CRF that is defaulted to *Y* when the event is pre-specified and *null* for event reported as free-text during the SDTM-based dataset creation. However, it can be used as a derived variable, for example, to solicited adverse event for the AE form in the Standard Form ALS where the value *YES* is derived and implies the adverse event should be asked to the subject *Did you experience this AE?* –PRESP variables in GLIB ALS include a codelist that contains *YES* and other values. LPOs should attach a subset codelist that contains only the *YES* value in their study build.
--ORREF		--ORREF is reference result for continuous measurements in original units. The reference value is collected when the test is continuous. For example, when spirometry tests are performed, the subject usually makes several efforts. Each test produces results, but only the best result is used in analyses. Spirometry test values are compared to a predicted value, rather than a normal range. Predicted values are represented in REORREF.
COEVAL	SDTM Variable	The COEVAL variable is intended to use controlled terminology even though a codelist is not specified in the SDTMIG. LPOs are requested by the CDISC team to use the EVAL codelist in their study build since this codelist is expected in a future version of the SDTMIG.

Variable/Item	Description	Recommendations/Instructions
AEACN	CDASH Variable	Variable AE_AEACN (if used by LPOs) is expected for SDTM reporting and should be included in the dataset for reporting. It should be used to capture a single study treatment action. To capture AE Action at the agent level, new custom variables (AE_AEACN1, AE_AEACN2, AE_AEACN3) should be added for each agent and the labels for these fields can be agent specific such as Action taken (Agent 1), Action Taken (Agent 2).
PPSTINT PPENINT	Planned Start/End of Assessment Interval	These variables are intended to show in the SDTM Pharmacokinetics Parameters (PP) dataset the planned start and end of an interval relative to the time point reference which is derived from another domain or from another PP record. Further, they must be in ISO 8601 Period format.
--TSTDTL	Measurement, Test or Examination Detail	--TSTDTL is a relatively new SDTM concept, and the Submissions Data Standards (SDS) team has not yet fully scoped/defined this variable and they do not recommend its use outside of the specific use cases/examples provided in the published SDTMIG documentation (i.e., FA, MB, MI, MS, OE, and UR Findings class domains).
RSPRIOR FTPRIOR		Most of the time if the Finding measurement/result is done before the study, the date of the result would be collected and would not need the --PRIOR field. PRIOR would only be used to indicate that the measurement was done prior to the study, or prior to some other timepoint, **if an actual date was not available.**

4. CDISC Implementation Workflow

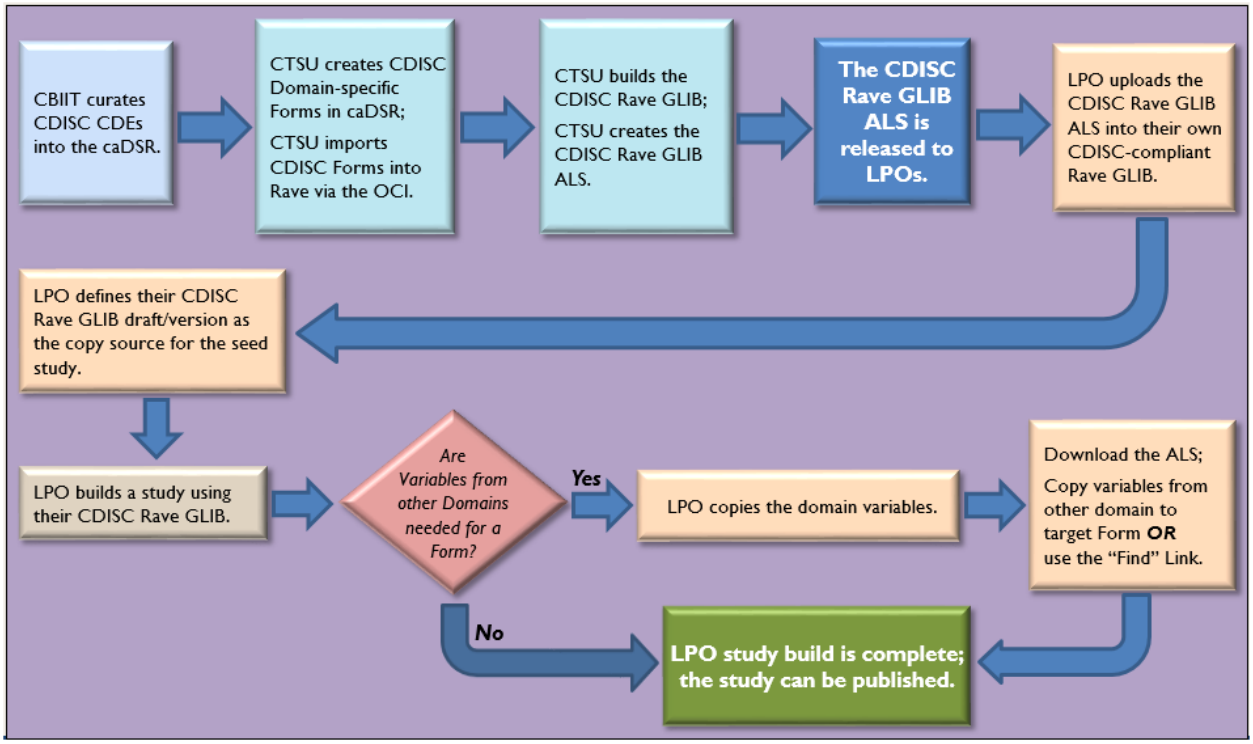


Figure 1: CDISC Implementation Workflow

5. How to Use the CDISC-compliant Rave GLIB ALS for Study Build

The CDISC implementation workflow identifies that LPOs will use the CDISC-compliant Rave GLIB ALS to build their own CDISC-compliant Rave GLIBs to be used as the copy source for the seed studies.

5.1 Creating the LPO CDISC-compliant GLIB in Rave

Using the CDISC-compliant GLIB ALS provided by CTSU, LPOs build the CDISC-compliant GLIB draft in the Rave Global Library Volume (GLV).

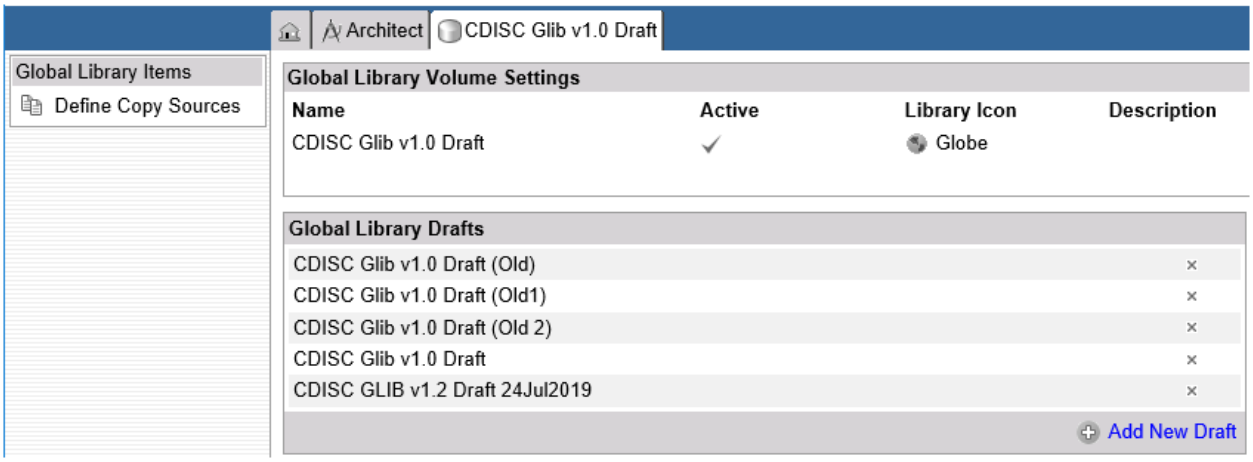


Figure 2: CDISC-compliant GLIB Draft in GLV

Before any objects can be copied from a GLV to a particular project, the GLV must be defined as a Copy Source.

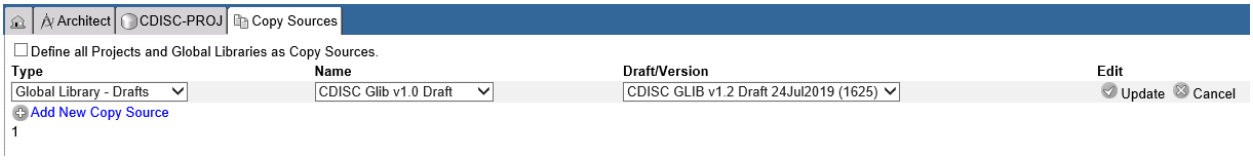


Figure 3: Creating Copy Source CDISC Glib v1.0 Draft

5.2 Options for Building Study Forms Using the CDISC GLIB

There are two options for building study forms using the CDISC-compliant Rave GLIB, copying forms via copy sources in Rave, and copying forms via download ALS.

5.2.1 Copying Forms via Copy Sources in Rave

1) Create a blank project draft.
2) Use the Copy Wizard to copy objects from a copy source into a project draft. On the draft page within a Project, click the *Copy to Draft* link in the Global Library Wizards section of the sidebar.

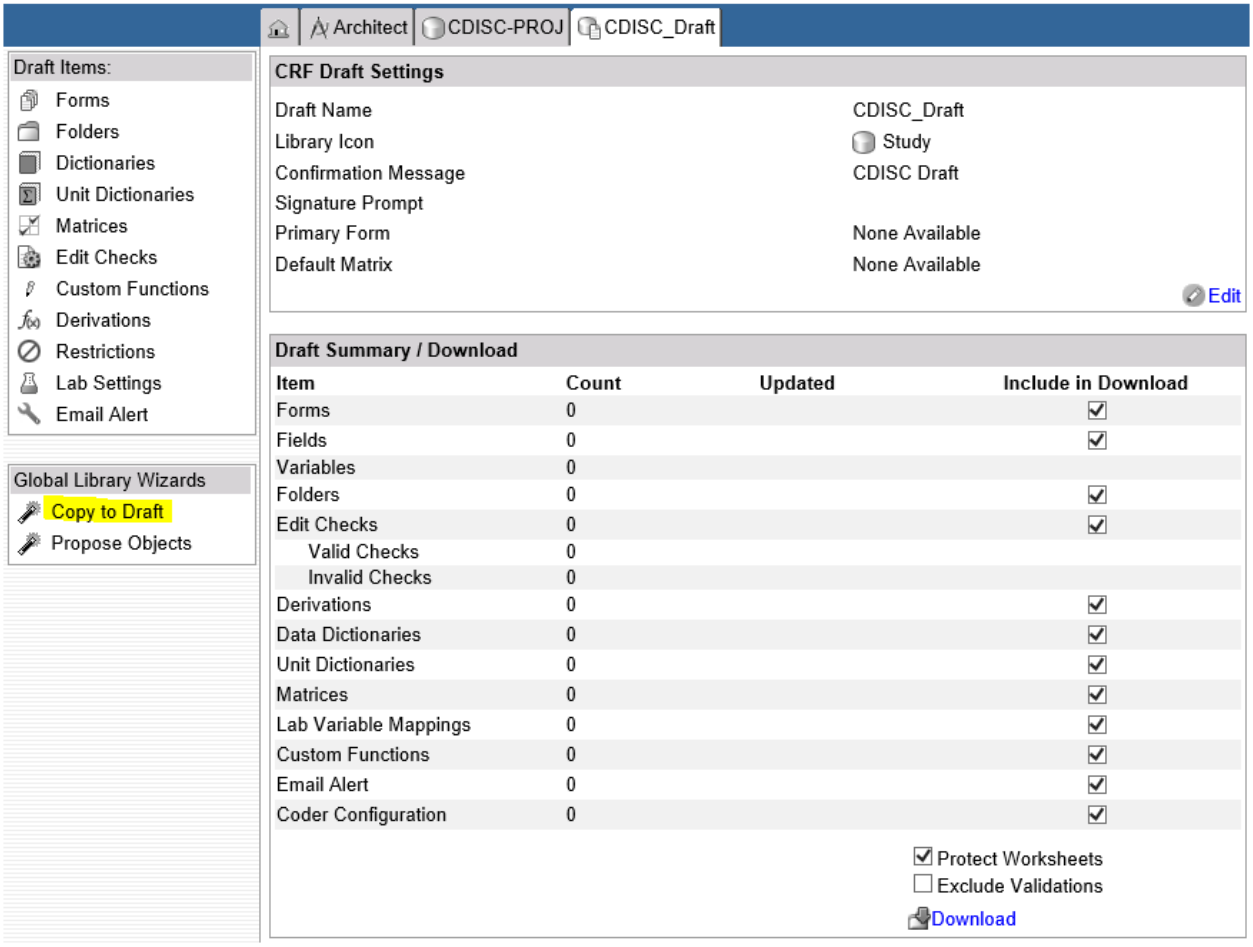


Figure 4: Copy to Draft Link in the Global Library Wizards

3) Select the draft from the copy source by checking the box to the left of a draft.

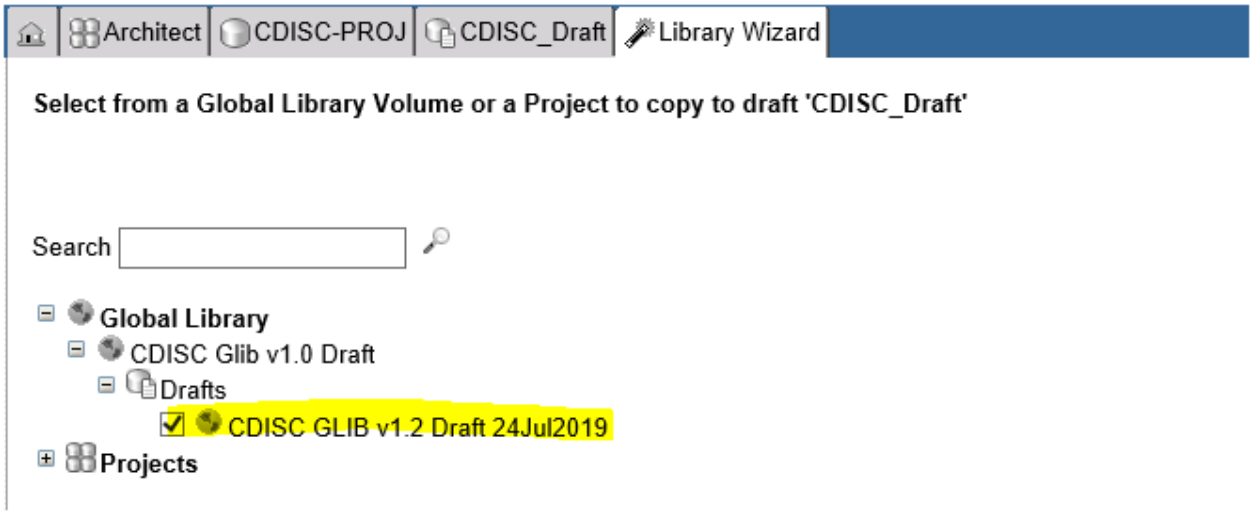


Figure 5: Selecting the Draft as the Copy Source

4) Within the copy wizard, each object type (e.g., Forms, Data Dictionaries) has its own page. Check the box next to the specific forms (e.g., AE, CM, DM, EC) or variable (e.g., DM_DTHDAT from the Disposition form) to copy into a project draft.

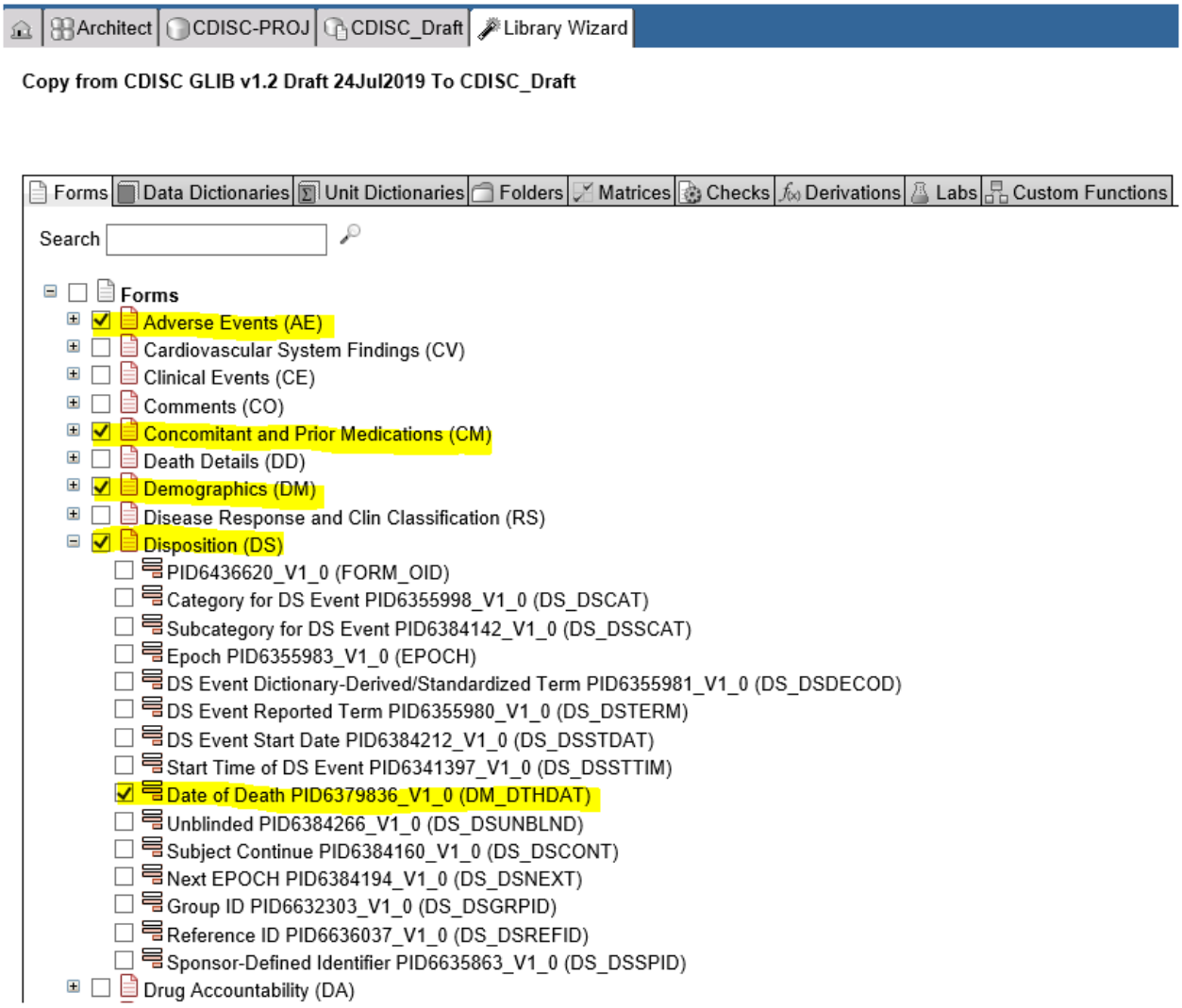


Figure 6: Copy Wizards

> *Note: If the Form box is checked, all children (e.g., variables and data dictionaries) are automatically selected to be copied into the target draft.*

5) Once the form(s) and data dictionaries are copied from the CDISC-compliant Rave GLIB into the project draft, modify the forms per the study requirements:

- Open the form in Rave and click the *Add New* link to add CDISC variables from other forms;
- Use the *Find* link on the variable definition screen to select the new variable and click *Apply* to copy; and

> *Note: The Find link copies only the Variable OID, Format and Data/Unit dictionary; the associated Field attributes are not copied.*

- Manually add any new custom variables that do not exist in the CDISC GLIB.

5.2.2 Copying Forms via Download ALS

1) Use the Copy Wizard to copy objects from a copy source into a project draft (see Section 5.2.1 Items 1 to 4 for instruction on copying objects).

2) Once the form(s) and data dictionaries are copied from the CDISC-compliant Rave GLIB into the project draft, download the ALS from the project draft.

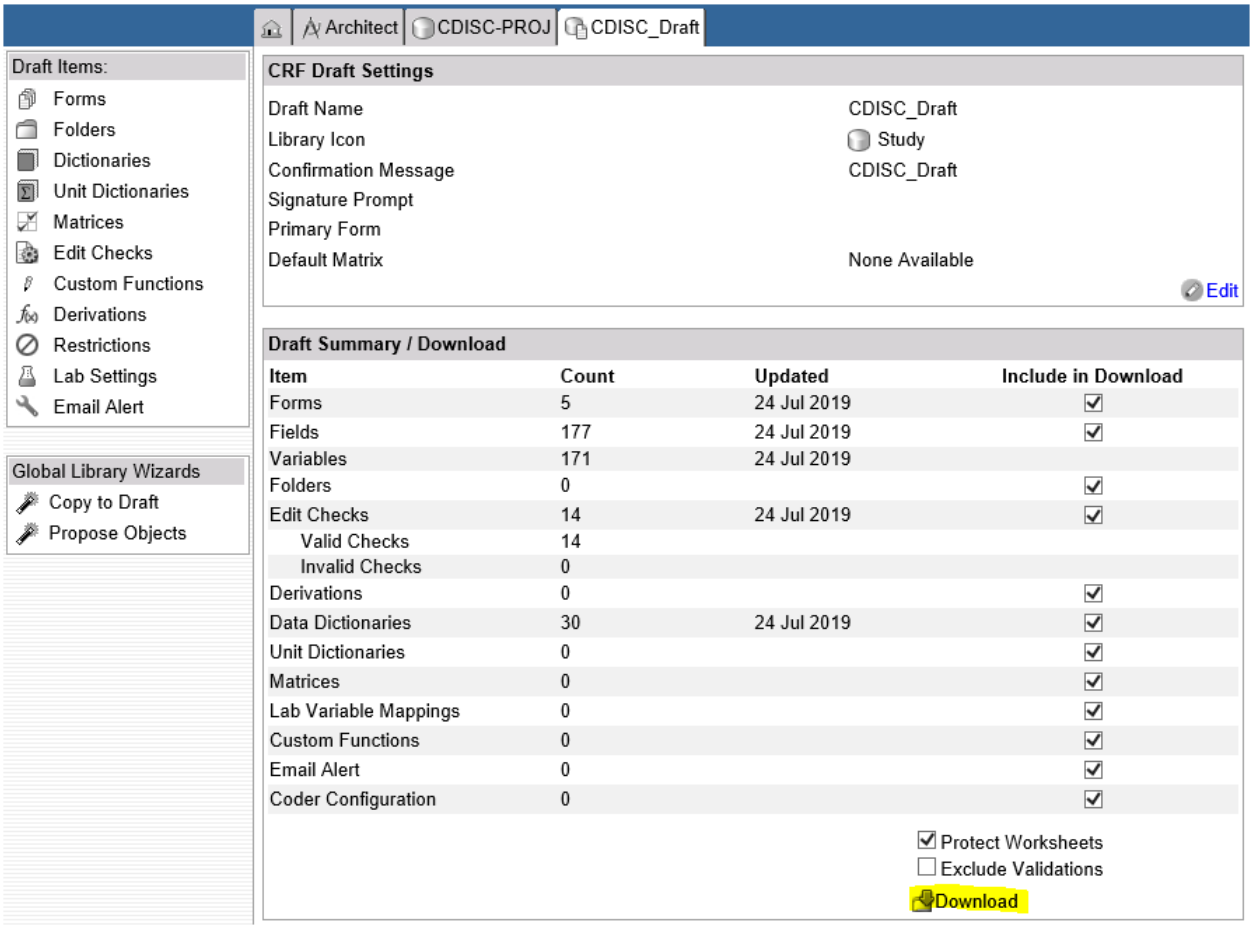
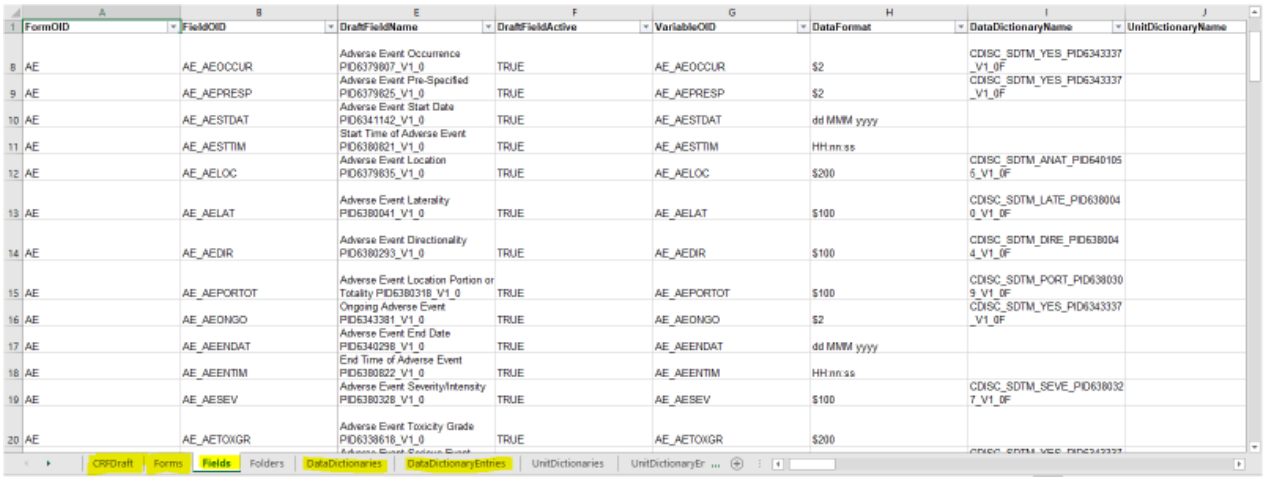


Figure 7: Downloading the CDISC-compliant GLIB ALS

3) Modify the ALS by copying and pasting variables and/or data dictionaries from other domains to the project draft, and/or deleting variables not required for the project *(Note: This document does not include instruction for adding custom variable to the ALS. The process for custom variable curation will be managed by the NCI NRDS committee)*. The following tabs in the ALS must be updated for the project-specific draft:

- CRFDraft
- Forms
- Fields

- DataDictionaries
- DataDictionaryEntries



	A	B	E	F	G	H	I	J
1	FormOID	FieldOID	DraftFieldName	DraftFieldActive	VariableOID	DataFormat	DataDictionaryName	UnitDictionaryName
8	AE	AE_AEOCCUR	Adverse Event Occurrence PID6379807_V1_0	TRUE	AE_AEOCCUR	$2	CDISC_SDTM_YES_PID6343337_V1_0F	
9	AE	AE_AEPRESP	Adverse Event Pre-Specified PID6379825_V1_0	TRUE	AE_AEPRESP	$2	CDISC_SDTM_YES_PID6343337_V1_0F	
10	AE	AE_AESTDAT	Adverse Event Start Date PID6341142_V1_0	TRUE	AE_AESTDAT	dd MMM yyyy		
11	AE	AE_AESTTIM	Start Time of Adverse Event PID6380821_V1_0	TRUE	AE_AESTTIM	HH:nn:ss		
12	AE	AE_AELOC	Adverse Event Location PID6379835_V1_0	TRUE	AE_AELOC	$200	CDISC_SDTM_ANAT_PID6401055_V1_0F	
13	AE	AE_AELAT	Adverse Event Laterality PID6380041_V1_0	TRUE	AE_AELAT	$100	CDISC_SDTM_LATE_PID6380040_V1_0F	
14	AE	AE_AEDIR	Adverse Event Directionality PID6380293_V1_0	TRUE	AE_AEDIR	$100	CDISC_SDTM_DIRE_PID6380044_V1_0F	
15	AE	AE_AEPORTOT	Adverse Event Location Portion or Totality PID6380318_V1_0	TRUE	AE_AEPORTOT	$100	CDISC_SDTM_PORT_PID6380309_V1_0F	
16	AE	AE_AEONGO	Ongoing Adverse Event PID6343381_V1_0	TRUE	AE_AEONGO	$2	CDISC_SDTM_YES_PID6343337_V1_0F	
17	AE	AE_AEENDAT	Adverse Event End Date PID6340298_V1_0	TRUE	AE_AEENDAT	dd MMM yyyy		
18	AE	AE_AEENTIM	End Time of Adverse Event PID6380822_V1_0	TRUE	AE_AEENTIM	HH:nn:ss		
19	AE	AE_AESEV	Adverse Event Severity/Intensity PID6380328_V1_0	TRUE	AE_AESEV	$100	CDISC_SDTM_SEVE_PID6380327_V1_0F	
20	AE	AE_AETOXGR	Adverse Event Toxicity Grade PID6338618_V1_0	TRUE	AE_AETOXGR	$200		

CRFDraft | Forms | Fields | Folders | DataDictionaries | DataDictionaryEntries | UnitDictionaries | UnitDictionaryEr ...

Figure 8: Project Draft ALS

4) Upload the ALS into the project draft in Rave.

5.3 Possible Scenarios in Form Build

5.3.1 Building Forms from Single Domain

All variables in a form are from a single domain (e.g., Concomitant Medication (CM)). There are two methods to copy a form for study build:

1) Copy the form via copy source:
   - Use copy wizard to select and copy the CM form, variables, and data dictionaries from the copy source to the project draft.

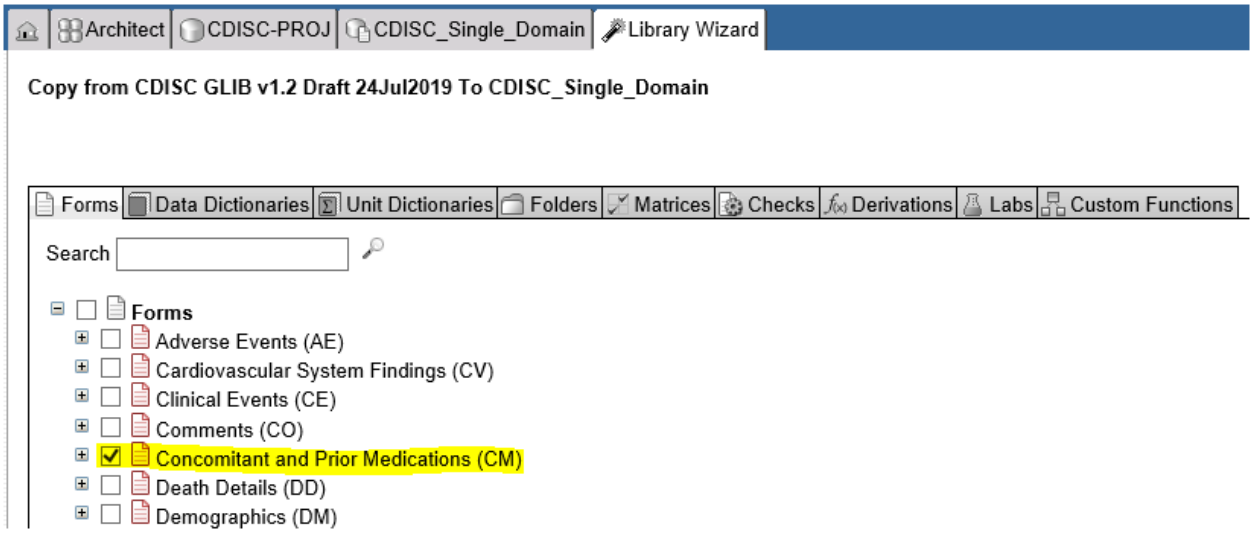
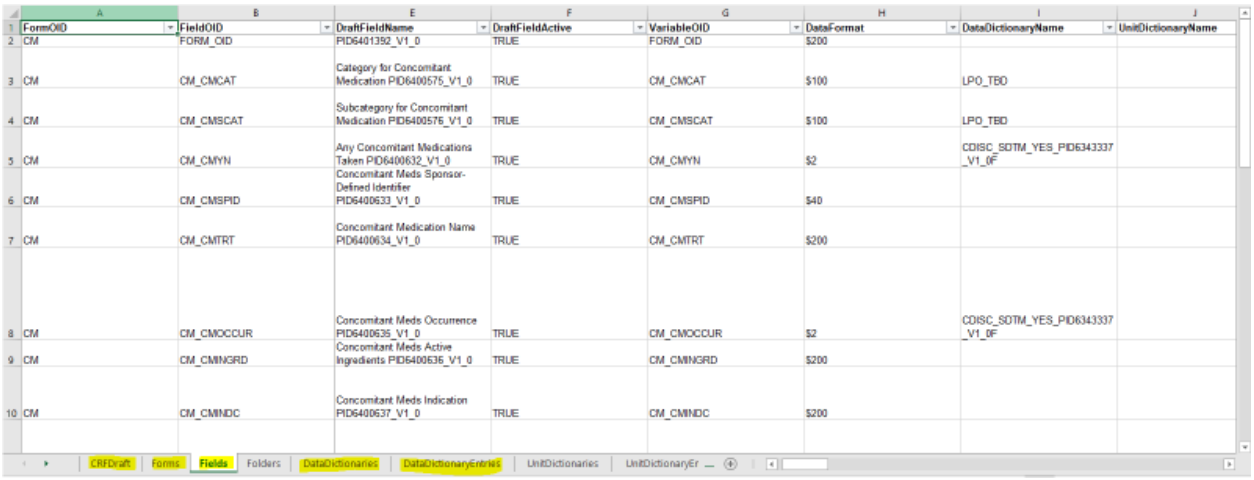


Figure 9: Copying Form from Single Domain via Copy Wizard

2) Copy the form via download ALS:

- Copy the CM form from the Form, Fields, DataDictionaries, and DataDictionaryEntries tabs in the ALS to the project draft ALS; and



	A	B	E	F	G	H	I	J
1	FormOID	FieldOID	DraftFieldName	DraftFieldActive	VariableOID	DataFormat	DataDictionaryName	UnitDictionaryName
2	CM	FORM_OID	PID6401352_V1_0	TRUE	FORM_OID	$200		
3	CM	CM_CMCAT	Category for Concomitant Medication PID6400575_V1_0	TRUE	CM_CMCAT	$100	LPO_TBD	
4	CM	CM_CMSCAT	Subcategory for Concomitant Medication PID6400576_V1_0	TRUE	CM_CMSCAT	$100	LPO_TBD	
5	CM	CM_CMYN	Any Concomitant Medications Taken PID6400632_V1_0	TRUE	CM_CMYN	$2	CDISC_SDTM_YES_PID6343337_V1_0F	
6	CM	CM_CMSPID	Concomitant Meds Sponsor-Defined Identifier PID6400633_V1_0	TRUE	CM_CMSPID	$40		
7	CM	CM_CMTRT	Concomitant Medication Name PID6400634_V1_0	TRUE	CM_CMTRT	$200		
8	CM	CM_CMOCCUR	Concomitant Meds Occurrence PID6400635_V1_0	TRUE	CM_CMOCCUR	$2	CDISC_SDTM_YES_PID6343337_V1_0F	
9	CM	CM_CMINGRD	Concomitant Meds Active Ingredients PID6400636_V1_0	TRUE	CM_CMINGRD	$200		
10	CM	CM_CMINDC	Concomitant Meds Indication PID6400637_V1_0	TRUE	CM_CMINDC	$200		

CRFDraft | Forms | Fields | Folders | DataDictionaries | DataDictionaryEntries | UnitDictionaries | UnitDictionaryEr

Figure 10: ALS for the CM Domain

- After modification of the project ALS, upload the ALS into the project draft in Rave.

StudyDesign: Previous Therapy - General Medications (Previous Therapy - General Medications) [CM_UseCase1b]

Previous Therapy - General Medications Taken [igCM_UseCase1b_YN]

1.	Were any medications taken?	**[CMYN]** *[A:N]* ◯ No *[A:Y]* ◯ Yes

	CM Number	Medication	Start Date	End Date	Indication	MH ID	Reason for Discontinuation
2.							

Previous Medication Entry Entry [igCM_UseCase1b_D]

2.1*	What is the medication / treatment identifier? *[read-only]* [CM Number]	**[CMSPID]** N3
2.2*	What was the term for the medication / therapy taken? [Medication]	**[CMTRT]** A200
2.3	What was the start date of the medication / therapy? [Start Date]	**[CMSTDAT]** NReq/Unk / NReq/Unk / NReq (2012-2014)
2.4	What was the end date of the medication / therapy? [End Date]	**[CMENDAT]** NReq/Unk / NReq/Unk / Req (2012-2014)
2.5	For what indication was the medication / therapy taken? [Indication]	**[CMINDC]** A50
2.6	What was the ID of the medical history condition(s) for which the medication was taken? [MH ID]	**[CMMHNO]** N4
2.7	What was the reason for medication / therapy discontinuation [Reason for Discontinuation]	**[CMNCOMPLT]** [clCMNCOMPLT]

Key: [*] = Item is required

Figure 11: Example of Form Annotation - CM (Concomitant Medication)

5.3.2 Building Forms from Multiple Domains

A form contains variables from multiple domains (e.g., Exposure as Collected (EC) form with variables from EC and AE domains). There are two methods to copy the forms and selected variables into a project draft:

1) Copy the forms via copy source:

- Use copy wizard to copy the EC and AE forms to the project draft;

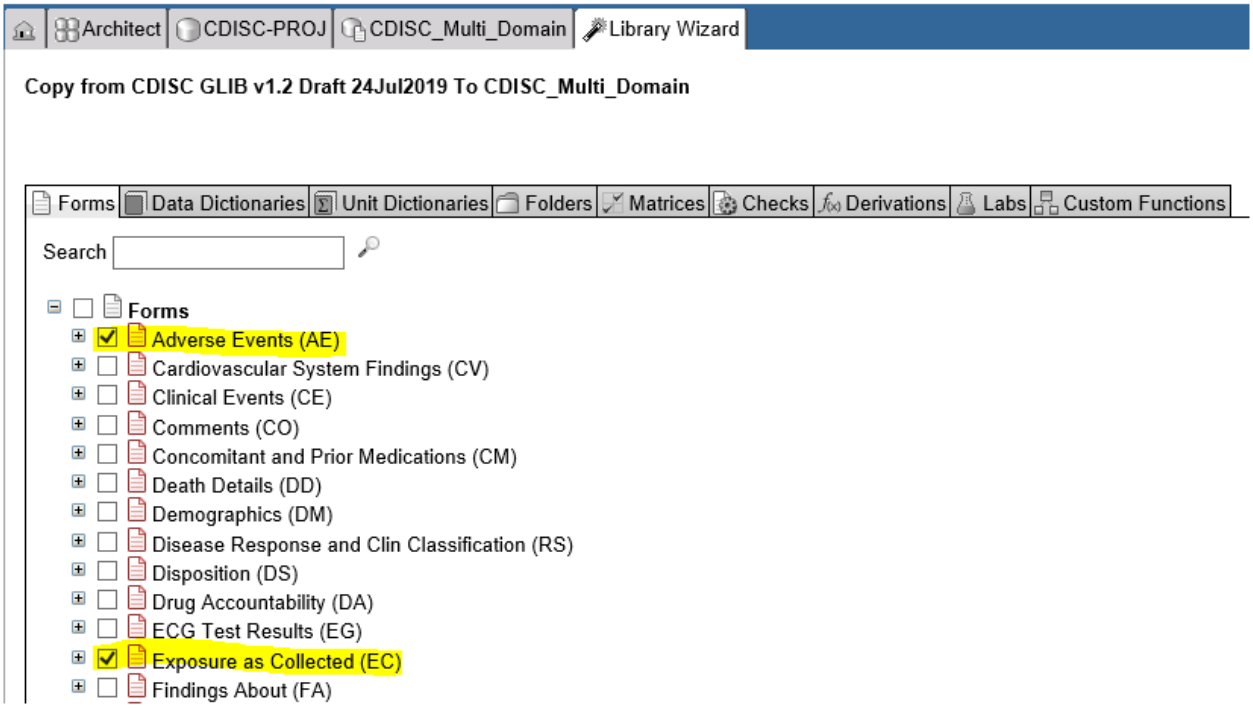


Figure 12: Copying Forms from Multiple Domains via Copy Wizard

- Follow the same steps outlined in *Building Form from Single Domain* (see Section 5.3.1) to build the EC form with variables from EC and AE domains;
- Open the EC form in Rave and click the *Add New* link to add the AE_AESPID variable from the AE form or other forms; and
- Use the *Find* link on the variable definition screen to select the AE_AESPID variable and click *Apply* to copy.

2) Copy the forms via download ALS:

- In the ALS copy the objects from the Forms, Fields, DataDictionaries, and DataDictionaryEntries tabs to the project draft ALS and upload the ALS into the project draft in Rave;
- For example, to copy variable AE_AESPID from form AE to form EC, go to the Fields tab, filter for FormOID = AE and copy the entire row FieldOID = AE_AESPID. Then, filter for FormOID = EC and insert the AE_AESPID row to the EC form; and
- To delete forms or variables that are not required for the project draft, update the Forms, Fields, DataDictionaries, and DataDictionaryEntries tabs to delete the appropriate objects.

5.3.3 Building Forms with Custom Domain and Variables

For a new custom domain (e.g., XV) or variables, update the ALS to add the XV domain to the Forms tab, and add new custom variables to the Fields tab. If a codelist is required for the custom variable, add the codelist to the DataDictionaries tab and the codelist values to DataDictionaryEntries tab. Refer to the

NCI's CDISC Best Practices document for guiding principles and standards on implementing custom domains and variables.

6. Other Modifications per LPO Study Requirements

6.1 Custom Domains

The following information is provided to assist with the creation of custom variables. Please note that all the custom domains/variables are required to follow the NCI CDISC Harmonization WG process.

Use the published, standard domains as much as possible. If necessary, create custom domains based on one of the three General Observation Classes in the CDASH Model. Reference SDTMIG Section 2.6 for additional details on this process.

Process for creating a custom domain:

1) Determine that the data requires a domain that does not already exist in CDASH or SDTM. Always use an existing domain if the data meets the definition for that domain.
2) Access the CDASH Model and choose the appropriate General Observation Class for the data. There are three General Observation Classes available:
   - Interventions: for data about substances the subject takes into their body (This set of metadata includes variables that hold the name of the substance, how much was used, units of consumption, dosage forms, etc.);
   - Events: for data about other things that happen, independent of the study design, but which are being tracked because they are of interest for the study; and
   - Findings: for the results of planned tests, measurements, and other observations. These may have standard reference ranges, methods of testing, lab names, and other similar concepts.
3) From the CDASH Model, add in any Identifiers and Timing variables that are needed for the domain.
4) Create a unique two-character domain code for the custom domain. Domain codes cannot conflict with existing domain codes. To eliminate any risk of using a name that CDISC later determines to have a different meaning, domain codes beginning with the letters X, Y, or Z have been reserved for the creation of custom domains. Any letter or number may be used in the second position. Consider assigning one of these letters to each General Observation Class to facilitate the SDTM programming. E.g., X for Interventions, Y for Events, and Z for Findings.
5) Add the domain code as a prefix to all the custom domain variables.
6) Update the ALS to add the custom domain.

6.2 Custom Variables

For any new variables not included in the CDISC GLIB for a domain, and not already published in an SDTMIG or CDASHIG domain, create variable names that follow the CDISC naming convention using CDASH Model or SDTM Model root variable names prefixed by two-letter domain name code. These are considered standard variables if they are used within a domain based on the same General Observation Class. If a root variable is used from a different General Observation Class, it will become a Supplemental Qualifier in the SDTM domain.

If there is a need to create a truly new, custom variable (called non-standard variables in SDTM) use the variable naming fragments in Section 6.2.2 and prefix them with the domain code. These will become Supplemental Qualifiers in SDTM.

6.2.1 Variable Following the General Observation Classes

The CDASH Model is based on the same three General Observation Classes as the SDTM Model. The CDASH Model organizes data into classes, which represent meaningful groupings of data in clinical research. It defines general observation class variables (Interventions, Events, and Findings), CDASH metadata for identifier variables, and timing variables. The presence of two hyphens before the variable name (e.g., --TRT) is used to indicate the required use of a prefix based on the two-character domain code (e.g., EC). The domain code is used as a variable prefix to minimize the risk of difficulty when merging or joining domains for reporting purposes. Refer to Section 2 of the CDASH Model Version 1.0 for list of the observation class variable attributes to use for implementing each of the domain classes and create custom domains.

6.2.2 Create Custom Variables from CDISC Naming Fragments

Table 8: *List of Fragments* contains a standard list of fragments (published in the SDTMIG Appendix D and used in CDASH) to use as a guide when naming custom variables. Use the fragment(s) that best convey the meaning of the variable meeting the eight-character variable naming limit of SAS transport files.

Table 8: List of Fragments

Fragment	Keyword
ACN	ACTION
ACPT	ACCEPTANCE
ACT	ACTUAL
AD	ANALYSIS DATASET
ADJ	ADJUSTMENT
AGE	AGE
ALIQ	ALIQUOT
ALT	ALERT
AMT	AMOUNT
AN	ANALYZED, ANALYSIS
ANT	ANATOMICAL
APT	ACCEPTED
ARM	ARM
AS	ASSAY
ASY	ASSAY
BL	BASELINE
BOD, BDY, BD	BODY
BRTH	BIRTH
C	CHARACTER
C	COLLECTED

Fragment	Keyword
C	CONSCIOUSNESS
CAN	CANCER
CAT	CATEGORY
CD	CODE
CL	CLINICAL
CLAS	CLASS
CND	CONDITION
CODE	CODE
COM	COMMENT
CON	CONCOMITANT
CONG	CONGENITAL
CONT	CONTACT
COUNTRY	COUNTRY
DECOD	DECODE
DESC, DES	DESCRIPTION
DEV	DEVICE
DIR	DIRECTIONALITY
DISAB	DISABILITY
DOMAIN	DOMAIN
DOS, DOSE	DOSE, DOSAGE
DRG	DRUG
DRV	DERIVED
DTC	DATE TIME CHARACTER
DTH	DEATH
DTL	DETAIL
DUR	DURATION
DX	DIAGNOSIS
DY	DAY
EL	ELAPSED
ELEMENT	ELEMENT
EM	EMERGENT
END, EN	END
EPOCH	EPOCH
ET	ELEMENT

Fragment	Keyword
ETHNIC	ETHNICITY
EVAL	EVALUATOR
EVL, EV	EVALUATION
EXT	EXTENT
FAST	FASTING
FL	FLAG
FN	FILENAME
FRM	FORMULATION, FORM
FRQ	FREQUENCY
GR	GRADE
GRP, GP, G	GROUP
HI	UPPER LIMIT
HL	HIGH LEVEL
HOSP	HOSPITALIZATION
ID	IDENTIFIER
IN	INTRINSIC
IND	INDICATOR
INDC	INDICATION
INT	INTERVAL
INTP	INTERPRETATION
INV	INVESTIGATOR
L	LOWER
LAT	LATERALITY
LEAD	LEAD
LIFE	LIFE-THREATENING
LL	LOWEST LEVEL
LNK	LINK
LO	LOWER LIMIT
LOC	LOCATION
LOINC	LOINC CODE
LOQ	LEVEL OF QUANTIFICATION
LOT	LOT
METH, METHOD, MTH	METHOD
MIE	MEDICALLY-IMPORTANT EVENT

Fragment	Keyword
MODIFY	MODIFIED
MOL	MOLECULAR
MOOD	MOOD
MU	MUTATION
N	NON-HOST
N	NUMERIC
NAM, NM	NAME
ND	NOT DONE
NR	NORMAL RANGE
NST	NON-STUDY THERAPY
NUM	NUMBER
OBJ	OBJECT
OCCUR	OCCURANCE
OD	OVERDOSE
ONGO	ONGOING
OR	ORIGINAL
ORD	ORDER
ORIG	ORIGIN
OTH, O	OTHER
OUT	OUTCOME
P	PREFERRED
PARM	PARAMETER
PARTY, PRTY	PARTY
PATH	PATHOLOGY
PATT	PATTERN
POOL	POOL
POP	POPULATION
POR	PORTION
POS	POSITION
PRC	PROCESSED, PROCESS
PRESP	PRE-SPECIFIED
QUAL	QUALIFIER
RACE	RACE
REAS	REASON

Fragment	Keyword
REF, RF, R	REFERENCE
REG	REGION
REL	RELATIONSHIP, RELATED
REL, R	RELATED, RELEVANT
RES	RESULT
RGM	REGIMEN
RI	REGION OF INTEREST
RL	RULE
ROUTE	ROUTE
RPT	REPORT
RUN	RUN
S, SUB	SUB
S, SER	SERIOUS
SCAT	SUBCATEGORY
SENT	SENT
SEQ	SEQUENCE
SEV	SEVERITY
SEX	SEX
SHIP, SH	SHIPPING
SIG	SIGNIFICANT
SITE	SITE
SOC	SYSTEM ORGAN CLASS
SP	SPONSOR
SPCES	SPECIES
SPEC, SPC	SPECIMEN
SPL	SAMPLE
ST	START
ST, STD	STANDARD
STAT	STATUS
STATE	STATE
STRG	STRENGTH
STRN	STRAIN
STUDY	STUDY
SUBJ, SUB	SUBJECT

Fragment	Keyword
SUPP	SUPPLEMENTAL
SYS, SY	SYSTEM
T	TERM
TAETORD	TAETORD
TERM	TERM
TEST, TST	TEST
TM	TIME
TOT, T	TOTAL
TOX	TOXICITY
TPT, T	TIME POINT
TRANS	TRANSITION
TRT	TREATMENT
TXT	TEXT
TYPE, TYP	TYPE
U	UNIQUE
U	UNIT
UP	UNPLANNED
V	VEHICLE
VAL	VALUE
VAR	VARIABLE
VR	VERSION
VISIT	VISIT
X	EXTERNAL
X	TEXT
X	TREATMENT
XFN	EXTERNAL FILE NAME

6.3 CDASH De-normalized Convention Types

Adding new collection fields is often constrained by business rules, clinical data management processes, and EDC systems. The naming conventions and other variable creation recommendations in CDASHIG are designed to consistently facilitate transforming the collected data into submission datasets. These conventions and recommendations are described in the following sections.

6.3.1 Changing Data Structure from Vertical to Horizontal

If a study requires capturing data in horizontal (de-normalized) structure that is different from the SDTM vertical (normalized) structure, concatenate the variable (e.g., VS_VSTEST_HR, VS_VSORRES_HR) where

part of the variable (e.g., VSTEST, VSORRES) is taken from the SDTM specs and is concatenated with the test name (in this case HR for Heart Rate), using controlled terminology from the appropriate codelist.

De-normalized variables must follow a pattern-based naming convention as outlined in Section 3.2.2. The pattern-based naming convention is provided to assist with the mapping of the SDTM datasets.

In some cases it is also useful or necessary to have the value of the TOPIC variable for the SDTM record in order to create a unique variable name, as in the case of Findings Observation Class domains where the topic is the short name of the **test** and there may be multiple tests on one form. The example in Figure 13: *Example of Creating De-normalized Variables for the Vital Signs Form* illustrates this where we have created multiple Original Result fields that are test-specific.

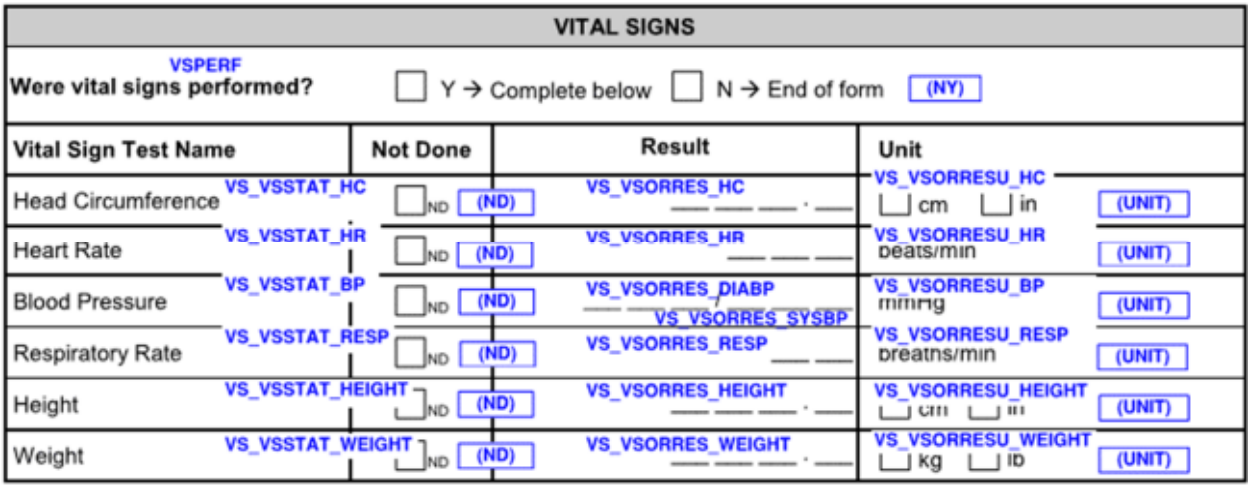


Figure 13: Example of Creating De-normalized Variables for the Vital Signs Form

6.3.2 Changing Data Type

If a study requires a variable to collect a data type that differs from the SDTM data type (i.e., change data type for variable FAORRES from character to numeric), create a new variable name that follows CDASH convention by adding a *C* for *Collected* in front of the variable root (e.g., FACORRES).