USDM Handbook 1: Leveraging USDM Metadata to Automatically Construct the Foundational CDISC Trial Design Model Domains (USDM-HB1 v1.0)

Version 1.0 (Final)

Prepared by the USDM Team

Notes to Readers

This is Version 1.0 of the USDM-WB1 v1.0.

Revision History

Date	Version
2026-06-01	1.0 Final

See Appendix D for Representations and Warranties, Limitations of Liability, and Disclaimers.

CONTENTS

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

1 Introduction

In the evolving landscape of clinical research, the transition from protocol concept to structured data has historically been a manual, fragmented process. The emergence of the Unified Study Definitions Model (USDM; https://www.cdisc.org/ddf) marks a pivotal shift toward a "digital data flow" (DDF) ecosystem. For business stakeholders and data architects, the challenge is no longer just about collecting data—it is about automating the downstream data flow. One of these use cases is the generation of Study Data Tabulation Model (SDTM; https://www.cdisc.org/standards/foundational/sdtm) Trial Design Model (TDM) domains directly from the "single source of truth" documented in a digitized format in the study design phase.

This Implementation Handbook provides a strategic and implementation framework for leveraging USDM metadata to automatically construct the TDM domains defined in the SDTM Implementation Guide (SDTMIG; https://www.cdisc.org/standards/foundational/sdtmig), including:

- TA (Trial Arms)
- TE (Trial Elements)
- TV (Trial Visits)
- TI (Trial Inclusion/Exclusion)
- TS (Trial Summary)

These are the domains usually required for regulatory submissions. Other Trial Design domains, if required, can be created using similar methods. For clinical studies, these include:

- Trial Disease Assessments (TD), which was invented for oncology studies, where "disease assessments" are disease response assessments that can be built up from multiple assessments which may not occur on the same date or in the same encounter.
- Trial Disease Milestones (TM), which defines triggers for performing assessments by an event or finding, rather than scheduled at planned times.

More information, regarding these trial design domains and their requirements can be found in the SDTM Implementation Guide (SDTMIG; https://www.cdisc.org/standards/foundational/sdtmig).

The Business Value Proposition

Transitioning to a USDM-based workflow for SDTM generation is not merely a technical upgrade; it is a business imperative. The traditional manual workflow, which often relies on spreadsheet-based data entry and custom SAS import scripts, is inherently fragile. This process fails when a study protocol is revised or amended, requiring extensive manual rework that is neither repeatable nor dependable. Since TDM datasets are among the first items regulatory reviewers examine to understand study structure, this form of manual technical debt creates significant compliance risk. There are significant benefits to automating the mapping between the study definition and the trial design domains, including:

- **Accelerated time-to-submission:** By defining the trial structure digitally at the protocol stage, the downstream SDTM setup happens in parallel with the study build, rather than as a reactive post-processing step.
- **Robust data traceability:** Automation creates a direct link between the clinical protocol and the regulatory submission package. This interoperability simplifies auditing and reduces the risk of inconsistencies and noncompliance.
- **Operational efficiency:** Eliminating manual "transcription" of study designs into SDTM domains reduces human error, minimizes rework, and frees up clinical programmers to focus on complex analysis rather than structural formatting.

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

Navigating the Digital Transformation

Although the technical mapping is a core component of this Implementation Handbook, the focus is on the integration of people, processes, and technology. Successful implementation requires a shift in how study teams collaborate. The protocol is no longer just a document; it is a structured dataset that informs the entire clinical lifecycle. This Implementation Handbook focuses on the SDTM Trial Design domain use case. However, similar approaches including requirements and mapping transformations can be used for regulatory registry data to be submitted.

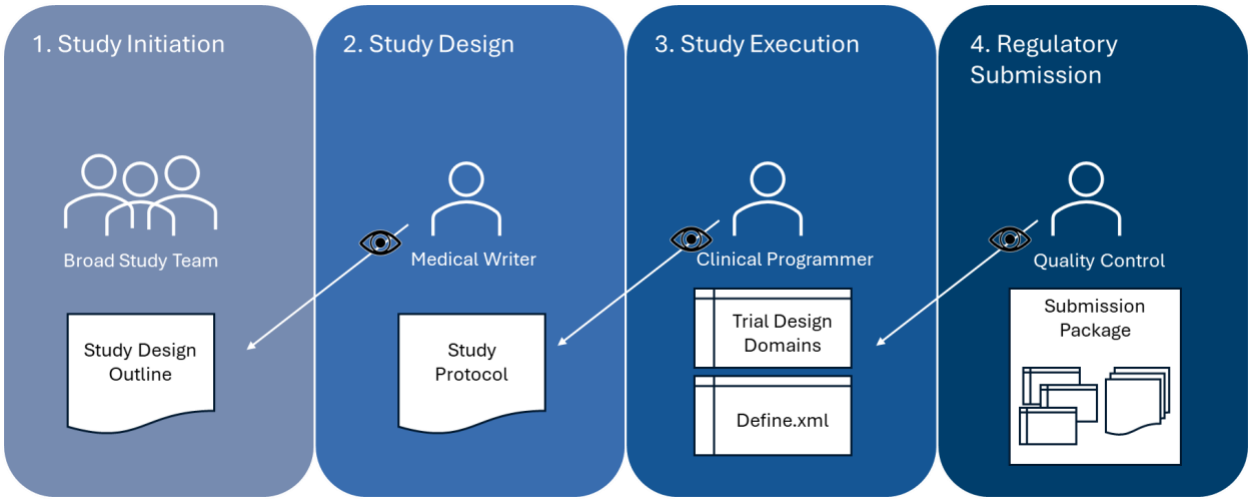
> ***Key Insight:*** *In a USDM-driven environment, the Trial Design domains are the digital blueprint of the study. If the blueprint is accurate at the source, the foundation of the clinical data remains rock-solid through submission.*

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

2 Relevant Workflows and Processes

2.1 Business Level

Current Business Flow



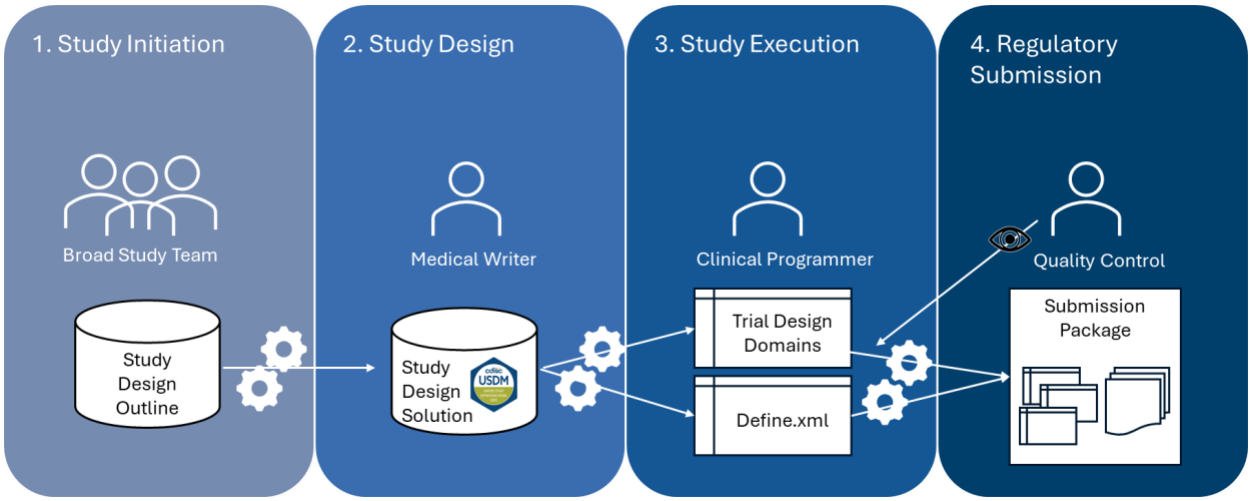
This figure illustrates the typical current business flow from study design to SDTM-formatted regulatory data submission package.

1. The study team decides on the global design of the study (i.e., indication, objectives, endpoints) and creates a study design outline document.
2. The study design outline is interpreted and copied manually (👁) and serves as the starting point for the complete study design described in the protocol by the medical writer.
3. The study design details that are noted in the final study protocol are then again manually copied and interpreted (👁) by a clinical programmer or similar function who is responsible for creating the Trial Design domains. The corresponding metadata needs to be submitted in the define.xml file.
4. The content of the Trial Design domains is validated manually (👁) by a second programmer or specialist before being included in the regulatory submission package.

The creation of SDTM Trial Design domains typically does not start before the actual start of the study. The clinical programmer needs to retrieve all the design details from the protocol document; this is typically stored in a dedicated and prestructured spreadsheet. The study design needs to align with controlled terminology, including external coding dictionaries; the medical expert usually verifies these medical-related codes. Finally, although more than 95% of the information needed for the Trial Design domains is from the study design, some last execution details like actual start and end dates of the study are added based on information captured in the electronic data capture (EDC) system (data flow not included in Figure 1). This structuring and verification process is a highly manual and time-consuming process. Depending on the complexity of the study and current efficiencies within the company, this process typically takes 8-32 hours of defining and programming and 2-3 weeks turnover time for alignment.

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

Future Business Flow



This figure shows the envisioned future DDF of trial-design information from study design to submission.

1. The study team decides on the global design of the study (i.e., indication, objectives, endpoints). The information is stored directly in a digitized format, either in a study operational management solution or in a study design solution. This information is already aligned with the USDM format.
2. If stored in a separate solution, the study design outline is automatically (⚙) exchanged via an API with the study design solution. The design is then further completed by the medical writer resulting in the full study design stored in the study design solution. This study design solution stores the complete design of the study; enables coding for therapeutic areas, interventions, and indications; and it allows direct verification by the medical experts with electronic signatures. This avoids the clinical programmer-medical expert feedback loop which usually happens weeks, months, or even years later. The solution provides the study design information in a study protocol view as well as in user-specific views. The study design solution stores and exchanges the study design information in USDM format.
3. Based on the standard USDM format and company standards, the study design information is automatically (⚙) transformed to the trial design information represented in the SDTM Trial Design domains. The clinical programmer reviews the results and can adjust the output.
4. Quality control is executed manually (👁) on the final digital version and verification confirmed by electronic signatures before the submission package is generated automatically (⚙). This package includes both the Trial Design domains as well as the metadata submitted in the define.xml file.

An alternative digital data flow route via the EDC system (not included in Figure 2) enables the automated addition of the few additional operational parameters needed for the Trial Design domains (e.g., study start and end dates).

This envisioned approach is likely to improve the efficiency in a number of downstream use cases. For the use case described in Figure 1, this approach may minimize the effort of the clinical programmer to 1 or 2 hours for a final check before validation can be performed. Further, it reduces the time for creating and/or validating the define.xml with regard to the Trial Design domains.

Although not specifically addressed in this Implementation Handbook, the same information, logic, and approach for creating SDTM Trial Design domains can be applied to create regulatory registry submissions, for example to the European Medicines Agency Clinical Trials Information System (https://www.ema.europa.eu) or the US ClinicalTrials.gov (https://clinicaltrials.gov), adding to the efficiencies enabled by the digitized design in USDM format.

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

Transitioning and Process Change

As described in the TransCelerate white paper on implementing DDF,[1] the following methods for the digital protocol have surfaced:

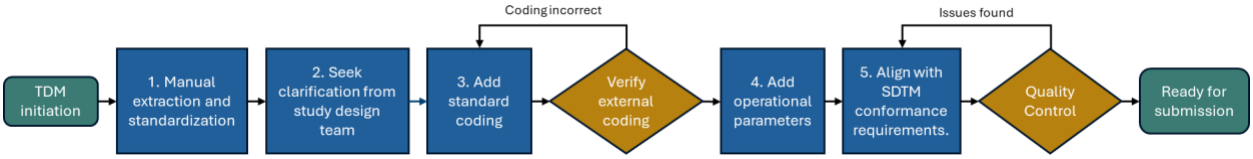
1. the protocol digitization, which enables the conversion of an analog protocol to a digital format; and
2. the protocol digitalization, which is a deeper transformation that harnesses digital technology to create and manage digital protocols from their inception.

Both approaches are included in the concept of a study design solution as described in this Implementation Handbook. The protocol digitization approach is very helpful for legacy studies and for studies which have already started based on the current process, and is likely to be more beneficial for future studies. With the new process, all study design information is recorded in a digital format from the start onwards, requiring a shift in processes and the mindset of the study design team members and medical writers. A hybrid approach might be beneficial for the transition period to allow for testing and familiarization with new and optimal solutions and procedures. As described in the TransCelerate framework for implementation, one size does not fit all; individual use cases and user needs and requirements will determine the best solution selected or developed based on the common future-state vision. Achieving agreement on a suggested solution by involving stakeholders in this process is crucial. Finally, the new approach described here will also have an effect on the company and organizational standard operating procedures involved in this process. This needs to be addressed in the change process as well.

2.2 User Level

Current Process

From a user and a functional perspective, the creation of Trial Design domains, initiated by the clinical programmer, follows a multistep process to prepare the TDM package for submission.



First, the clinical programmer reviews the protocol and extracts and standardizes information, including

- Titles, identifiers, indication, and general study characteristics
- Objectives, endpoints, and estimands
- Intervention details
- Study planning scheme (periods, arms, branching)
- Study population characteristics
- Eligibility criteria
- Visit/encounter information

This study design information is transformed to a standardized spreadsheet (1). This process may or may not include some automated features for replacing special symbols, adding additional required information, and combining information according to SDTM conformance requirements, regulatory requirements, and internal standard procedures. This manual interpretation of protocol documents is inherently prone to inefficiency.

The programmer seeks clarification from the design team as needed to represent the protocol information correctly (2). Critical design details are often scattered or undocumented. Since this process typically occurs late in the study life cycle, it creates a disconnected workflow that relies on memory rather than a structured "single source of truth."

The programmer adds standard coding (3) aligning to CDISC Controlled Terminology and consults with the assigned medical expert to verify non-CDISC codes for the study indication, therapeutic area, and interventions. The

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

programmer also adds additional parameters (4) from the study execution database (e.g., start and end date of the study).

Alignment with SDTM conformance requirements (5) is tested and content is validated during the quality-control phase by a second programmer or expert. Identified issues need to be resolved.

The Trial Design domains are added to the SDTM submission package. Corresponding metadata needs to be added manually to the input for the define.xml file which includes all the metadata of SDTM domains submitted to regulatory agencies.

Future Process

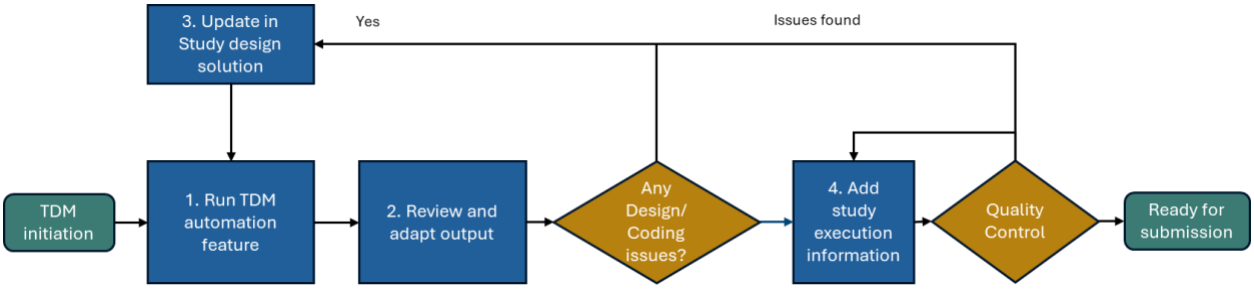
The User Requirements section of this Implementation Handbook describes what features are necessary for the study design solution (including information to be stored) to automatically run the creation of Trial Design domains.

In this solution, study design information is entered by the medical writer, including:

- Defining subject pathways by arm, preferably using visual tooling which generates pathway pictures to be included in the protocol
- Defining the visits and the logical ordering
- Selecting eligibility criteria from a template library with automatic linkage to study-specific features and the ability to add study-specific criteria
- Defining the study population (and, optionally, corresponding cohorts)
- Selecting objectives and endpoints from a template library with automatic linkage to study-specific features and the ability to add study-specific objectives and endpoints
- Defining study-specific design elements (e.g., title, identifier, indication, phase, trial type, blinding structure)
- Specifying intervention details (e.g., dose, frequency, administration route)

Coding of standardized elements is automatically added in the background. This is based on the SDTM and external codelists that are required for submission. Coding for indications and pharmacological class can directly be defined and verified by the medical expert in the study design phase. Moreover, if stored in a metadata repository, the design can be reused for other studies that are performed for the same indication or intervention.

The study design solution typically makes the study design available via the clinical study protocol as an official document representing the study design, and allows for creating different study versions by means of study amendments.



In the future process, the creation of the trial design is still initiated by the clinical programmer. However, instead of having to thoroughly go through the full protocol, the programmer can just run the tooling (1) and review the output (2). If the output is not yet complete, the programmer can verify with the protocol and adapt the information that is still missing. This step keeps the human in the loop. If this review reveals design or coding issues, then the corresponding information can be updated in the study design solution (3), ensuring consistency and one source of truth for all design information. At the final stages of the study, the last study execution parameters can be added manually (4) or by an automation feature which directly links to the EDC output (not covered in this Implementation

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

Handbook). After a final quality control, the information is ready for submission—including both the actual Trial Design domains and the corresponding metadata that is exchanged with the regulatory agencies via the define.xml file.

By this direct link to the study design information stored in the study design solution, any study design update affecting the Trial Design domains will be automatically propagated into the Trial Design domains.

2.3 Technical level

Current Process

Currently, most companies creating TDM use small tools to create SDTM-compliant content for the Trial Design domains. This may include the following functionalities:

- Transform text to ASCII and replacing symbols
- Align TA and TE domain content
- Check specified codes against CDISC Controlled Terminology
- Split text after 200 characters
- Check minimum set of required parameters are included in the TS domain

However, these features are usually stand-alone macros or other functionalities based on manual input from the clinical programmer, retrieved from a PDF of the clinical study protocol.

Future Process

All study design information is entered directly in a study design solution by the study design team, which includes the medical writer. This study design information can be represented as a clinical study protocol, and it includes the information that is relevant for the creation of the Trial Design domains.

A trial design automation feature is either included in the study design solution or in another downstream solution that can ingest the study design in USDM format. The feature is being used to automatically create the Trial Design domains based on the digitized USDM-formatted study design information. This feature includes the following functionalities:

- Automatically create the Trial Design domains based on the USDM output
- Automatically check completeness and conformance with standards and regulatory requirements
- Review functionality for the clinical programmer and quality control person
- Option to change or add information when needed in a controlled environment. This includes a method to track and review all changes performed, including the reason why changes were performed. If design details from the protocol are adapted, then the rationale for this is also verified during quality control.
- Option to ingest new versions of the study in USDM format without losing the additions and validation of information that were not changed
- Enable company and therapeutic area-specific formatting requirements for visit naming, visit numbering, and parameter grouping
- Option to ingest and/or add manually the few study execution-related elements

See the User Requirements and Technical Implementation sections for additional details and discussion.

Note that there is an overlap of information needed for the trial design domains and information needed for other downstream processes. For example, the same information is also used for registry submissions. A similar approach as for creating the Trial Design domains can be used to create such submissions, which increases the process efficiencies.

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

2.4 Maturity Matrix

In the process of transforming to a full digital data flow, some study design solutions may not yet cover all information or features that are needed to automatically create Trial Design domains. Simple trial design information pieces (e.g., study phase, population characteristics) are likely to be covered first; more complex concepts (e.g., epochs, branching) are included in more sophisticated solutions. The following table indicates the maturity of a solution, from both the upstream study design solution perspective and the downstream TDM automation feature perspective.

	Feature implementation	Low	Low-medium	Medium	Medium-high	High
Study design solution	Low-level features as described in Upstream Study Design Solution Implementation Requirements	Partly covered	Fully covered	Fully covered	Fully covered	Fully covered
	Medium-level features as described in Upstream Study Design Solution Implementation Requirements		Partly covered	Fully covered	Fully covered	Fully covered
	High-level features as described in Upstream Study Design Solution Implementation Requirements				Partly covered	Fully covered
TDM automation feature	Creation of TV	Direct mapping	Direct mapping	+ full numbering features	+ full numbering features	+ design comparison for potential arm allocation
	Creation of TI	Direct mapping	Direct mapping	+ version information	+ version information	+ subcategory information
	Creation of TE	Direct mapping	Direct mapping	Direct mapping	Direct mapping	+ duration information
	Creation of TA		Construction and mapping of study pathways based on arms, elements and epochs	+ full numbering features	+ branching information	+ transition rule information
	Creation of TS	Direct mappings	+ Boolean and ISO variables	+ version information	+ arrays and groupings	+ version comparison and user evaluation
	All domains					+ full validation functionality including formal verification.

For the **upstream study design solution**, the maturity level indications in the table can be interpreted as follows:

- Low: Started with the digitized collection of information needed for Trial Design domains.
- Medium: Most study information for Trial Design domains is captured within the solution, with the exception of some complex concepts that need to be added manually later.
- High: All study design information for Trial Design domains is captured within the solution and additional features are included to enable easy configuration of pathways and tagging of criteria, objectives, and endpoints.

For the **downstream TDM automation feature,** the maturity level indications in the table can be interpreted as follows:

- Low: Basic direct mapping from information stored in USDM to TDM
- Medium: Additional more complex derivation features included (e.g., version comparison, numbering)
- High: Full mapping and complex logical derivations included as well as full validation and track-changes functionality

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

3 User Requirements

This section includes the user requirements for both the upstream study design solution and downstream TDM automation. These requirements can be used as a starting point to either build a solution or to select a commercially available study design solution. Although this list of requirements is focused on the TDM build, it can also be used as a starting point for the full requirements of the study design solution, as many aspects are applicable to other downstream use cases.

3.1 Relevant Stakeholders

Role	Responsibility	Viewpoint	High-level interests
Management	Improve efficiency in processes	Business-level processes	Process flow from a business perspective High-level regulatory needs Key performance indicators
Clinical programmer	Create Trial Design domains	User-level data flow	Regulatory TDM requirements Required controlled terminology Automated creation of TDM
Quality control	Validate quality of Trial Design domain information	Requirements and output	Regulatory TDM requirements Required controlled terminology Easy validation and sign off
System developer	Create tooling for automation	Modeling and development	Data structures Input/output requirements Technical implementation details Data lineage
Medical writer	Enter or link relevant study design information	Study design documentation	Correct and consistent representation of study designs Seamless study design features Flexibility in writing
Medical expert	Review study design information and add/validate coding for indications and medications	Study design documentation	Study design review options Study protocol representation

3.2 Upstream Requirements for the Study Design Solution

Study Identifier

The study identifier used for submission needs to be collected in the tool, stored in the USDM, and correctly linked to the sponsor organization. For the exact requirements and an example instance diagram for storing the study identifier, see the USDM Implementation Guide (https://github.com/cdisc-org/DDF-RA).

Eligibility Criteria

As a minimum, the solution needs to include a means to specify criteria or select criteria texts from a library, link them to the corresponding study design, and indicate the ordering and whether they are inclusion or exclusion criteria. The following table summarizes information needed for this.

Variable	System requirement	USDM location	SDTM TI variable
Incl/Excl Criterion Short Name	Store test code/short name for each selected criterion. Enable definition per study. Often numbering is included in here.	EligibilityCriterion/@identifier	IETESTCD
Incl/Excl Criterion	Select criterion from list and/or enable writing of new criteria, enable adjustment of criterion text	EligibilityCriterionItem/@text	IETEST
Incl/Excl Criterion Category	Require for each criterion the indication whether it is an inclusion or exclusion criterion. Standardize according to SDTM codelist C66797.	EligibilityCriterion/@category	IECAT
Incl/Excl Criterion SubCategory	Optionally enable an indication of a subcategory. This can be done according to an internal codelist.	EligibilityCriterion/@notes	IESCAT

Eligibility criteria may change due to study amendments. If subjects were already enrolled when a previous version of the study was effective, then it needs to be clear what criteria were effective when these subjects were enrolled. As a result, the corresponding version identifier needs to be stored in the USDM.

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

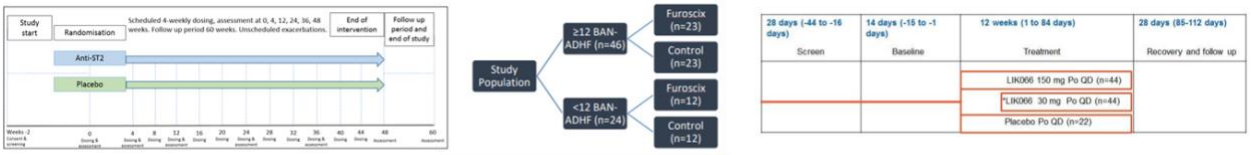
Variable	System Requirement	USDM Location	SDTM TI variable
Protocol Criteria Versions	Enable the storage of different versions of criteria and store the corresponding version identifier. Enable export or utilization of different versions (not only the last one) in USDM format.	StudyVersion/@versionIdentifier	TIVERS

Based on this requirement, when multiple versions in USDM are applicable and available, the trial design tooling can compare the versions and automatically indicate differences between them including for what version(s) specific criteria were effective.

More advanced study design solutions may provide tagging and digitization functionality for eligibility criteria. This could be valuable in linking criteria to screening assessments specified in the schedule of activities and/or for automatic selection of potential eligible subjects from downstream sources. Although this is not needed for the creation of the TDM, when it is included then the downstream TDM automation feature needs to be able to process the tagged information.

Study Pathway/Experimental Design Utility

Study pathways are in many protocols presented as schematic figures showing the different paths that subjects in different arms of a study can take, such as:



Via search on clinicaltrials.gov: NCT03615040, NCT05528588, and NCT03205150

These pathways are the basis of the arms, epochs, and elements structure in the SDTM and the USDM. A solution needs to provide a way to link the arms to different phases (epochs) of the study and to indicate what happens (elements) for each arm in each of the phases. (See SDTMIG Section 7.2, Experimental Design (TA and TE), for requirements for defining arms, epochs, and elements (https://www.cdisc.org/standards/foundational/sdtmig); for how this information is stored in USDM, see Chapter 2 of the Understanding USDM training.)

Visual tooling for creating and presenting these pathways, included in more advanced study design solutions, will aid both the understanding of the pathways as well as the presentation in protocols and related documentation. Creating and visualizing the pathways is even more challenging for complex studies like the modular, platform, or basket trial approaches that are frequently applied for current oncology studies. However, these visualizations are very helpful for understanding these complex trial designs. In addition, apart from being able to automatically create the TDM and registry submissions, the digitalization of this information aids the automatic creation of subviews targeted to a specific audience.

Encounters/Visits Utility

The tool must be able to create a list of encounters which can be used for the creation of the TV domain. Encounters may be physical visits but could also be virtual encounters or inquiries. As a minimum the tooling must be able to specify the following information for each visit:

Variable	User requirement
Visit name	Store the visit name as it is presented in the schedule of activities
Planned study day of visit	Store the planned study day of the visit in the corresponding timing included in the corresponding schedule of activities
Visit start rule	A description of the start of the visit. This can be a standard text according to company procedures linked to the standard visit name. However, it should be adjustable for specific study needs.
Visit end rule	A description of the end of the visit. This can be a standard text according to company procedures linked to the standard visit name. However, it should be adjustable for specific study needs.

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

In some cases, the visits and corresponding descriptions may differ for different arms of the study. In that case, the corresponding schedule of activities might be considerably different as well and then it is feasible to store the arm as a separate study design. The visits/encounters are then defined separately for each arm.

Objectives/Endpoints Utility

Each study design includes a list of objectives and endpoints. As a minimum, the solution should enable the specification of a list of distinctive objectives and related endpoints including the level of the objectives. Optionally, the tool can provide the tagging and digitization functionality for the objectives and endpoints ensuring consistency within the design and enabling initial set-up of the schedule of activities.

Study Interventions Utility

The minimum requirement is to specify all relevant study interventions including the investigational therapy, the comparative treatment, and current therapy if applicable. For each of these treatments the following needs to be specified:

- Intervention type
- Administration frequency
- Administration dose and units
- Administration route
- Drug pharmacological class
- Drug dose form
- Other study intervention features (e.g., minimum response duration, administration duration)

Note that standard SDTM and DDF vocabularies are applicable to a number of these items (see CDISC Controlled Terminology for SDTM, https://evs.nci.nih.gov/ftp1/CDISC/SDTM/, and for DDF, https://evs.nci.nih.gov/ftp1/CDISC/DDF/). Storing the corresponding standard terminology information (e.g. code, decode, code system, code system version) in the background enables automated generation of the TS domain.

Population and Study Indication

The minimum requirement is to specify the study population and indication of the study. For the population, the age range, gender, and planned number of subjects to be enrolled are required; an indicator whether the population includes healthy subjects also needs to be present. For more complex studies, it is advisable to enable the generation of cohorts. Cohorts saved in USDM have the same features as the study population and are nested within the total population. Gender and age ranges are required for at least the whole study design population and/or all the study cohorts.

One or more study indications need to be specified for a study, and for each of these indications it needs to be indicated whether it is a rare disease. The corresponding coding is required for the Trial Design domains and needs to be verified by a medical expert. This verification needs to be an official dated sign-off with a computerized signature.

For submissions to the EMA, a MedDRA code is expected (see https://www.meddra.org/). In other cases, SNOMED controlled terminology or other dictionaries for standardizing conditions may be needed. Note that there should be 1 instance for each indication while optionally more codes from multiple coding systems are allowed for each indication.

General Study Design Features

In addition to the objectives/endpoints, interventions, and population requirements, a number of general design features are required to be stored to enable the generation of the trial design domains. These include but may not be limited to:

- Sponsor name
- Study title and acronym

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

- Study phase
- Study type
- Trial scope
- Study design model
- Trial intent type
- Whether the study is randomized
- Study blinding indicator
- Whether the study is an adaptive design
- Whether the study is an extension trial
- Intervention model
- Control type
- Therapeutic area
- Other study identifiers (e.g., registry number, submission number)

A number of these items should be coded in the background according to the corresponding SDTM terminology in order to enable automated creation of the Trial Design domains. In addition, the coding of the therapeutic area needs to be verified by the medical expert.

3.3 Downstream Requirements for the TDM Automation Feature

The TDM automation feature should automatically generate the Trial Design domains based on the information entered and stored in the study design solution in USDM format and the corresponding mappings from USDM to TDM. The additional requirements noted in the following table will make the output more accurate and efficient. These requirements can be tailored to the company's or sponsor's specific needs.

Requirement	Description
Numbering formats	The tool will automatically generate the numbering of epochs within arms (TAEORD), and the visits (VISITNUM) based on the previous and next attributes stored in the corresponding USDM classes. In addition, especially for visit numbers, company- and/or therapeutic area-specific formats might be applicable (e.g. 00100, 00101, 00102) and may include study epoch and encounter information in the numbering logic as well. The user should be able to add or adjust this sponsor-specific numbering format. For the Trial Summary (TS) domains, sequence numbering is also applicable (TSSEQ) which is based on the instance ordering in the USDM. For this variable, company-specific formats might be applicable as well.
Required parameters and null-flavor indicators	Some TS parameters are required. These requirements may differ based on the regulatory body to which data are submitted, the study type, and the study design. Users should be able to indicate the parameters that are required for a TS domain. This may vary per use case. Users should also be able to indicate what null flavor should be added, in case of missing values. The corresponding null-flavor coding will then be automatically added to the output in case of missing values.
Naming of grouped items	Multiple features of a concept in the TS domain need to be grouped with a common id (TSGRPID). These can be based on the corresponding instance and nesting within the USDM API file, but the user needs to be able to apply the company-specific naming and formatting to the identified groups.
Eligibility criteria versioning	With amendments to a study design, the criteria may vary. Therefore, it is required to ingest all the different official versions of the study design, compare them, and create a list of unique criteria per version. An additional requirement may be that when the study is completed, it is checked whether all the different versions were applicable to at least 1 study subject. If not, then the corresponding version information may be omitted from the output.
Resolve tags	In case digitization of eligibility criteria, if objectives and/or endpoints are applied using the USDM syntax template feature, then the tags need to be resolved and the HTML formatting needs to be removed.
ASCII requirements	Regulatory bodies may require the information to be submitted in ASCII format. In that case special symbols need to be replaced by simple ASCII-formatted text representations. The user must be able to indicate special symbols and their replacement values in the settings of the solution.
Eligibility criteria text length	After resolution of tags and replacement of special symbols, per SDTMIG criteria texts which are longer than 200 characters need to be shortened manually for submission in the TI domain. The complete version of the criteria needs to be included as metadata in the define.xml file. AI functionality could be beneficial to suggest the shorter eligibility text.

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

Requirement	Description
Trial summary parameter text length	If the text value for trial summary parameters is longer than 200 characters, per SDTMIG the string needs to be split into multiple strings, each with a maximum of 200 characters. Note that the string may not be split in the middle of a word. So, the last space before the 200-character length is the splitting point.
Study execution parameters	There are additional study execution parameters that are needed for the TS domain (e.g., actual study start date, study end date). This could be either required as a manual addition or based on information retrieved from the study database including the collected data. However, some of these additional parameters are operational (e.g., versions of relevant implementation guides); such information is more likely to be added manually.
Audit trail and data lineage	For flexibility reasons, the clinical programmer must be able to add and/or alter information that is automatically created. However, if information entered in the study design solution is changed, this should be either (1) flagged and a rationale be given, or (2) required to be changed in the upstream study design solution. All changes performed by the programmer must be tracked and dated to enable the full data lineage from design to submission.
Review functionality	The clinical programmer and quality control person should be able to review and verify the content of automatically generated TDM. The verified information should stay verified if it is not changed, even if a new version of the study is uploaded in USDM format.

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

4 Technical Implementation

4.1 Upstream Study Design Solution Implementation Requirements

The upstream study design solution must be at least capable of storing and mapping the study design information that is required for the downstream TDM automation feature. The following table further specifies this from a technical perspective; although it includes the main requirements for the Trial Design domains, it is not exhaustive. Requirements can differ between submissions and new parameters may be added when needed. For each requirement, the implementation level is indicated which indicates the difficulty level for implementation. These levels can be interpreted as follows:

- Required information with **low complexity** is stored in USDM format as single attributes which can be directly mapped to the TDM domains. This requires the following skills:
  - Basic understanding of CDISC USDM classes and SDTM trial design domains
  - Ability to perform 1:1 data mapping between source system fields and USDM attributes
  - Familiarity with using controlled terminology dictionaries (e.g., CDISC, SNOMED, MED-RT)
- Required information with **medium complexity** is stored in multiple classes in USDM with relational logic, cross-references and applied content logic. This requires the following skills:
  - Extended understanding of USDM classes and SDTM Trial Design domain content
  - Understanding of relational data architecture structures and how to handle cross-referencing between classes (e.g., linking the EligibilityCriterion class to the StudyDesignPopulation class)
  - Ability to implement content logic that governs how different data elements interact across the study version
- **High complexity** information requires advanced automation and validation, such as
  - Understanding of Schedule of Activity timeline USDM features
  - Ability to implement metadata repository features for reuse and consistency within and between studies
  - Rule engineering expertise, including in implementing USDM Core Rules for real-time validation and feedback loops
  - Advanced semantic modeling ability to implement complex syntax templates and tags for automated validation (e.g., linking assessments to biomedical concepts)

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

Functional requirement	Relevant SDTM domain	Relevant USDM attributes	USDM nested in	Implementation notes	Implementation level		
					Low	Medium	High
Store the study identifier in the scope of the sponsor organization	all	StudyIdentifier/@text	study version	Ensure that the sponsor study identifier is stored with a scope of the sponsor organization. See the USDM-IG for examples (https://github.com/cdisc-org/DDF-RA/).		X	
Store criterion category and identifier	TI	EligibilityCriterion/@identifier EligibilityCriterion/@category StudyDesignPopulation/@criteria	study design	The EligibilityCriterion class includes the attributes for indicating the category (inclusion/exclusion) and identifier (short name) of the criterion. This information is nested within the corresponding study design and is referenced to from out of the study design population (StudyDesignPopulation class) or 1 or more cohorts, if they are defined. Note that renaming the identifier (short name) of criteria for new versions of a study is not recommended. Because of lineage and traceability reasons it is recommended to keep the combination of criterion identifier (short name) and text unique and add sub numbering to the identifier in case of changes due to amendments. See the SDTM-IG for TI assumptions. The ordering in USDM for protocol presentation reasons should be represented by the EligibilityCriterion previous and next attributes.	X		
Store criterion item text as a separate repository.=	TI	EligibilityCriterionItem/@text	study version	The actual text items of the criteria are stored at a higher level in the USDM structure to allow for reuse across designs. They are referred to from the EligibilityCriterion class in the applicable study design.	X		
Refer to the criteria from cohorts	TI	StudyCohort/@criteria	study design population	Enable the definition of cohorts and enable the cross-reference from a cohort to the cohort specific criteria instead of from the complete study design population when applicable.		X	
Enable tagging of assessments in the criteria: to cross-refer to corresponding screening assessments as biomedical concepts (This will allow automated validation and set-up of the schedule of activities), and/or to set up a criterion repository with templates that are automatically adapted according to the study specific details	TI	SyntaxTemplate/@dictionary ParameterMap/@tag	study version	The implementation of syntax templates is described in the USDM-IG. Tags could, for example, link the age criterion to the corresponding age range stored in the StudyDesignPopulation class, automatically linking both together and ensuring consistency. Note: All tags need to be defined in the ParameterMap class in order for the downstream functionality to work correctly.			X
Enable export of previous versions	TI	StudyVersion		Eligibility criteria are one of the most likely aspects of a design that changes by amendments across versions. For the submission,		X	

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

Functional requirement	Relevant SDTM domain	Relevant USDM attributes	USDM nested in	Implementation notes	Implementation level		
					Low	Medium	High
of the study in USDM format				different versions need to be compared in order to get the right output.			
Store the distinct activities that are defined for a specific period and for a specific arm (e.g., SCREENING, TREATMENT X MG, TREATMENT Y MG, FOLLOW-UP)	TE/TA	StudyElement/@label StudyElement/@description	study design	These are building blocks for building the specific pathway for a specific arm.	X		
Specify the transition start and end rules for study elements	TE	StudyElement/@transitionStartRule StudyElement/@transitionEndRule	study design	The transition rules need to be defined according to the TransitionRule class. The corresponding text attribute will hold the actual transition rule text.		X	
Specify the planned duration of an element	TE	StudyElement/@studyInterventions Administration/@duration	study design	The duration is not (yet) included in USDM v4.0 on an element level. However, for the treatment elements, it can be derived from the study intervention administration duration when the corresponding treatment is referred to from out of the element. On an element level it can be added as an USDM extension with the same duration complex datatype that is used form administration duration.			X
Store the distinctive study periods as epochs	TA	StudyEpoch/@label StudyEpoch/@description	study design	These are building blocks for building the specific pathway for a specific arm.	X		
Store the distinctive arms	TA	StudyArm/@label StudyArm/@description	study design	These are building blocks for building the specific pathway for a specific arm.	X		
Create the study pathways by relating the arms, periods (epochs), and elements to each other	TA	StudyCell/@arm StudyCell/@epoch StudyCell@elements	study design	Combining the building blocks enables representation of the study pathways.		X	
Visually show and enable configuration of study pathways	TA	StudyCell/@arm StudyCell/@epoch StudyCell@elements	study design	Enable users to easily configure patient pathways and use of visualizations for presentation purposes in protocols and related documents.			X
Include transition information as decision instances in the schedule of activities	TA	ScheduledDecisionInstance/ @conditionAssignments ScheduledDecisionInstance/@epoch ScheduledActivityInstance/@epoch	schedule timeline	In cases where a transition is described in the study design and included in the timeline logic, the transition information can be automatically derived. For example, the decision to go to end-of study/follow-up when a patient is no longer fit to continue can be modeled as such in USDM. See the USDM-IG and Understanding USDM training (https://learnstore.cdisc.org/product?catalog=Understanding-USDM) for more information on modeling transitions and alternative routes using the ScheduledDecisionInstance class.			X
Specify study visits/encounters	TV	Encounter/@label	study design	Enable users to specify the visit/encounter information as presented in the schedule of activities.	X		

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

Functional requirement	Relevant SDTM domain	Relevant USDM attributes	USDM nested in	Implementation notes	Implementation level		
					Low	Medium	High
Enable ordering of visits/encounters	TV	Encounter/@previous Encounter/@next	study design	Enable users to specify the order of the visits as presented in the scheduled of activities.		X	
Specify the timing of visits/encounters	TV	Encounter/@scheduledAt Timing/@valueLabel	study design / schedule timeline	The timing should be related to the schedule of activities. See the USDM-IG and Understanding USDM training regarding how this can be defined.		X	
Specify the transition start and end rules of visits/encounters	TV	Encounters/@transitionStartRule Encounters/@transitionEndRule	study design	The transition rules need to be defined according to the TransitionRule class. The corresponding text attribute will hold the actual transition rule text.		X	
Enable design variations for different arms	TV/TA		study design	In cases where arms follow a clearly different pathway and/or timeline, storing them as separate designs in USDM enables clear distinction and presentation of the different pathways. All common information may be shared in the solution but needs to be exchanged separately by design in the USDM format.			X
Specify study titles and acronyms	TS	StudyTitle/@text StudyTitle/@type	study version	Use DDF controlled terminology (https://evs.nci.nih.gov/ftp1/CDISC/DDF/) to correctly indicate the acronym and official study title.	X		
Specify distinct objectives and endpoints including their level	TS	Objectives/@text Objectives/@level Endpoints/@text	study design	Define the objectives level according to the DDF controlled terminology and store the corresponding endpoints as nested in these objectives.	X		
Specify non-sponsor study identifiers	TS	StudyIdentifer/@text	study version	Enable the specification of all related identifiers for a study within the scope of their corresponding organization (e.g., NCT number, CTIS number).		X	
Specify general study design characteristics	TS	StudyDesign/@characteristics StudyDesign/@therapeuticAreas StudyDesign/@studyType StudyDesign/@studyPhase StudyDesign/@model StudyDesign/@subTypes StudyDesign/@blindingSchema StudyDesign/@intentTypes	study design	Use CDISC definitions and controlled terminology as specified in the DDF controlled terminology list.	X		
Specify observational study design characteristics	TS	StudyDesign/@model StudyDesign/@subTypes StudyDesign/@timePerspective StudyDesign/@samplingMethod	study design	Use CDISC definitions and controlled terminology as specified in the SDTM controlled terminology list. Note that the codelists for model and subTypes differ between interventional and observational designs.		X	
Specify study intervention details	TS	StudyIntervention/@role StudyIntervention/@label StudyIntervention/@minimumResponseDuration	study version	The role should be coded according to DDF controlled terminology and is critical for distinguishing the values for the different related TS parameters. The interventions are referred to out of the study design and can be reused across designs.	X		
Specify study intervention administration details	TS	Administration/@dose Administration/@route Administration/@frequency Administration/@duration	study intervention	Use CDISC definitions and controlled terminology as specified in the SDTM controlled terminology list.		X	
Specify administrable product details	TS	AdministrableProduct/@administrableDoseForm AdministrableProduct/@pharmacologicClass	study intervention	Use CDISC definitions and controlled terminology as specified in the SDTM and DDF controlled terminology lists for dose form and pharmacologic class respectively.		X	

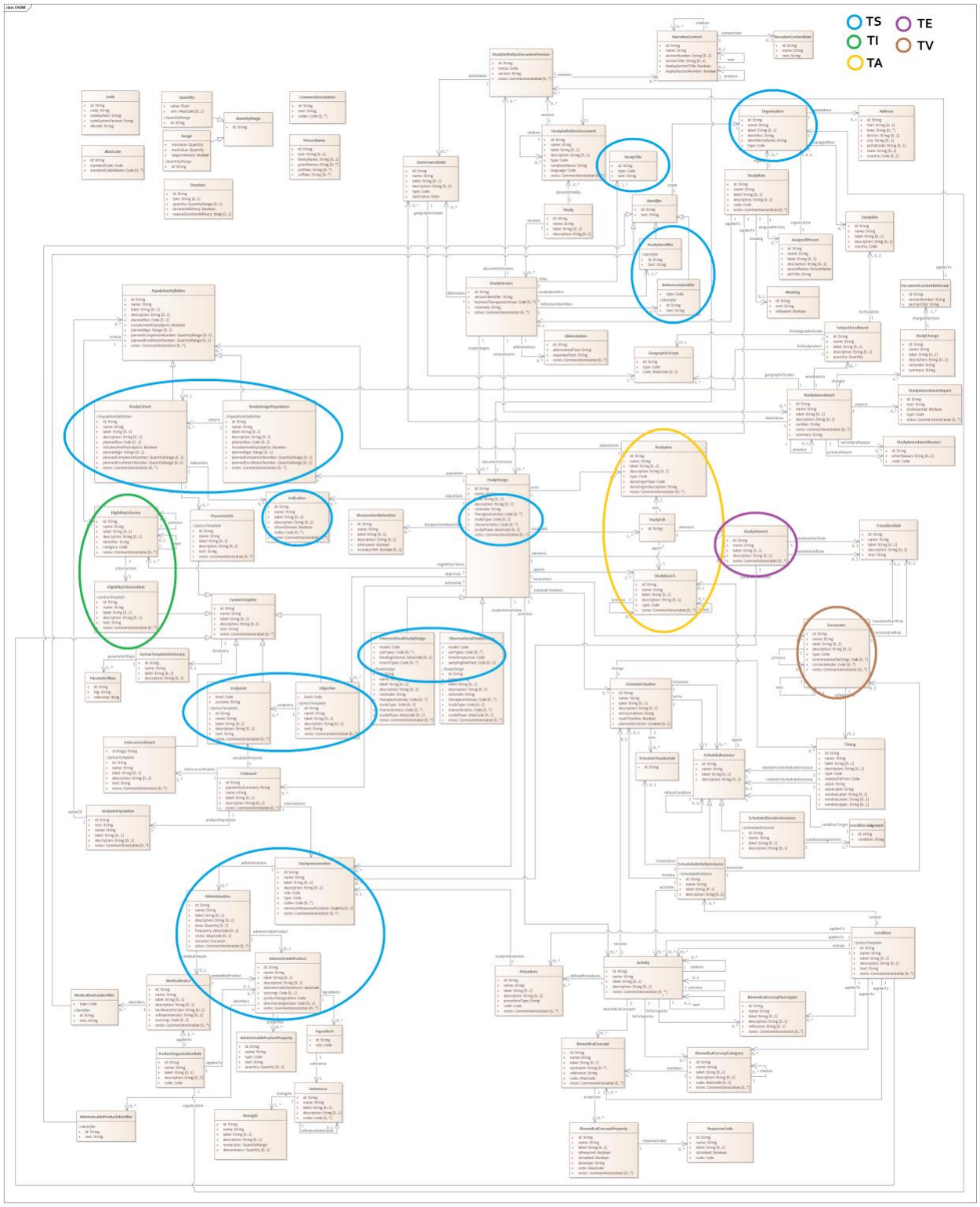
© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

Functional requirement	Relevant SDTM domain	Relevant USDM attributes	USDM nested in	Implementation notes	Implementation level		
					Low	Medium	High
Enable coding and reuse of products administrable product pharmacological class	TS	AdministrableProduct/@pharmacologicClass	study intervention	The coding for pharmacologicCLass does not follow CDISC terminology and needs to be validated or verified by a content knowledge expert.			X
Enable the specification of the study indication	TS	Indication/@label Indication/@isRareDisease	study design		X		
Enable coding of study-specific indication(s)	TS	Indication/@codes	study design	The preferred coding dictionary to use depends on company preference. The coding needs to be validated or verified by a content knowledge expert.			X
Enable specification of the study design population	TS	StudyDesignPopulation/@plannedSex StudyDesignPopulation/@includesHealthySubjects StudyDesignPopulation/@plannedAge StudyDesignPopulation/@plannedEnrollmentNumber	study design	Use CDISC definitions for plannedSex as specified in the the SDTM controlled terminology list. Note that in USDM both sexes are defined as separate codes rather than using BOTH as a coded value.	X		
Enable specification of the study cohorts	TS	StudyCohort/@plannedSex StudyCohort/@includesHealthySubjects StudyCohort/@plannedAge StudyCohort/@plannedEnrollmentNumber	study design	Use CDISC definitions for plannedSex as specified in the the SDTM controlled terminology list. Note that in USDM both sexes are defined as separate codes rather than using BOTH as a coded value.		X	
Validate the content using the applicable USDM Core rules as specified in the Validation section	all			Enable the run of the applicable core rules and provide feedback if they are not met in the form of warnings and/or suggestions.			X

4.2 Mapping from USDM to TDM

The USDM is officially represented in the USDM UML (https://github.com/cdisc-org/DDF-RA/USDM_UML.png), which gives an overview of classes and relationships included in the model. This is further explained in the introduction to the Understanding USDM training (https://Understanding-USDM). The study design information needed for creation of the TDM is scattered over the USDM data model as presented in following figure provide a UML overview. The information for the TV domain is mainly stored in the USDM Encounter class; the information for the TI domain is stored in the USDM EligibilityCriterion and EligibilityCriterionItem class; the information for the TA domain is mainly stored in the USDM StudyArm and StudyEpoch class; and the information for the TE domain is mainly stored in the StudyElement class. The TS domain includes a variety of information pieces located in the studyTitle, Organization, StudyIdentifier, StudyDesign, StudyDesignPopulation, Objectives, Endpoints, StudyIntervention, and Administration classes.

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved



The latest version of the mapping file for USDM-to-TDM variables (https://github.com/cdisc-org/sdtm_mapping.xlsx) indicates the class and attribute of each of the relevant items and also includes selection conditions and notes when applicable. However, to actually implement these mappings, a transformation is needed from the mappings in the mapping file to the actual programming logic.

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

For example, the mapping in the mapping sheet for the investigational therapy or treatment is:

```
StudyVersion/@studyDesigns/StudyDesign/@studyInterventions/StudyIntervention/@label
```

with the corresponding condition:

```
StudyIntervention/@role /code/@code = "C41161"
```

These mappings need to be translated to actionable code. This can be done, for example in JSONata, which is a lightweight query and transformation language for JSON data and can be used to query the USDM API file (which is in JSON format). JSONata queries can be run using a Python JSONata package. Translating the preceding statement into actual JSONata code, the path and condition from the mapping sheet are integrated in 1 statement:

```
study.versions.studyInterventions[role.code="C41161"].label
```

Note that a JSON path is only needed for the retrieval of information. Therefore, the class names are not included in this query.

Another option is to use the standard Python JSON package to retrieve the information from the indicated classes and attributes in the mapping. For example, the following code can be used to get the indication if healthy subjects are included in the population:

```
pop = study_design.get("population", {})
val = pop.get("includesHealthySubjects")
```

However, this approach might become quite complex and intensive in the case of complex trial summary parameters (TS).

See the Examples and Use Cases section for a discussion of 2 open-source tools for creating the Trial Design domains; both of these approaches are demonstrated there.

Getting the Correct Study Identifier

An important variable for all Trial Design domains is the study identifier. The mapping of the sponsor's study identifier needs a complex query; the study may have multiple identifiers scoped within different organizations. For the STUDYID variable, select the one that is scoped by the sponsor organization. The storage of this identifier is described in more detail in the USDM-IG (https://github.com/cdisc-org/USDM-IG.pdf).

In the preceding Trial Design mapping file, the study identifier mapping is indicated as:

```
StudyVersion/@studyIdentifiers/StudyIdentifier/@text
```

However, the corresponding condition is complex:

```
StudyRole/@code/Code/@code="C70793" | StudyRole/@organizations
| StudyVersion/@studyIdentifiers/StudyIdentifier/@scope
```

It implies that a study role should exist that is coded as sponsor and which points to the details of the sponsor stored in the Organization class. The corresponding study identifier then also needs to point to the same organization with its scope. In the USDM API file, these are all cross-references, so the reference naming in the API will be adapted to organizationIds and ScopeId. The corresponding JSONata query then would look like this:

```
(study.versions)@$sv.$sv.studyIdentifiers[scopeId in
$sv.roles[code.code="C70793"].organizationIds[0]].text
```

4.3 Downstream TDM Feature Implementation Requirements

As discussed in Mapping from USDM to TDM, the information in the mapping file can be transformed to actual queries and programming code which then can action upon the study design stored in the USDM API file. Specific programming features that are needed to create a fully compliant TDM output are described in this section.

Creation of the TA and TE Domains

The TE domain is the domain that can be directly based on the USDM StudyElement class, as described in the mapping file (https://github.com/cdisc-org/sdtm_mapping.xlsx). The content for the TEDUR variable is, however,

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

less straightforward. If an element is referring to a study intervention with its studyIntervention relationship, then the corresponding administration duration can be used. In all other cases, the timing details in the schedule of activities might work. The timing difference between the visit at the start of the epoch and the visit at the start of the next epoch can then be used to calculate the trial element duration. For this, a fully digitized schedule of activities is needed.

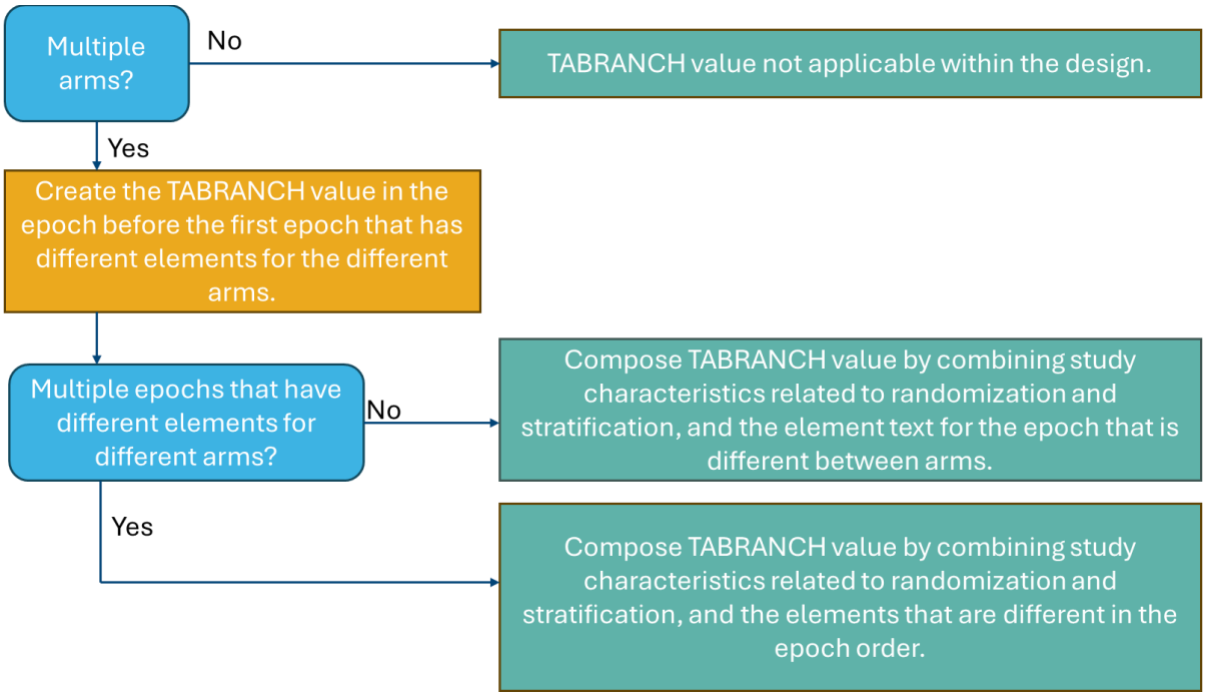
As discussed in the User Requirements section, study pathways are a combination of epochs and study elements allocated per epoch for a specific arm. The TA domain is representing these pathways and is therefore a combination of instances from the USDM classes StudyArm, StudyEpoch, and StudyElement, bound together by the instances in the StudyCell class. For each arm, the following steps need to be automated:

- Identify the arm label and description to map to, respectively, ARMCD and ARM.
- Identify epochs related to the study arm via the StudyCell class and define the order (TAEORD) as described below, based on the next (or previous) attribute.
- Identify the elements related to the arm and corresponding epoch via the StudyCell class to map to, respectively, ETCD and ELEMENT.

The TABRANCH and TATRANS variables can also be derived, as described next.

TABRANCH

The TABRANCH variable in the TA domain is not yet available in the mappings to USDM published for Version 4.0. It depends on the study design implementation whether and how the corresponding branching information is stored in the study design solution. However, the TDM automation feature can automatically evaluate at what epoch the elements are starting to differ between arms. That is the indication that a branch is starting. Usually this is true for the first treatment epoch. Further, in the study design characteristics, it is indicated whether the study is randomized and/or stratified. This is information that is usually included in the TABRANCH description. With this combined information, the TABRANCH value can be suggested to the users. The following decision tree indicates the evaluation for this suggested content of the TABRANCH variable based on information stored in USDM.



© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

The following is an example of how the pre-wording for study-design characteristics can be defined.

Study design characteristics		Branch description
Code	Decode	Starts with:
C46079	Randomized	Randomized to
C147145	Stratified Randomisation	Stratified Randomisation to
C25689	Stratification	Stratified to

The following is an example of a study with the characteristic "Randomized" and multiple treatment phases with different elements.

TAEORD	Epoch	Arm 1 Element	Arm 1 Branch description (TABRANCH)	Arm 2 Element	Arm 2 Branch description (TABRANCH)
1	SCREENING	Screening		Screening	
2	RUN-IN	Run-In	Randomized to Drug X - Placebo regimen	Run-In	Randomized to Placebo - Drug X regimen
3	TREATMENT 1	Drug X		Placebo	
4	TREATMENT 2	Placebo		Drug X	
5	FOLLOW-UP	Follow-up		Follow-up	

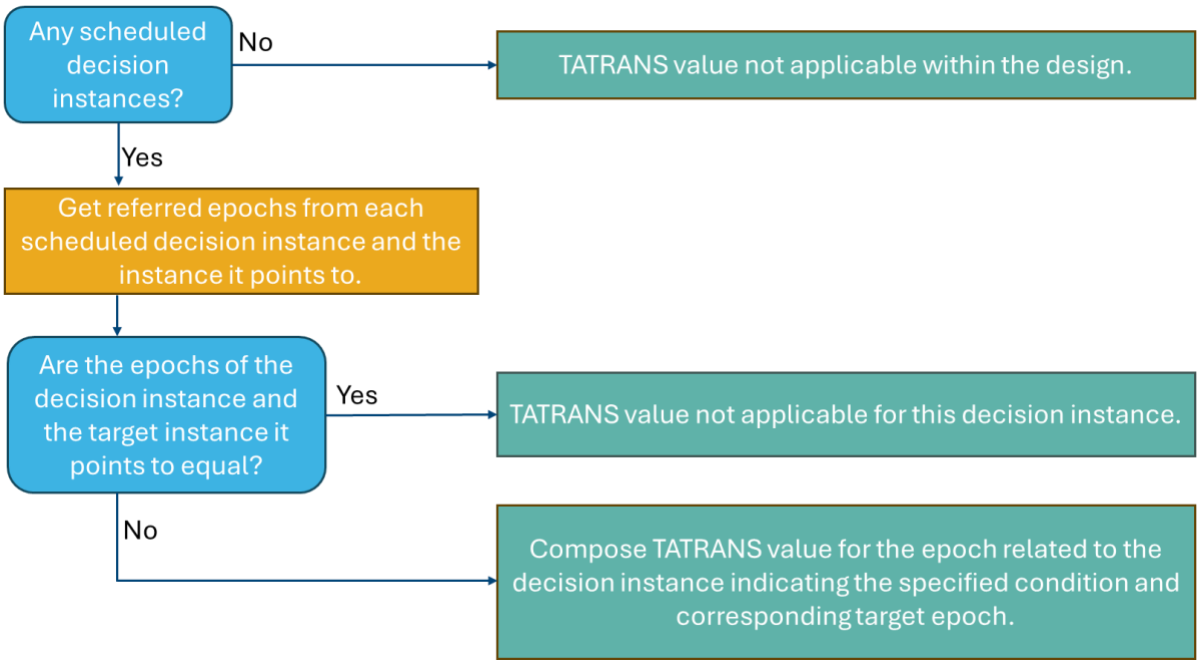
This approach of defining the TABRANCH value is demonstrated in the ta.py code of the ClinLine open-source tool for transforming USDM to Trial Design domains(https://github.com/ClinLine/USDM_SDTM_mapper). See also the Examples and Use Cases section. Note that for more complex studies, branching can occur at multiple phases (epochs) in the study.

TATRANS

The TATRANS value can be based on values stored in the conditionAssignment class when the following requirements are met:

- The conditionAssignment is defined within a scheduled decision instance in the main timeline.
- The scheduledDecisionInstance in which the conditionAssignment is defined points to the corresponding epoch in the TA domain record.
- The conditionAssignment's condition target points to another epoch included for the reported arm.

For automation, the following evaluation needs to be performed in order to define the content for the TATRANS value.



© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

Note that the specified condition and target epoch indicated in the ConditionAssignment class together are forming the text for the TATRANS parameter. This example from the ClinLine tool shows how the results from the following JSONata code —

```
study.versions.studyDesigns.scheduleTimelines.($GetEpochId := function($id) {(instances[id=$id].epochId)};
instances[instanceType="ScheduledDecisionInstance"].
{epochId:
conditionAssignments.condition,"targetEpoch":$GetEpochId(conditionAssignments.conditionTargetId)})
```

is stored in a variable called resultTrans; this is further processed in Python to evaluate whether the scheduled decision instance in the timeline points to another epoch, and creates the final TATRANS text for the epoch that is related to the decision instance based on the condition text and the epoch to which it points. The corresponding Python example code is:

```
resultTrans=definition.Parse_jsonata(codeSnip=TatransCodeSnip,data=data)
definition.string_to_nested_list(resultTrans, resultFromTrans, resultToTrans)  # convert the json string - save epoch/condition and target/epochId as separate strings
for i in range(len(resultFromTrans)):
        fromId, resultFromTrans2 = definition.get_ID(resultFromTrans[i])  # extracting the ID from the string
        toId, resultToTrans2 = definition.get_ID(resultToTrans[i])  # extracting the ID from the string
        if fromId != resultToTrans2: # check if there is actually a transformation to another epoch (epoch Ids from and to are not equal)
            for row in rowIds:          # save in the pre-defined array of output variables for the corresponding epoch
                if row["EpochId"] == fromId:
                    row["tatrans"] =  f"{resultFromTrans2}: transition to " + resultToTrans2  # add the transformation information to the relevant rows
```

Note: The actual text for TATRANS based on USDM content needs to be a suggestion and the user must be able to adjust that according to the preferred output format. Further, this functionality is only possible when a digitized schedule of activities including decision instances for potential transitions between epochs is available.

Creation of the TV Domain

The creation of the TV domain is straightforward based on the information stored in the USDM Encounter class. The numbering of visits based on the USDM previous and next attributes is explained in the following subsection. Per the SDTMIG, ARM and and ARMCD should only be added when the visits are different for different arms. In such cases, the arms are likely to be stored as separate study designs in USDM. Companies can decide to combine different designs in a single submission or create separate submissions for different designs. The user should be able to indicate what designs stored in USDM are to be included for a specific submission.

Creation of the TS domain

The TS domain includes a wide variety of parameters that are either required or not. Per the SDTMIG, missing values are not allowed in the output. From a functional perspective, the user should be able to indicate per submission type what parameters are required, and what parameters should only be included when data is available. From a technical perspective, if parameters are required but no information is available then it should be verified with the user how to handle that. A number of options might be helpful in this case:

- add the information in the study design solution and get a new export,
- enter the corresponding information manually in the study design domain tooling, or
- specify a null flavor value for the specific parameter and confirm the missingness of the value when detected.

TS Array Results

A number of mappings for the TS domain may result in an array of results instead of 1 single value. This may be, for example, a list of multiple indications or a list of objectives and endpoints. Another example is the trial type for which the mapping is:

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

`StudyVersion/@studyDesigns/InterventionalStudyDesign/@subTypes/Code/@decode`

Retrieving this information from USDM using the JSONata query "study.versions.studyDesigns.subTypes.decode" it may result in a list like:

{'Efficacy Study', 'Safety Study', 'Pharmacokinetic Study'}

For the TS domain, all the values in the array need to be stored as separate rows and a counter needs to be added (TSSEQ) to indicate that more items are present for the same parameter, for example:

STUDYID	DOMAIN	TSSEQ	TSGRPID	TSPARMCD	TSPARM	TSVAL
H2Q-MC-LZZT	TS	1		TTYPE	Trial Scope; Trial Type	Efficacy Study
H2Q-MC-LZZT	TS	2		TTYPE	Trial Scope; Trial Type	Safety Study
H2Q-MC-LZZT	TS	3		TTYPE	Trial Scope; Trial Type	Pharmacokinetic Study

The cardinality for the USDM attributes and relationships indicates whether an array result is to be expected. This is presented in the USDM UML diagram. In this case, the subTypes attribute in the InterventionalStudyDesign class has a cardinality of 0..* indicating that zero-to-many values are allowed for the study design subTypes.

TS Grouping

The TSGRPID variable in the TS domain is used to tie together a group of related records. Examples for this are records related to a specific intervention like dose, dosing frequency, and route. All these features are rooted/nested within the corresponding study intervention in USDM. Therefore, the grouping is implicitly available via the USDM structure. The following output shows how the StudyIntervention name attribute is used to group the parameters related to the same intervention together. The tooling can include company standards to add better naming and/or the user needs to be able to adjust these group names.

STUDYID	DOMAIN	TSSEQ	TSGRPID	TSPARMCD	TSPARM	TSVAL
H2Q-MC-LZZT	TS	2	XINONILINE25	CRMDUR	Confirmed Response Minimum Duration	1 Day
H2Q-MC-LZZT	TS	2	XINONILINE25	DOSE	Dose Level; Dose per Administration	54.0,
H2Q-MC-LZZT	TS	2	XINONILINE25	DOSFRQ	Dosing Frequency	QD
H2Q-MC-LZZT	TS	2	XINONILINE25	DOSU	Dose Units	Milligram
H2Q-MC-LZZT	TS	2	XINONILINE25	INTTYPE	Intervention Type	DRUG
H2Q-MC-LZZT	TS	2	XINONILINE25	PTRTDUR	Planned Treatment Duration	24 Week
H2Q-MC-LZZT	TS	2	XINONILINE25	ROUTE	Route of Administration	ORAL
H2Q-MC-LZZT	TS	2	XINONILINE25	TRT	Investigational Therapy or Treatment	Xinomiline

The following table presents the 2 most frequent groupings applied in the TS domain. The grouping principle may also be applicable to other parameters.

Grouping	TS parameter codes (start with)	USDM grouping level	Notes
Objectives/Endpoints	OBJ..., OUTMS...	Objective	
Interventions	TRT, CMPTRT, CURTRT, CRMDUR, DOS, DOSU, DOSFRM, DOSFRQ, ROUTE, INTTYPE, PCLAS, PTRTDUR,	StudyIntervention	Multiple administrations are allowed per intervention, which may cause duplication of information. Deduplication is needed in these cases and for the treatment duration the duration of the different administrations probably need to be added up.

TS Flagging of Missing Data

As indicated in the User Requirements section, it should be possible to indicate what TS parameters are required and, if they are missing, what the corresponding "null flavor" should be. The null flavors are to be coded according to ISO 21090 (as described in the SDTMIG) and should be stored in the TSVALNF variable. The most likely used null flavors for the TS domain are:

Null flavor code	Description
NI	No information
NA	Not applicable
PINF	Positive infinity

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

The PINF null flavor is to be used in case the upper age range is not available. However, if a value of 120 or higher years is entered in USDM then this likely means positive infinity. The TSVAL variable can then be set to missing and the TSVALNF value to PINF.

TS Duration Formats

As described in the SDTMIG, duration and age values in the TS domain are required to be in ISO 8601 format. The solution needs to transform the value and corresponding units stored in USDM to the corresponding ISO format. A JSONata query for creating the ISO duration format of the planned treatment duration parameter (PTRTDUR) could look like this:

```
study.versions.studyInterventions{name: administrations.duration.quantity.("P"&
value & $substring(unit.standardCode.decode,0,1))}
```

However, then one is assuming that the duration is minimally expressed in days. More robust implementations will also conditionally add the "T" value after the "P" in case of an hour or minute duration.

TS Coding Variables

The TS domain includes variables for the parameter value (TSVAL), code (TSVALCD), code system (TSVCDREF), and code system version (TSVCDVER). In most cases using standardized TS parameters, the corresponding information is stored in USDM according to the Code or AliasCode complex datatype. The following table illustrates how attributes of those datatypes can be used to fill the corresponding TS variables.

USDM path within Code datatype	USDM path within AliasCode datatype	TS variable
decode	standardCode.decode	TSVAL
code	standardCode.code	TSVALCD
codeSystem	standardCode.codeSystem	TSVCDREF
codeSystemVersion	standardCode.codeSystemVersion	TSVCDVER

Note: The codeSystem specification for CDISC CT is validated for USDM to be "http://www.cdisc.org/"; for the TDM, "CDISC" is expected as the value for the TSVCDREF variable. The tool should transform this information accordingly when creating the TSVCDREF variable.

TS Y/N Variables

A number of parameters in the TS domain are expecting a boolean result and should be filled with Y or N. Examples of these are ADAPT (Adaptive Study Design Indicator) and DMCIND (Data Monitoring Committee Indicator). The corresponding information in USDM is either stored as a boolean variable or as the existence of a specific coded value from a codelist. If the mapping refers to a USDM boolean variable then the TS domain TSVAL variable should be filled with "Y" if the result in USDM is true. See, for example, the JSONata code for the healthy subject indicator:

```
$exists(study.versions.studyDesigns.population[includesHealthySubjects=true]) ?
"Y" :
($exists(study.versions.studyDesigns.population.cohorts[includesHealthySubjects
=true]) ? "Y" : "")
```

If the mapping refers to a certain response code stored in USDM, then it needs to be evaluated whether the corresponding code is specified in the corresponding mapped variable. See, for example, the JSONata code for checking the extension trial indicator:

```
$exists(study.versions.studyDesigns.characteristics[code="C207613"]) ? "Y" : ""
```

Note: If the Y/N values are being used for TSVAL, the corresponding code (TSVALCD), code system (TSVCDREF), and code system version (TSVCDVER) need to be added according to SDTM codelist C66742.

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

TS Count Variables

The number of arms and number of cohorts are not stored in USDM as exact numbers but can be based on counts of instances specified in the USDM for, respectively, the StudyArm class and StudyCohort class. For example, the number of arms can be queried with JSONata as follows:

```
$count(study.versions.studyDesigns.arms)
```

Creation of the TI domain

Eligibility Criteria Versioning

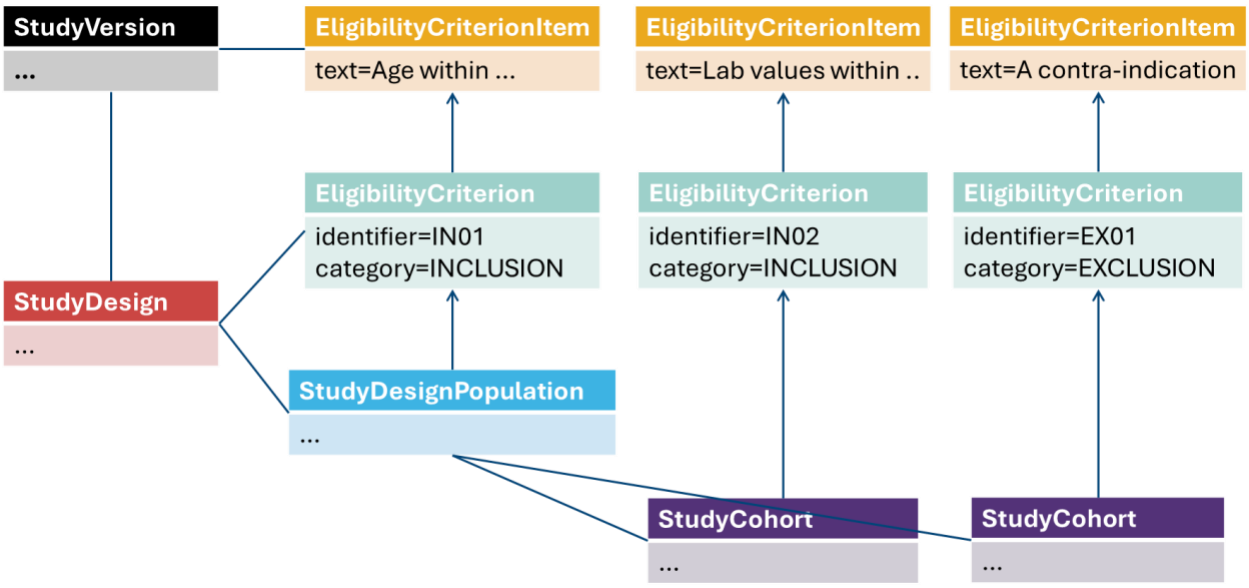
The TI domain variable TIVERS maps to the versionIdentifier attribute in the USDM StudyVersion class. Alternatively, the protocol version identifier might have been stored in the StudyDefinitionDocumentVersion domain as protocol version. Both options should be checked, as shown with this JSONata query:

```
$exists(study.versions.versionIdentifier) ? study.versions.versionIdentifier :
study.documentedBy[type.code="C70817"].versions.version
```

Further, per SDTMIG it is important to report all criteria that have been effective for any subjects enrolled in the study. When criteria are changed based on an amendment, the criteria applicable for each effective version need to be specified in the TI domain. However, the USDM API includes only 1 version of a study. In case of amendments, the tooling needs to compare all different versions exchanged via API calls and check whether there were any changes in the eligibility criteria. After completing the execution part of the study, it can be checked whether any subjects were enrolled for a specific study design version. When a version has not been effective for any included subject, the corresponding criteria can be disregarded for the submission in the TI domain output.

Eligibility Criteria Item Texts

The actual text of eligibility critera is stored in the USDM EligibilityCriterionItem class. This text can be utilized for multiple designs. The identifiers and eligibility-criteria categories specific to the design are stored in the EligibilityCriterion class. They point to the corresponding item text via the criterionItem relationship. Note that criteria may be applicable to the whole study design population or only to a specific cohort. The corresponding StudyDesignPopulation and StudyCohort class needs to reference those criteria that are applicable via their criteria relationships.



© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

The following JSONata code shows how the eligibility criteria text selected based on the criterionItemId reference in the API. This text can then be combined with the other information needed for the TI domain via its id value.

```
study.versions.($EligTxt := function($id) {(eligibilityCriterionItems[id=$id].text)};
studyDesigns.eligibilityCriteria{id: $EligTxt(criterionItemId)})
```

Transforming Text Features

Syntax Template Text Variables

The USDM includes the capability to tag information in text strings and link that to the corresponding digitized items stored elsewhere in the USDM. (This syntax template capability is described in the USDM-IG.) There are multiple USDM classes using this capability. For the TDM the relevant ones are Objective, Endpoint, and EligibilityCriterionItem. The corresponding text attribute used for the TDM is HTML formatted to allow for tagging. The tags, if any, need to be resolved and the value of the mapped attribute needs to be put in place of the tag. For this, the following actions are needed as exemplified by the ClinLine open-source USDM to trial design domain tool:

1. Retrieve the tag names, identified by the HTML text string. The following is a Python code example from the ClinLine open source USDM-to-SDTM Trial Design domain tool:

```
def ResolveTag(Txt,data):
# <usdm:tag name="min_age"/>
  while Txt.find("<usdm:tag") != -1:
      m = re.search(r'.*<usdm:tag name="([^"]*)"/>', Txt, re.DOTALL | re.IGNORECASE)
      if m:
         attrs = m.group(1)
         # print (attrs, m.end(0), m.start(1))
         NewTxt=Get_TagValue(attrs,data)
         Txt2=Txt[0:m.start(1)-16] + str(NewTxt) + Txt[m.end(0):len(Txt)]
         # print(Txt2)
         Txt=Txt2
  result=Get_plainText(Txt)
  return result
```

2. Search the tag name in the ParameterMap class and retrieve the reference to the corresponding referenced item.

3. Parse the reference information and search the corresponding information in the JSON file. Note that the InstanceType attribute, indicating the original class name, is needed to correctly get the information.

```
def Get_TagValue(tag,data):
   jsonataString = "study.versions.dictionaries.parameterMaps[tag='" + tag + "'].reference"
   expr = jsonata.Jsonata(jsonataString)
   reference = expr.evaluate(data)
   if reference is None:
      value="//TAG NOT IN DICTIONARY//"
   else:
      if reference[0] != "<":
         value=reference
      else:
         try:
            m = re.search(r'.*klass="([^"]*)"', reference, re.DOTALL | re.IGNORECASE)
            klass = m.group(1)
            m = re.search(r'.*id="([^"]*)"', reference, re.DOTALL | re.IGNORECASE)
            id = m.group(1)
            m = re.search(r'.*attribute="([^"]*)"', reference, re.DOTALL | re.IGNORECASE)
            attr = m.group(1)
            jsonataString3 = "study.**[instanceType='" + klass + "'  and id='" + id + "']." + attr
```

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

```
            # print(jsonataString2)
            expr2 = jsonata.Jsonata(jsonataString3)
            value = expr2.evaluate(data)
            # print("value: ", value)
        except:
            value="//TAG REFERENCE PARSING ERROR//"
    return value
```

4. Replace the tag with the retrieved information.

5. Before use for the Trial Design domains, the HTML formatting needs to be removed.

Maximum Length of Values (TS/TI)

As described in the SDTMIG the TSVAL variable in the TS domain and the IETEST variable in the TI domain have a maximum length of 200 characters. In cases where the mapping results in a string longer than 200 characters, multiple variables may be added to contain the remainder of the string. So, for the TS domain TSVAL2, TSVAL3, and so on can be additionally created; for the TI domain IETEST2, IETEST3, and so on. The SDTMIG describes rules for splitting texts in multiple strings (e.g., not in the middle of a word).

Per SDTMIG v3.4 (https://www.cdisc.org/standards/foundational/sdtmig), only 1 IETEST variable is allowed and criteria need to be abbreviated for submission if longer than 200 characters. In such cases, an AI agent could be used to suggest the abbreviated text, as long as the abbreviated text is confirmed by a user (i.e., human in the loop). The full text also needs to be stored, as it has to be specified in the corresponding metadata provided via the define.xml.

Note: This abbreviation requirement does not hold for the TSVAL variable containing the text value in the TS domain. Typical TS parameters that may contain longer strings include titles, objectives, and endpoints.

Special Characters and Symbols

Although not described in the SDTMIG, special characters in the text strings (e.g., ±, ≥, ≤, ©, ®) are often replaced by more simple text representations in ASCII to ease exchange of information and utilization in the submission package. This is also a requirement for FDA submissions; the *Study Data Technical Conformance Guide*[2] Section 3.1.5 indicates that variable values are the most broadly compatible with software and operating systems when they are restricted to ASCII text codes (printable values below 128).

Numbering and ordering

Actual visit numbers and ordering for epochs are not stored as numbers in USDM. Instead, the order is indicated by the previous and next reference attributes in the Encounter and StudyEpoch class, respectively. The mappings in the USDM to SDTM trial design domain mapping file for VISITNUM in the TV domain and TAEORD in the TA domain are referring to this next attribute. The instances in the corresponding classes need to be ordered according to this next reference and add the corresponding visit and epoch numbering accordingly. A method for ordering the epochs according to their next reference is included in the ta.py code of the ClinLine open source USDM-to-SDTM Trial Design domain tool (def sort_row_ids_by_epoch).

VISITNUM

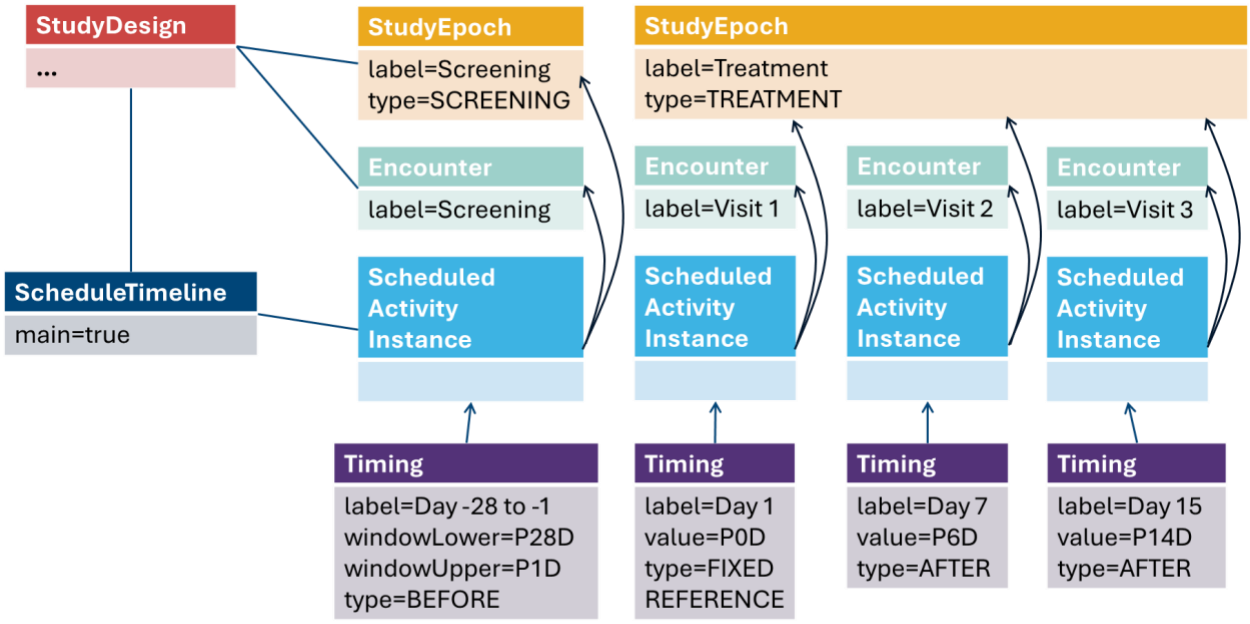
Additional company or sponsor rules may apply for VISITNUM values. This VISITNUM value may include the epoch and timing (day) information as well as the general visit/encounter information. The following example shows how the VISITNUM value could be created based on the combined information of these 3 timing aspects.

StudyEpoch	Encounter	Timing	VISITNUM
Screening	Screening		10001
Treatment	Visit 1	Day 1	20101
Treatment	Visit 2	Day 7	20207
Treatment	Visit 3	Day 15	20315
Follow-up	FU 1	Day 21	30121
Follow-up	FU 2	Day 42	30242

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

Although stored in the USDM, getting the corresponding logic into this format needs full understanding of the scheduleTimeline part of the USDM; this is further explained in the USDM-IG and the Understanding USDM training. This USDM part is representing the schedule of activities and refers to both the epochs and encounters.

The encounters in USDM may refer to the corresponding timing instance with the scheduledAt reference. However, if that reference is not included, it can also be retrieved via the corresponding timepoints stored as scheduledActivityInstance. The following schema shows how a chain of scheduled activity instances in the main timeline refer to both the corresponding encounter and the epoch. The instances in the timing class define the timing between 2 scheduledInstances.



Note: The pictured reference from the Timing instances to the ScheduledActivity instance is the relativeFromScheduledInstance reference. The relativeToScheduledInstance will in this case refer to the fixed Visit 1 scheduled activity instance.

Outputs

Output formats

The open-source tools described in the Examples and Use Cases section create TDM exports in csv and Excel format. Another option is to export the information as SAS Transport files or in Dataset-JSON format. Enabling the output in this format makes it easier to include the TDM data in a future aligned submission package. Examples, like the following for the TA arm, are provided in the CDISC Dataset-JSON repository (https://github.com/cdisc-org/DataExchange-DatasetJson).

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

```
{"datasetJSONCreationDateTime":"2024-11-11T15:09:18",
"datasetJSONVersion":"1.1.0",
"fileOID":"www.cdisc.org/StudyMSGv2/1/Define-XML_2.1.0/2024-11-11/ta",
"dbLastModifiedDateTime":"2020-08-21T09:14:26",
"originator":"CDISC SDTM MSG Team",
"sourceSystem":
   {"name":"SAS on X64_10PRO",
   "version":"9.0401M7"
   },
"studyOID":"cdisc.com/CDISCPILOT01",
"metaDataVersionOID":"MDV.MSGv2.0.SDTMIG.3.3.SDTM.1.7",
"metaDataRef":"define.xml",
"itemGroupOID":"IG.TA",
"records":8,
"name":"TA",
"label":"Trial Arms",
"columns":[
   {"itemOID":"IT.TA.STUDYID","name":"STUDYID","label":"Study Identifier","dataType":"string","length":12,"keySequence":1},
   {"itemOID":"IT.TA.DOMAIN","name":"DOMAIN","label":"Domain Abbreviation","dataType":"string","length":2},
   {"itemOID":"IT.TA.ARMCD","name":"ARMCD","label":"Planned Arm Code","dataType":"string","length":8,"keySequence":2},
   {"itemOID":"IT.TA.ARM","name":"ARM","label":"Description of Planned Arm","dataType":"string","length":28},
   {"itemOID":"IT.TA.TAETORD","name":"TAETORD","label":"Planned Order of Element within Arm","dataType":"integer","keySequence":3},
   {"itemOID":"IT.TA.ETCD","name":"ETCD","label":"Element Code","dataType":"string","length":7},
   {"itemOID":"IT.TA.ELEMENT","name":"ELEMENT","label":"Description of Element","dataType":"string","length":26},
   {"itemOID":"IT.TA.TABRANCH","name":"TABRANCH","label":"Branch","dataType":"string","length":200},
   {"itemOID":"IT.TA.TATRANS","name":"TATRANS","label":"Transition Rule","dataType":"string","length":200},
   {"itemOID":"IT.TA.EPOCH","name":"EPOCH","label":"Epoch","dataType":"string","length":9}],
"rows":[
   ["CDISCPILOT01","TA","PLACEBO","Placebo",1,"SCREEN","Screening","Randomized to Placebo","","SCREENING"],
   ["CDISCPILOT01","TA","PLACEBO","Placebo",2,"PLACEBO","Placebo","","","TREATMENT"],
   ["CDISCPILOT01","TA","ZAN_LOW","Zanomaline Low Dose (54 mg)",1,"SCREEN","Screening","Randomized to Zanomaline Low Dose","","SCREENING"],
   ["CDISCPILOT01","TA","ZAN_LOW","Zanomaline Low Dose (54 mg)",2,"LOW","Zanomaline 54 mg","","","TREATMENT"],
   ["CDISCPILOT01","TA","ZAN_HIGH","Zanomaline High Dose (81 mg)",1,"SCREEN","Screening","Randomized to Zanomaline High Dose","","SCREENING"],
   ["CDISCPILOT01","TA","ZAN_HIGH","Zanomaline High Dose (81 mg)",2,"TITRATE","Zanomaline 54 mg Titration","","","TREATMENT"],
   ["CDISCPILOT01","TA","ZAN_HIGH","Zanomaline High Dose (81 mg)",3,"HIGH","Zanomaline 81 mg","","","TREATMENT"],
   ["CDISCPILOT01","TA","ZAN_HIGH","Zanomaline High Dose (81 mg)",4,"TITRATE","Zanomaline 54 mg Titration","","","TREATMENT"]]}
```


Define xml

The define.xml file provides the metadata for all the SDTM datasets submitted (see https://www.cdisc.org/standards/data-exchange/define-xml). This includes the definitions of the Trial Design domains. To be efficient, the tooling should be able to create the SDTM metadata related to the trial design domains in order to include them in the define.xml file. When doing so,

- Specify the main structure as specified for define.xml
- Specify domains as ItemGroupDef, for example:

```
<ItemGroupDef OID="IG.TS" Name="TS" Domain="TS" Purpose="Tabulation" SasDatasetName="TS" Repeating="No" IsReferenceData="Yes" Structure="One
record per trial summary parameter value" ArchiveLocation="LF.TS" StandardOID="MDV1">
  <Description>
    <TranslatedText xml_lang="en">Trial Summary</TranslatedText>
  </Description>
</ItemGroupDef>
```

- Reference to variables within each ItemGroupDef as ItemRef, for example:

```
<ItemRef ItemOID="IT.STUDYID" Mandatory="Yes" OrderNumber="1"/>
<ItemRef ItemOID="IT.TA.DOMAIN" Mandatory="Yes" OrderNumber="2"/>
<ItemRef ItemOID="IT.TA.ARMCD" Mandatory="Yes" OrderNumber="3"/>
<ItemRef ItemOID="IT.TA.ARM" Mandatory="Yes" OrderNumber="4"/>
<ItemRef ItemOID="IT.TA.TAETORD" Mandatory="Yes" OrderNumber="5" MethodOID="MT.TAETORD"/>
<ItemRef ItemOID="IT.TA.ETCD" Mandatory="Yes" OrderNumber="6"/>
<ItemRef ItemOID="IT.TA.ELEMENT" Mandatory="No" OrderNumber="7"/>
<ItemRef ItemOID="IT.TA.TABRANCH" Mandatory="No" OrderNumber="8"/>
<ItemRef ItemOID="IT.TA.TATRANS" Mandatory="No" OrderNumber="9"/>
<ItemRef ItemOID="IT.TA.EPOCH" Mandatory="Yes" OrderNumber="10"/>
```

- Specify each variable as ItemDef, for example:

```
<ItemDef OID="IT.TS.TSPARMCD" Name="TSPARMCD" DataType="text" Length="8" SASFieldName="TSPARMCD">
  <Description>
    <TranslatedText xml:lang="en">Trial Summary Parameter Short Name</TranslatedText>
  </Description>
  <CodeListRef CodeListOID="CL.TSPARMCD"/>
  <def:Origin Type="Assigned" Source="Sponsor"/>
</ItemDef>
```

- Add definitions for derived elements (e.g., TAEORD, VISITNUM, TSGRPID, TSSEQ) and refer to these definitions from the ItemRef, for example:

```
<MethodDef OID="MT.TAETORD" Name="TAETORD Derivation Method" Type="Computation">
  <Description>
    <TranslatedText xml_lang="en">Sequential number identifying the order of epochs within an arm. Based on the previous and next epoch start date
    and time. The first epoch of an arm is assigned a TAETORD of 1.</TranslatedText>
  </Description>
</MethodDef>
```

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

- Add the used codes in variable specific codelist definitions and refer to the definitions from the ItemDef, for example:

```
<CodeList OID="CL.TSPARMCD" Name="Trial Summary Parameter Test Code" DataType="text" def:StandardOID="STD.3" SASFormatName="$TSPARMC">
  <CodeListItem CodedValue="ADDON">
    <Decode>
      <TranslatedText xml:lang="en">Added on to Existing Treatments</TranslatedText>
    </Decode>
    <Alias Context="nci:ExtCodeID" Name="C49703"/>
  </CodeListItem>
  <CodeListItem CodedValue="AGEMAX">
    <Decode>
      <TranslatedText xml:lang="en">Planned Maximum Age of Subjects</TranslatedText>
    </Decode>
    <Alias Context="nci:ExtCodeID" Name="C49694"/>
  </CodeListItem>
  <CodeListItem CodedValue="AGEMIN">
    <Decode>
      <TranslatedText xml:lang="en">Planned Minimum Age of Subjects</TranslatedText>
    </Decode>
```

- Add any necessary comment references and specify, for example:

```
<def:CommentDef OID="COM.COUNTRY">
  <Description>
    <TranslatedText xml:lang="en">This is a subset of the ISO-3166 (Country Codes) codelist</TranslatedText>
  </Description>
</def:CommentDef>
```

4.4 Validation

CDISC has specified a list of validation rules that can be applied to a specific standard. Both USDM and SDTM rules might be applicable to the creation of SDTM trial design domains. In addition, US FDA rules were created which validate FDA-specific requirements for submitted data in SDTM format. The USDM Implementation Handbook addresses the rules that are specifically focussed on the TDM (from both the USDM and the SDTM perspective). Most of the rules are implemented and can be run using the CDISC Open Rules Engine (CORE; https://www.cdisc.org/core).

USDM Validation Rules Related to TDM

The USDM CORE rules (https://github.com/cdisc-org/USDM_CORE_Rules.xlsx) include a number of general rules that validate the schema, including attributes and relationships, as well as rules that are content specific. The USDM CORE rules include those that check relevant SDTM terminology, ordering of epochs and encounters, and objective-endpoint relationships. The following table lists USDM v4.0-specific CORE rules that are applicable to the SDTM Trial Design domains. They are a subset of the full list published for Version 4.0 (https://www.cdisc.org/ddf/usdm-errata-3-4). Future versions may include changes and/or additions.

CORE rule ID	Rule description (USDM v4.0)	ERROR/ WARNING	Relevant USDM classes	Relevant USDM attributes	Relevant SDTM domain variables
DDF00023	To ensure consistent ordering, when both previous and next attributes are available within an entity the previous id value must match the next id value of the referred instance.	ERROR	EligibilityCriterion, Encounter, StudyEpoch	previous, next	TI ordering, TV.VISITNUM, TA.TAETORD
DDF00027	To ensure consistent ordering, the same instance must not be referenced more than once as previous or next.	ERROR	EligibilityCriterion, Encounter, StudyEpoch	previous, next	TI ordering, TV.VISITNUM, TA.TAETORD
DDF00179	An administrable dose form must be specified according to the extensible Pharmaceutical Dosage Form (C66726) SDTM codelist (e.g., an entry with a	ERROR	AdministrableProduct	administrableDoseForm	TS.DOSFRM

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

CORE rule ID	Rule description (USDM v4.0)	ERROR/ WARNING	Relevant USDM classes	Relevant USDM attributes	Relevant SDTM domain variables
	code or decode used from the codelist should be consistent with the full entry in the codelist).				
DDF00178	If a dose is specified then a corresponding frequency must also be specified.	ERROR	Administration	dose	TS.DOSE, TS.DOSFRQ
DDF00176	If an administration's dose is specified then a corresponding route is expected and vice versa.	WARNING	Administration	route	TS.DOSE, TS.ROUTE
DDF00073	Only 1 version of any code system is expected to be used within a study version.	WARNING	Code	codeSystemVersion	TS.TSVCDREF, TS.VCDVER
DDF00155	For CDISC codelist references (where the code system is "http://www.cdisc.org/"), the code system version must be a valid CDISC CT release date in ISO 8601 date format.	ERROR	Code	codeSystemVersion	TS.VCDVER
DDF00110	An eligibility criterion's category must be specified using the Category of Inclusion/Exclusion (C66797) SDTM codelist.	ERROR	EligibilityCriterion	category	TI.IECAT
DDF00246	Any parameter name referenced in a tag in the text should be specified in the data dictionary parameter maps.	ERROR	EligibilityCriterionItem, Objective, Endpoint	text	TI.IETEST, TS.OBJPRIM, TS.OBJSEC, TS.OBJEXP, TS.OUTMSPRI, TS.OUTMSSEC, TS.OUTMSEXP
DDF00247	Syntax template text is expected to be HTML formatted.	WARNING	EligibilityCriterionItem, Objective, Endpoint	text	TI.IETEST, TS.OBJPRIM, TS.OBJSEC, TS.OBJEXP, TS.OUTMSPRI, TS.OUTMSSEC, TS.OUTMSEXP
DDF00041	Within a study design, there must be at least 1 endpoint with level primary.	ERROR	Endpoint	level	TS.OUTMSPRI
DDF00096	All primary endpoints must be referenced by a primary objective.	ERROR	Endpoint	level	TS.OBJPRIM, TS.OUTMSPRI
DDF00148	An endpoint level must be specified using the Endpoint Level (C188726) DDF codelist.	ERROR	Endpoint	level	TS.OUTMSPRI, TS.OUTMSSEC, TS.OUTMSEXP
DDF00217	A study design's blinding schema must be specified according to the extensible Trial Blinding Schema Response (C66735) SDTM codelist (e.g., an entry with a code or decode used from the codelist should be consistent with the full entry in the codelist).	ERROR	InterventionalStudyDesign	blindingSchema	TS.TBLIND
DDF00214	An interventional study design's intent types must be specified according to the extensible Trial Intent Type Response (C66736) SDTM codelist (e.g., an entry with a code or decode used from the codelist should be consistent with the full entry in the codelist).	ERROR	InterventionalStudyDesign	intentTypes	TS.TINDTP
DDF00222	Within a study design, if more intent types are defined, they must be distinct.	ERROR	InterventionalStudyDesign	intentTypes	TS.TINDTP
DDF00213	If the intervention model indicates a single group design then only one intervention is expected. In all other cases more interventions are expected.[a]	WARNING	InterventionalStudyDesign	model	TS.INTMODEL
DDF00216	A study design's intervention model must be specified according to the extensible Intervention Model Response (C99076) SDTM codelist (e.g., an entry with a code or decode used from the codelist should be consistent with the full entry in the codelist).	ERROR	InterventionalStudyDesign	model	TS.INTMODEL
DDF00215	An interventional study design's subtypes must be specified according to the extensible Trial Type Response (C66739) SDTM codelist (e.g., an entry with a	ERROR	InterventionalStudyDesign	subTypes	TS.TTYPE

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

CORE rule ID	Rule description (USDM v4.0)	ERROR/ WARNING	Relevant USDM classes	Relevant USDM attributes	Relevant SDTM domain variables
	code or decode used from the codelist should be consistent with the full entry in the codelist).				
DDF00218	A study design's characteristics must be specified according to the extensible study design characteristics (C207416) DDF codelist (e.g., an entry with a code or decode used from the codelist should be consistent with the full entry in the codelist).	ERROR	InterventionalStudyDesign, ObservationalStudyDesign	characteristics	TS.ADAPT, TS.EXTTIND, TS.RANDOM, TS.MSEUTIND, TS.SSEUTIND *more?*
DDF00219	Within a study design, if more characteristics are defined, they must be distinct.	ERROR	InterventionalStudyDesign, ObservationalStudyDesign	characteristics	TS.ADAPT, TS.EXTTIND, TS.RANDOM, TS.MSEUTIND, TS.SSEUTIND *more?*
DDF00154	A study design must not be characterized as both "Single-Centre" and "Multicentre".	ERROR	InterventionalStudyDesign, ObservationalStudyDesign	characteristics	TS.MSEUTIND, TS.SSEUTIND
DDF00258	A study design is not expected to have more than one of the following characteristics: "Randomized", "Stratification", "Stratified Randomisation".	WARNING	InterventionalStudyDesign, ObservationalStudyDesign	characteristics	TS.RANDOM
DDF00220	Within a study design, if more subtypes are defined, they must be distinct.	ERROR	InterventionalStudyDesign, ObservationalStudyDesign	subTypes	TS.TTYPE
DDF00221	Within a study design, if more therapeutic areas are defined, they must be distinct.	ERROR	InterventionalStudyDesign, ObservationalStudyDesign	therapeuticAreas	TS.THERAREA
DDF00084	Within a study design there must be exactly 1 objective with level "Primary Objective".[b]	WARNING	Objective	level	TS.OBJPRIM
DDF00147	An objective level must be specified using the Objective Level (C188725) DDF codelist.	ERROR	Objective	level	TS.OBJPRIM, TS.OBJSEC, TS.OBJEXP
DDF00223	A study design's observational model must be specified according to the extensible Observational Study Model (C127259) SDTM codelist (e.g., an entry with a code or decode used from the codelist should be consistent with the full entry in the codelist).	ERROR	ObservationalStudyDesign	model	TS.OBSMODEL
DDF00225	An observational study design's sampling method must be specified according to the extensible Observational Study Sampling Method (C127260) SDTM codelist (e.g., an entry with a code or decode used from the codelist should be consistent with the full entry in the codelist).	ERROR	ObservationalStudyDesign	samplingMethod	TS.OBSTSMM
DDF00232	An observational study (including patient registries) is expected to have a study phase decode value of "NOT APPLICABLE".[c]	WARNING	ObservationalStudyDesign	studyPhase	TS.TPHASE
DDF00224	An observational study design's time perspective must be specified according to the extensible Observational Study Time Perspective (C127261) SDTM codelist (e.g., an entry with a code or decode used from the codelist should be consistent with the full entry in the codelist).	ERROR	ObservationalStudyDesign	timePerspective	TS.OBSTIMP
DDF00229	A study design's study phase must be specified according to the extensible Trial Phase Response (C66737) SDTM codelist (e.g., an entry with a code or decode used from the codelist should be consistent with the full entry in the codelist).	ERROR	ObservationalStudyDesign, InterventionalStudyDesign	studyPhase	TS.TPHASE
DDF00230	A study design's study type must be specified using the Study Type Response (C99077) SDTM codelist.	ERROR	ObservationalStudyDesign, InterventionalStudyDesign	studyType	TS.STYPE
DDF00124	Referenced items in a parameter map must be available elsewhere in the data model.	ERROR	ParameterMap	reference	TI.IETEST, TS.OBJPRIM, TS.OBJSEC, TS.OBJEXP,

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

CORE rule ID	Rule description (USDM v4.0)	ERROR/ WARNING	Relevant USDM classes	Relevant USDM attributes	Relevant SDTM domain variables
					TS.OUTMSPRI, TS.OUTMSSEC, TS.OUTMSEXP
DDF00137	References must be a fixed value or a reference to items stored elsewhere in the data model which must be specified in the correct format. They must start with '<usdm:ref', end with either '/>' or '></usdm:ref>', and must contain 'klass="klassName"', 'id="idValue"', and 'attribute="attributeName"/>' in any order (where "klassName" and "attributeName" contain only letters in upper or lower case).	ERROR	ParameterMap	reference	TI.IETEST, TS.OBJPRIM, TS.OBJSEC, TS.OBJEXP, TS.OUTMSPRI, TS.OUTMSSEC, TS.OUTMSEXP
DDF00030	At least the text or the family name must be specified for a person name.	ERROR	PersonName	text, familyName	TS.CONNAME
DDF00233	A unit must be coded according to the extensible unit (C71620) SDTM codelist (e.g. an entry with a code or decode used from the codelist should be consistent with the full entry in the codelist).	ERROR	Quantity	unit	TS.DOSU
DDF00042	The range specified for a planned age is not expected to be approximate.	WARNING	Range	isApproximate	TS.AGEMIN, TS.AGEMAX
DDF00241	If the unit is the same (or missing) for both the minimum and maximum value, then the minimum value must be less than the maximum value.	ERROR	Range	minValue, maxValue	TS.AGEMIN, TS.AGEMAX
DDF00071	A study cell must only reference an arm that is defined within the same study design as the study cell.	ERROR	StudyCell	arm	TA
DDF00069	Each combination of arm and epoch must occur no more than once within a study design.	ERROR	StudyCell	arm, epoch	TA
DDF00243	Each StudyArm is expected to have 1 StudyCell for each StudyEpoch.	WARNING	StudyCell	arm, epoch	TA
DDF00040	Each study element must be referenced by at least 1 study cell.	ERROR	StudyCell	elements	TA, TE
DDF00047	A study cell must only reference elements that are defined within the same study design as the study cell.	ERROR	StudyCell	elements	TA, TE
DDF00072	A study cell must only reference an epoch that is defined within the same study design as the study cell.	ERROR	StudyCell	epoch	TA, TE
DDF00097	Within a study design, the planned age range must be specified either in the study population or in all cohorts.	ERROR	StudyDesignPopulation, StudyCohort	plannedAge	TS.AGEMIN, TS.AGEMAX
DDF00237	The unit of a planned age is expected to be specified using terms from the Age Unit (C66781) SDTM codelist.	ERROR	StudyDesignPopulation, StudyCohort	plannedAge	TS.AGEMIN, TS.AGEMAX
DDF00132	Within a study design, if a planned completion number is defined, it must be specified either in the study population or in all cohorts.	ERROR	StudyDesignPopulation, StudyCohort	plannedCompletionNumber	TS.PLANSUB
DDF00235	A unit must not be specified for a planned completion number.	ERROR	StudyDesignPopulation, StudyCohort	plannedCompletionNumber	TS.PLANSUB
DDF00098	Within a study design, the planned sex must be specified either in the study population or in all cohorts.	ERROR	StudyDesignPopulation, StudyCohort	plannedSex	TS.SEXPOP
DDF00141	A planned sex must be specified using the Sex of Participants (C66732) SDTM codelist.	ERROR	StudyDesignPopulation, StudyCohort	plannedSex	TS.SEXPOP
DDF00188	A planned sex must ether include a single entry of male or female or both female and male as entries.	ERROR	StudyDesignPopulation, StudyCohort	plannedSex	TS.SEXPOP
DDF00088	Epoch ordering using previous and next attributes is expected to be consistent with the order of corresponding scheduled activity instances according to their specified default conditions.	WARNING	StudyEpoch	previous, next	TA.TAETORD
DDF00024	An epoch must only reference epochs that are specified within the same study design.	ERROR	StudyEpoch	previous, next	TA.TAETORD
DDF00022	An instance of a class must not refer to itself as its next instance.	ERROR	StudyEpoch, Encounter	next	TA.TAETORD, TV.VISITNUM

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

CORE rule ID	Rule description (USDM v4.0)	ERROR/ WARNING	Relevant USDM classes	Relevant USDM attributes	Relevant SDTM domain variables
DDF00021	An instance of a class must not refer to itself as its previous instance.	ERROR	StudyEpoch, Encounter	previous	TA.TAETORD, TV.VISITNUM
DDF00172	There must be exactly 1 sponsor study identifier (i.e., a study identifier whose scope is an organization that is identified as the organization for the sponsor study role).	ERROR	StudyIdentifier	scope	TS.STUDYID, TI.STUDYID, TV.STUDYID, TA.STUDYID, TE.STUDYID
DDF00210	An administrable product's product designation must be specified using the Product Designation (C207418) DDF codelist.	ERROR	StudyIntervention	productDesignation	TS.PCLAS
DDF00112	A study intervention's role must be specified using the Study Intervention Role (C207417) DDF codelist.	ERROR	StudyIntervention	role	TS.TRT, TS.CURTRT, TS.COMPTRT, TS.TCNTRL
DDF00128	A study intervention's type must be specified using the Intervention Type Response (C99078) SDTM codelist.	ERROR	StudyIntervention	type	TS.INTTYPE
DDF00259	A study role code must be specified according to the Study Role Code (C215480) DDF codelist (e.g., an entry with a code or decode used from the codelist should be consistent with the full entry in the codelist).[d]	ERROR	StudyRole	code	TS.STUDYID, TI.STUDYID, TV.STUDYID, TA.STUDYID, TE.STUDYID
DDF00146	A study title type must be specified using the Study Title Type (C207419) DDF codelist.	ERROR	StudyTitle	type	TS.TITLE
DDF00100	Within a study version, there must be no more than 1 title of each type.	ERROR	StudyVersion	titles	TS.TITLE
DDF00115	Every study version must have a title of type "Official Study Title".	ERROR	StudyVersion	titles	TS.TITLE

[a]This rule limited by the errata published December 2025 (https://www.cdisc.org/ddf/usdm-errata-3-4).
[b]This rule changed to warning by the errata published December 2025.
[c]According to the errata published December 2025, validation is based on the code corresponding to "Not Applicable" and not on the decode value.
[d]Adjusted rule text indicating that the codelist is extensible according to the errata published December 25.

SDTM Validation Rules Related to TDM

The set of SDTM CORE rules (https://www.cdisc.org/standards/) include the rule set according to different versions of the SDTMIG. The following table provides a list of executable rules related to the SDTM Trial Design domains from SDTMIG Version 3.4 that can be executed on the resulting Trial Design domains described in this Implementation Handbook. Based on these rules, some additional requirements might be needed for the tooling.

Rule ID	Cited guidance/rule (SDTMIG v3.4)	ERROR/ WARNING	Relevant USDM classes	Relevant USDM attributes	Relevant SDTM domain parameters
CG0153	ARMCD is limited to 20 characters and does not have the character restrictions that apply to --TESTCD.	ERROR	StudyArm	label or alternative mapped item	TA.ARMCD
CG0154	ELEMENT and ETCD have a one-to-one relationship.	ERROR	StudyElement	label and description or alternative mapped items	TA.ETCD, TA.ELEMENT, TE.ETCD, TE.ELEMENT
CG0246	ETCD (the companion to ELEMENT) is limited to 8 characters.	ERROR	StudyElement	label or alternative mapped item	TA.ETCD, TE.ETCD
CG0247	Number that gives the order of the Element within the Arm (TAETORD is unique within an ARM).	ERROR	*Not applicable; validation of TDM automation feature functionality*		TA.TAEORD
CG0248	TAEORD is an integer.	ERROR	*Not applicable; validation of TDM automation feature functionality*		TA.TAEORD

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

Rule ID	Cited guidance/rule (SDTMIG v3.4)	ERROR/ WARNING	Relevant USDM classes	Relevant USDM attributes	Relevant SDTM domain parameters
GC0249	If an Element does not end with a decision that could lead to a shortened path within the Arm, then TATRANS will be blank.	ERROR	*Not applicable; validation of TDM automation feature functionality*		TA.TATRANS
CG0250	EPOCH may be used as a timing variable in other datasets, such as EX and DS, and values of EPOCH must be different for different epochs.	ERROR	StudyEpoch	label or alternative mapped item	TA.EPOCH
CG0256	If a criterion changes, it should be treated as a new criterion, with a new value for IETESTCD (IETESTCD unique in dataset).	ERROR	EligibilityCriterion	identifier	TI.IETESTCD
CG0257	TSPARMCD (the companion to TSPARM) is limited to 8 characters and does not have special character restrictions.	ERROR	*See mappings for TS domain. These include standard CDISC CT values. However, extensions specified by users should comply to this rule.*		TS.TSPARMCD
CG0258	The value in TSPARM cannot be longer than 40 characters.	ERROR	*See mappings for TS domain. These include standard CDISC CT values. However, extensions specified by users should comply to this rule.*		TS.TSPARM
CG0259	TSVAL can only be null when TSVALNF is populated.	ERROR	*See mappings for TS domain and instructions in this Implementation Handbook.*		TS.TSVAL, TS.TSVALNF
CG0260	Null flavor (TSVALNF) for the value of TSPARM, to be populated if and only if TSVAL is null.	ERROR	*See mappings for TS domain and instructions in this Implementation Handbook.*		TS.TSVAL, TS.TSVALNF
CG0261	Text over 200 characters can be added to additional columns TSVAL1-TSVALn. (TSVAL ^= null when TSVAL1 ^= null)	ERROR	*Not applicable; validation of TDM automation feature functionality*		TS.TSVAL, TS.TSVAL1-TS.TSVALn
CG0262	Text over 200 characters can be added to additional columns TSVAL1-TSVALn. (TSVALn ^= null when TSVAL(n+1) ^= null)	ERROR	*Not applicable; validation of TDM automation feature functionality*		TS.TSVAL1-TS.TSVALn
CG0265	TSVALCD and TSVAL have a one-to-one relationship (condition: TSVAL ^= null and TSVALCD ^= null).	ERROR	*Not applicable; validation of TDM automation feature functionality*		TS.TSVAL, TS.TSVALCD
CG0266	The version number of the Reference Terminology, if applicable. (TSVCDREF ^= null when TSVCDVER ^= null)	ERROR	code	codeSystem, codeSystemVersion	TS.TSVCDREF, TS.TSVCDVER
CG0268	TSSEQ is unique for each distinct value of TSPARMCD (when TSPARMCD is not unique)	ERROR	*Not applicable; validation of TDM automation feature functionality*		TS.TSPARCD, TS.TSSEQ
CG0270	When the trial summary parameter is AGEMAX, then TSVAL should have a value expressed as an ISO8601 time duration (e.g., P43Y for 43 years old or P6M for 6 months old).	ERROR	*Not applicable; validation of TDM automation feature functionality*		TS.TSVAL
CG0288	TSVALCD = valid code in the version identified in TSVCDVER when TSVCDREF='CDISC'	ERROR	code	code, codeSystem, codeSystemVersion	TS.TSVALCD, TS.TSVCDREF, TS.TSVCDVER
CG0289	TSVCDVER = valid published version (date) when TSVCDREF="CDISC"	ERROR	code	codeSystem, codeSystemVersion	TS.TSVCDREF, TS.TSVCDVER
CG0291	TSVAL not populated with values or synonyms of values in the ISO 21090 null flavor codelist (or other terms that can be represented as null flavors)	WARNING	*See mappings for TS domain and instructions in this Implementation Handbook.*		TSVAL
CG0297	ARMCD is limited to 20 characters and does not have special character restrictions.	ERROR	StudyArm	label or alternative mapped item	TV.ARMCD
CG0307	TSPARM and TSPARMCD have a one-to-one relationship	ERROR	*Not applicable; validation of mapping.*		TS.TSPARM, TS.TSPARMCD
CG0325	The combination of ELEMENT, TESTRL, TEENRL, and TEDUR is unique for each ETCD	ERROR	StudyElement	label and description or alternative mapped items, transitionStartRule.text, transitionEndRule.text	TE.ETCD, TE.ELEMENT, TE.TESTRL, TE.TEENRL, TE.TEDUR
CG0328	TEENRL ^= null when TEDUR = null	ERROR	StudyElement	transitionEndRule.text	TE.TEENRL, TE.TEDUR
CG0329	TEDUR ^= null when TEENRL = null	ERROR	StudyElement	transitionEndRule.text	TE.TEENRL, TE.TEDUR
CG0372	'--TESTCD <= 8 chars and contains only letters, numbers, and underscores and can not start with a number	ERROR	EligibilityCriterion	identifier	TI.IETESTCD

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

Rule ID	Cited guidance/rule (SDTMIG v3.4)	ERROR/ WARNING	Relevant USDM classes	Relevant USDM attributes	Relevant SDTM domain parameters
CG0649	TSVAL not populated with values or synonyms of values in the ISO 21090 null flavor codelist (or other terms that can be represented as null flavors), unless the term is included in published terminology.	ERROR	*Not applicable; validation of TDM automation feature functionality*		TS.TSVAL

Other Validation Rules Related to TDM

The TDM-related standard rules presented in this section were published for USDM and SDTM. However, some custom checks might be needed to identify additional submission requirements, validate codings from external codelist (e.g., therapeutic area, indications), and verify whether the content aligns to company procedures.

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

5 Challenges and Considerations

Regulatory Requirements for Trial Summary Parameters

The list of standardized TS parameters is documented in the CDISC SDTM CT dictionary (https://evs.nci.nih.gov/ftp1/CDISC/SDTM/). However, the parameters required for a submission may be specific to a regulatory authority; for example, the FDA *Study Data Technical Conformance Guide*[2] indicates the list of parameters that should be submitted to the FDA.

In addition, the expectations for submitted TS parameters and other information may change over time, along with what is covered by the USDM structure and corresponding solutions. Therefore, it is important that the tooling is flexible in enabling:

- additional mappings to USDM content,
- the addition of custom newly created content variables without direct link to the study design, and
- change in mappings to USDM content.

The example ClinLine USDM-to-trial-design domain tool, described in Examples and Use Cases, shows how most of the JSONata mappings are disconnected from the actual core solution of creating the domains. This makes it more flexible to change and add any mappings to USDM if and when needed.

Impact of Amendments

Most clinical studies are amended during the actual conduct of the study. This implies that part of the study subjects follow a (slightly) different study design after the study amendment became effective. All the official study versions need to be included for creating the Trial Design domains—the original design and all official amendments. A USDM JSON file representing a study design, by definition, only includes 1 study version. Therefore, it is important to upload the USDM files of each official study version and compare the trial design information that is needed to create the SDTM trial design domains. The Technical Implementation section of this Implementation Handbook describes how changes over time should be processed and represented for the TI domain. The content of this domain is most likely subject to change based on amendments. Any other changes that occur to information that is needed for the Trial Design domains need to be verified to decide if and how it should be included in the final submission.

Study Design Complexities

The increasing complexity of study designs makes the creation of study designs more challenging. Complexities that might affect the creation of the design include but are not limited to:

- Adaptive designs
- Multiphase studies
- Decentralized trials
- External control arm studies
- Complex dosing schemes

From a complex study design perspective, different designs, phases, and arms could be stored as one design or, if they are clearly different, in separate study designs in the USDM format. From the submission perspective, a company may choose to submit the different parts, arms, or phases as separate SDTM packages or combine these into one submission. If the information is stored in separate designs in USDM but has to be submitted as one package, then the information of the separate designs needs to be compared and any differences affecting the Trial Design domains need to be verified. If information stored in one design has to be submitted as separate packages then some design features might need to be split. For example, if the design indicates a study phase of "Phase 1/Phase 2", then one submission might be applicable for phase 1 and another for phase 2, which means that the study phase needs to be defined more specifically for the corresponding submission.

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

Dependency on Study Execution

Although most information to be submitted in the Trial Design domains is based on the study design, some of the parameters are operational or based on the study execution. Examples of this are:

- Actual Number of Subjects (ACTSUB)
- CDISC therapeutic area user guide (CTAUG)
- Study Start date and Study End date (SSTDTC and SENDTC)
- SDTM version (SDTMVER)
- SDTMIG version (SDTIGVER)

Further, during execution, there is still a possibility that additional amendments may be needed. Therefore, although a preliminary set of TS domains can be created and validated early during study execution, the final quality control cannot take place before the end of the study. For efficiency reasons, a partial validation of the fixed study design information can be executed and corresponding information locked. This would cover over 80% of the information, meaning that only changed and added information needs to be validated at the end. This approach only works when, in case of amendments, a good and validated comparison of study design information takes place that identifies any changes that occurred between versions and indicates the revalidation needed for those changes.

Cycles and Unscheduled Visits

For oncology studies, in principle, the number of cycles that a patient can go through is limitless. The USDM provides the ability to define cycles as presented in a schedule of activities (e.g., "Cycle 2-4", "Cycle 3 and following"). As the number of cycles is limitless, it is not possible to define the total number of actual cycles in the design phase. However, in the submission, the actual visits are presented in the TV domain and they are usually not grouped. This includes all the actual cycle visits that at least 1 patient has encountered during the study. A solution creating the Trial Design domains therefore needs to provide the option to copy the combined cycle information for each cycle and enable customization per actual cycle. This process may be automated based actual study execution information at the end of the study, if and when available.

Unscheduled visits can be also specified in the USDM. This is beneficial when the specific assessments of an unscheduled visit are predefined. These unscheduled visits are usually not linked to the main timeline in USDM, but specified as a subtimeline with a separate entry condition. Unscheduled visits are not to be specified in the TV domain; per the SDTMIG, this should only include planned visits. Therefore, unscheduled visits need to be excluded from the mapping for the Trial Design domains.

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

6 Examples and Use Cases

Study Design Solution

The CDISC COSA Repository Directory (https://cosa.cdisc.org/) provides information on open-source solutions that provide examples on how to implement the study design in USDM format. Examples of these include (1) the soa_workbench, a Python web app created within the CDISC 360i project, which focuses on study pathways and timelines; (2) the Study Definitions Workbench, which is focused on elements that align to M11; and (3) OpenStudyBuilder, which provides a solution to include the full study design which can be imported in USDM format.

SDTM Trial Design Domain Creation

There are 2 open-source tools available to show the actual creation of Trial Design domains.

cdisc-USDM-utils

This tool (https://github.com/pendingintent/cdisc-usdm-utils) was created in phase 1 of the CDISC 360i project. Although based on an earlier version of USDM, it provides some good examples of how information can be retrieved from USDM using the Python JSON module, logically processed, and how it can be exported in csv format. The tool creates an initial set of the Trial Design domains TA, TE, TI, TS, and TV. Future updates are expected to account for more complex mappings. Note that some of the mappings (e.g., to the name attributes of classes) might not be accurate. Verification with the actual USDM-to-trial design domain mapping file is needed for this (see https://github.com/cdisc-org/DDF-RA/sdtm_mapping.xlsx).

ClinLine/USDM_SDTM_mapper

This tool (https://github.com/ClinLine/USDM_SDTM_mapper) was created directly after the release of USDM v4.0. For the creation of the domains, the mappings in the published USDM-to-trial design domain mapping file are transformed to JSONata queries, stored in the original Excel file, and then picked up by the Python tool. The Python tool includes the creation of the TA, TE, TI, TS, and TV files, as well as a first start for the metadata stored in define.xml format. Although not complete, it shows most of the required features described in this Implementation Handbook.

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

7 Appendices

Appendix A: Implementation Handbook Team

Name	Institution/Organization
John Owen	Project Manager, CDISC
Berber Snoeijer	USDM Expert and Primary Technical Writer, CDISC
Darren Moreland	Co-technical Writer and Technical Reviewer, CDISC
Edwin van Stein	Co-technical Writer and Technical Reviewer, CDISC
Anthony Chow	Use Case SME, CDISC
Lorraine Sobson	Copyeditor, CDISC

CDISC would like to thank the following individuals for their review of the Implementation Handbook

Name
Neha Begum
Chris Damato
Bill Illis
Nicole Thorne
Pratiksha Wani
Diane Wold

This Implementation Handbook has been developed in partnership with TransCelerate Biopharma.

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

Appendix B: Glossary and Abbreviations

The following abbreviations and terms are used in this document. Additional definitions can be found in the sections of this Implementation Handbook, in the USDM-IG v4.0, and in the CDISC Glossary.

Term	Description
ADaM	Analysis Data Model
AI	Artificial intelligence
API	Application programming interface
ASCII	American Standard Code for Information Interchange
CDISC	Clinical Data Interchange Standards Consortium
CORE	CDISC Open Rules Engine
csv	comma-separated values
CT	(CDISC) Controlled Terminology
CTIS	Clinical Trials Information System; a registry for clinical trials in the European Union, facilitating the registration and management of trial data for sponsors and regulatory authorities
Dataset-JSON	Data exchange standard for sharing tabular data using JSON; designed to meet a wide range of data exchange scenarios, including regulatory submissions and API-based data exchange
DDF	Digital data flow
Define-XML	Define-XML transmits metadata in an xml format that describes any tabular dataset structure. When used with the CDISC foundational standards, it provides the metadata for human and animal model datasets using the SDTM and/or SEND standards and analysis datasets using ADaM.
EDC	electronic data capture
EMA	European Medicines Agency
ETL	Extraction, transformation, and loading (of data)
FDA	(US) Food and Drug Administration
HTML	Hypertext markup language
ISO	International Organization for Standardization
JSON	JavaScript Object Notation, a lightweight, text-based data-interchange format designed to be human-readable and easy for machines to parse and generate
JSONata	JSON query and transformation language
MedDRA	Medical Dictionary for Regulatory Activities
MED-RT	Medication Reference Terminology
NCT number	National Clinical Trial number assigned to clinical studies registered on ClinicalTrials.gov
Python	An interpreted, object-oriented, high-level programming language with dynamic semantics
SDTM	Study Data Tabulation Model
SDTMIG	SDTM Implementation Guide
SNOMED CT	Systematized Nomenclature of Medicine -- Clinical Terms: Controlled terminology providing a standardized way to represent clinical phrases captured by the clinician and enables automatic interpretation of these
SOA	Schedule of activities
Study design solution	Either custom built software or commercially available software that aids in digitally capturing the study design which can be represented as a clinical study protocol and can utilize/export the design in USDM format
TDM	Trial Design Model
TDM functionality	Software functionality to automatically create domains according to the TDM, which may either be included in a study design solution or separate downstream solutions
UML	Unified modeling language
USDM	Unified Study Definitions Model
USDM-IG	USDM Implementation Guide
USDM-HB1	USDM Implementation Handbook 1

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

Appendix C: References

1. TransCelerate. *Practical Approach to Implmementing Digital Data Flow: A Framework to Getting Started.* TransCelerate Biopharma, Inc. August 2025. Accessed May 7, 2026. https://github.com/transcelerate/ddf-home/blob/main/documents/white_paper/DDF_Practical_Approach_to_Implementation.pdf
2. US Food and Drug Administration. *Study Data Technical Conformance Guide - Technical Specifications Document.* Center for Drug Evaluation and Research. Published March 2026. Accessed May 7, 2026. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/study-data-technical-conformance-guide-technical-specifications-document

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved

Appendix D: Representations and Warranties, Limitations of Liability, and Disclaimers

CDISC Patent Disclaimers
It is possible that implementation of and compliance with this standard may require use of subject matter covered by patent rights. By publication of this standard, no position is taken with respect to the existence or validity of any claim or of any patent rights in connection therewith. CDISC, including the CDISC Board of Directors, shall not be responsible for identifying patent claims for which a license may be required in order to implement this standard or for conducting inquiries into the legal validity or scope of those patents or patent claims that are brought to its attention.

Representations and Warranties
"CDISC grants open public use of this User Guide (or Final Standards) under CDISC's copyright."

Each Participant in the development of this standard shall be deemed to represent, warrant, and covenant, at the time of a Contribution by such Participant (or by its Representative), that to the best of its knowledge and ability: (a) it holds or has the right to grant all relevant licenses to any of its Contributions in all jurisdictions or territories in which it holds relevant intellectual property rights; (b) there are no limits to the Participant's ability to make the grants, acknowledgments, and agreements herein; and (c) the Contribution does not subject any Contribution, Draft Standard, Final Standard, or implementations thereof, in whole or in part, to licensing obligations with additional restrictions or requirements inconsistent with those set forth in this Policy, or that would require any such Contribution, Final Standard, or implementation, in whole or in part, to be either: (i) disclosed or distributed in source code form; (ii) licensed for the purpose of making derivative works (other than as set forth in Section 4.2 of the CDISC Intellectual Property Policy ("the Policy")); or (iii) distributed at no charge, except as set forth in Sections 3, 5.1, and 4.2 of the Policy. If a Participant has knowledge that a Contribution made by any Participant or any other party may subject any Contribution, Draft Standard, Final Standard, or implementation, in whole or in part, to one or more of the licensing obligations listed in Section 9.3, such Participant shall give prompt notice of the same to the CDISC President who shall promptly notify all Participants.

No Other Warranties/Disclaimers. ALL PARTICIPANTS ACKNOWLEDGE THAT, EXCEPT AS PROVIDED UNDER SECTION 9.3 OF THE CDISC INTELLECTUAL PROPERTY POLICY, ALL DRAFT STANDARDS AND FINAL STANDARDS, AND ALL CONTRIBUTIONS TO FINAL STANDARDS AND DRAFT STANDARDS, ARE PROVIDED "AS IS" WITH NO WARRANTIES WHATSOEVER, WHETHER EXPRESS, IMPLIED, STATUTORY, OR OTHERWISE, AND THE PARTICIPANTS, REPRESENTATIVES, THE CDISC PRESIDENT, THE CDISC BOARD OF DIRECTORS, AND CDISC EXPRESSLY DISCLAIM ANY WARRANTY OF MERCHANTABILITY, NONINFRINGEMENT, FITNESS FOR ANY PARTICULAR OR INTENDED PURPOSE, OR ANY OTHER WARRANTY OTHERWISE ARISING OUT OF ANY PROPOSAL, FINAL STANDARDS OR DRAFT STANDARDS, OR CONTRIBUTION.

Limitation of Liability
IN NO EVENT WILL CDISC OR ANY OF ITS CONSTITUENT PARTS (INCLUDING, BUT NOT LIMITED TO, THE CDISC BOARD OF DIRECTORS, THE CDISC PRESIDENT, CDISC STAFF, AND CDISC MEMBERS) BE LIABLE TO ANY OTHER PERSON OR ENTITY FOR ANY LOSS OF PROFITS, LOSS OF USE, DIRECT, INDIRECT, INCIDENTAL, CONSEQUENTIAL, OR SPECIAL DAMAGES, WHETHER UNDER CONTRACT, TORT, WARRANTY, OR OTHERWISE, ARISING IN ANY WAY OUT OF THIS POLICY OR ANY RELATED AGREEMENT, WHETHER OR NOT SUCH PARTY HAD ADVANCE NOTICE OF THE POSSIBILITY OF SUCH DAMAGES.

Note: The CDISC Intellectual Property Policy can be found at
http://www.cdisc.org/system/files/all/article/application/pdf/cdisc_20ip_20policy_final.pdf

© 2026 Clinical Data Interchange Standards Consortium, Inc. All rights reserved