Unified Study Definitions Model Implementation Guide (USDM-IG)

Version 2.0 (Draft for Internal Review)

Prepared by the
DDF Team

Notes to Readers

- This is the draft version 2.0 of the Unified Study Definitions Model Implementation Guide (USDM-IG v2.0). It is intended for Internal Review only and is not a final version.

Revision History

Date	Version
2023-03-08	2.0 Draft for Internal Review

Contents

1 Introduction

CDISC, in collaboration with TransCelerate Biopharma and Accenture as a part of TransCelerate's Digital Data Flow Project, have developed a Study Definition Reference Architecture called the unified Study Definitions Model (USDM).

The aim of TransCelerate's DDF initiative is to optimize study start-up (SSU) processes and automate system configuration and readiness. The current state typically involves disconnected study design services and assets, and transcription or re-entry of the same information into many systems across Sponsors, Contract Research Organizations, and systems vendors. This inefficiency results in systems configuration falling onto the critical path for study start-up and adds risks for transcription errors and unnecessary delays.

Ideally, a solution would enable interoperability across multiple systems in a clinical study, improve efficiency and data quality, and reduce cycle times. That solution should capture protocol elements and present them in standardized formats to enable automated configuration of downstream systems and efficient consumption of protocol information across the study ecosystem.

The challenge is that study start-up system configuration workflow and asset creation is currently not automated, which makes it inefficient and increases the risk of error. Current workflows also include a number of redundant, manual activities. Sponsors are not able to utilize resources efficiently due to the siloed, document-based environment. Additional information can be found on the TransCelerate Digital Data Flow Solutions page.

The collaborative effort between TransCelerate, CDISC and Accenture has enabled the development of the USDM Reference Architecture in conjunction with development of a Study Definitions Repository (a reference Implementation of the USDM architecture. For more information on the Study Definitions Repository please visit the TransCelerate DDF GitHub site and the SDR Github Site.

1.1 Purpose

This USDM Implementation Guide is intended for companies and individuals involved in the set-up of clinical studies either as a Sponsor, or as stakeholders involved in upstream (protocol and content authoring tools) and downstream consumers of standardized digitized study definitions both systems (e.g., EDC, CTMS and TMF) and documents (e.g., Protocol, Clinical Study Reports, Statistical Analysis Plans)

This document provides users with sufficient information to understand the USDM and also it's potential implementations with the study design process by showing examples of the types of study definition information that can be represented in the USDM.

1.2 Organization of this Document

This document is divided into the following sections:

- Section 1, Introduction, provides an overall introduction to the purpose and goals of the USDM-IG
- Section 2, Fundamentals of the USDM, provides a boundary of the scope of this version of the USDM and what use cases this version was intended to support.
- Section 3, Relationships to Other Standards and Formats, describes at a high level how the USDM relates to other standards (both CDISC and Non-CDISC) and also to the TransCelerate Common Protocol Template
- Section 4, USDM Data Dictionary, provides a data dictionary listing of the classes, attributes and definitions of the USDM model.
- Section 5, Examples, provides examples of the types of information that can represented using the USDM including various study designs ranging in complexity.
- Section 6, USDM API, provides information on the USDM API
- Appendices provide additional background material and describe other supplemental material relevant to the USDM.

1.3 How to Read this Document

1. It is recommended to familiarise yourself with the Digital Data Flow project by reading resources from TransCelerate's Digital Data Flow Project and CDISC's Digital Data Flow information.
    a. If readers are new to Digital Data Flow it is recommended to watch the video presentation on the TransCelerate DDF video library
    b. Of particular interest will be the video on the USDM Overview
2. Next it is recommended to visit the DDF Internal Review page to download the Internal Review USDM review material from the CDISC DDF GitHub site. This page also provide additional information on the Internal Review process including how to make comments in the DDF JIRA.
    a. Note that this is a .zip file. The UML, CT and API deliverables can be located in the
3. Next, Read this guide all the way through (without skipping any sections) at least once.
    a. The USDM-IG compiled will provide you with a page by page view of the USDM-IG
    b. The USDM-IG sections will show each section of the USDM-IG - when making comments on the USDM-IG please use the USDM-IG sections as you will not be able to make comments on the compiled view.
4. Finally, revisit any sections of particular interest and provide your comments.

2 Fundamentals of the USDM

The Unified Study Definitions Model (USDM) standard comprises of 4 parts that are official CDISC standards. USDM v2.0 therefore consists of the following components:

1. **Unified Study Definitions Model (USDM) Class Diagram represented as a UML class diagram.**
2. **Application Programming Interface (API) Specification**
3. **CDISC Controlled Terminology**
4. **Unified Study Definitions Model Implementation Guide (USDM-IG)**

Please note that USDM v1.0 did not have a corresponding Implementation Guide. The USDM-IG is new for USDM v2.0 and therefore a USDM-IG v1.0 is not available.

Release of the Unified Study Definitions Model v1.0 in August 2022, provided a base model of structured study design. Building on this foundation, USDM v2.0 has been developed to satisfy an agreed set of use cases based around:

1. Updates to the USDM in order to enable greater population of study set-up elements and represent structured study design information for more complex trials
2. Updates to the USDM in order to support enablement of EDC automation
3. Updates to the USDM with selected structured eCPT data elements to demonstrate population of the TransCelerate Common Protocol Template (eCPT)

2.1 Support for more complex trials

V1.0 of the USDM provided a model for simple study designs. V2.0 has implemented additional elements that will be able to cater for more complex study designs to be represented in USDM format. One main area of development is the implementation of Study Timing within the model allowing for complex timing and visit structures to be represented. The USDM Features provides a overview of enhancements of the model to cater for increased trial complexity.

2.2 Support to enable EDC Automation

In order to support enablement of EDDC automation, the CDISC Biomedical Concept model was adapted and included as a sub-model in the USDM. The additional of biomedical concepts into the model will add in a machine readable "data" layer into the study design. This data layer can be used in a variety of ways to inform about what data relates to particular assessments within a study design. This biomedical concept model will not only assist in informing an EDC system as to the individual data items required for an assessment (e.g., automating identification of an form in an EDC library with the same/similar set of biomedical concepts) but also provide basic information required to build a new form should there be no EDC library, or a form that matches.
As Implementation of the biomedical concept model in the USDM also provides a machine readable data specification that can support other data source use cases such as digital health technologies, ePROs, and electronically supplied data (e.g., central lab or central ECG data).

2.3 Support to demonstrate population of selected structured CPT elements

Additional elements have been added to the model in order for a proof of viability exercise to be carried out to demonstrate that structured study design information can be moved from an upstream study design application, into USDM format and then used to populate the TransCelerate Common Protocol Template. Additional information on the USDM elements that are used for this POV can be found in the Use of USDM for Populating Protocol Content section. Note that only a selected set of CPT elements was chosen to be included for the POV and additional elements may be added to the USDM in a future release.

3 Relationships to Other Standards and Formats

The USDM covers a wide range of concepts related to study design that also appear in other published standards. As part of the development process these standards were used as input in order to try to ensure that, where possible, harmonization with these standards.

3.1 Relationship to Other CDISC Standards

The USDM development process relied on published CDISC standards and other products that served as references for modeling and naming conventions. To the extent possible, an effort was made to align or be compatible with these sources where the content was determined to be conceptually identical or closely related to those being developed for the USDM.

3.1.1 BRIDG

Biomedical Research Integrated Domain Group (BRIDG) is a CDISC, **HL7** and **ISO** standard for biomedical research concepts designed to support computable semantic interoperability. BRIDG can be used for various purposes: as a reference model, a data integration/mapping solution, an exchange format, an ontology and to create a BRIDG-based database. The use of BRIDG helps support the meaningful exchange of data between software systems and databases.
When BRIDG is used as a reference model to create or add new content to a standard, it can help ensure that relationships between and among biomedical research concepts represented using the standard are consistently modeled.

3.1.2 PRM

Protocol Representation Model (PRM) provides a standard for planning and designing a research protocol with focus on study characteristics such as study design, eligibility criteria, and requirements from the ClinicalTrials.gov, World Health Organization (WHO) registries, and EudraCT registries. PRM assists in automating CRF creation and EHR configuration to support clinical research and data sharing.
PRM was released in 2012 and although the standard includes some overlap with the USDM it is anticipated that the USDM will develop to be more content rich and implementable as a model and will therefore supersede the PRM model

3.1.3 SDTM and SDTM-IG

The Study Data Tabulation Model (SDTM) provides a standard for organizing and formatting data to streamline processes in collection, management, analysis and reporting. Implementing SDTM supports data aggregation and warehousing; fosters mining and reuse; facilitates sharing; helps perform due diligence and other important data review activities; and improves the regulatory review and approval process. Where the SDTM provides a standard model for organizing and formatting data for human and animal studies, the Study Data Tabulation Model Implementation Guide SDTMIG is intended to guide the organization, structure, and format of standard clinical trial tabulation datasets. It was developed to support data submitted to a regulatory authority, such as the US Food and Drug Administration (FDA), but is not restricted to use in regulated submissions. SDTM is one of the required standards that sponsors must use as specified in the FDA's Data Standards Catalog (see section II.C) for NDA, ANDA, and certain BLA submissions.
The SDTM-IG contains a section related to Trial Design Model Datasets. ICH E3, Guidance for Industry, Structure and Content of Clinical Study Reports (available at http://www.ich.org/products/guidelines/efficacy/article/efficacy-guidelines.html), Section 9.1, calls for a brief, clear description of the overall plan and design of the study, and supplies examples of charts and diagrams for this purpose in Annex IIIa and Annex IIIb. Each Annex corresponds to an example trial, and each shows a diagram describing the study design and a table showing the schedule of assessments. The Trial Design Model provides a standardized way to describe those aspects of the planned conduct

of a clinical trial shown in the study design diagrams of these examples. The standard Trial Design Datasets will allow reviewers to:

- Clearly and quickly grasp the design of a clinical trial
- Compare the designs of different trials
- Search a data warehouse for clinical trials with certain features
- Compare planned and actual treatments and visits for subjects in a clinical trial

Modeling a clinical trial in this standardized way requires the explicit statement of certain decision rules that may not be addressed or may be vague or ambiguous in the usual prose protocol document. Prospective modeling of the design of a clinical trial should lead to a clearer, better protocol. Retrospective modeling of the design of a clinical trial should ensure a clear description of how the trial protocol was interpreted by the sponsor.

Trial design concepts include:

- Trial Design
- Epoch
- Arm
- Study Cell
- Element
- Branch
- Treatments
- Visit

Although not a current use case for USDM v2.0, automated creation of the SDTM Trial Design Datasets may in the future be possible using data structed in USDM format. Therefore there is alignment between the USDM and the SDTM Trial Design and controlled terminology elements related to study design. The table below provides a list of published Trial Summary (TS) Parameters and their mapping to the USDM elements (Entities, Attributes, or Valid Values). The table includes only those parameters for which there is a mapping. The table is based on the SDTM Controlled Terminology codelist C66738, from SDTM Terminology Version 2022-12-16.

Code	Codelist Code	Codelist Extensible (Yes/No)	Codelist Name	CDISC Submission Value	CDISC Synonym(s)	CDISC Definition	NCI Preferred Term	USDM Entity Name	USDM Role	USDM Item Name
C101302	C66738		Trial Summary Parameter Test Code	THERAREA	Therapeutic Area	A knowledge field that focuses on research and development of specific treatments for diseases and pathologic findings, as well as prevention of conditions that negatively impact the health of an individual. (NCI)	Therapeutic Area	StudyDesign	Attribute	therapeuticAreas
C112038	C66738		Trial Summary Parameter Test Code	INDIC	Trial Disease/Condition Indication; Trial Disease/Condition Indication Description	The textual representation of the condition, disease or disorder that the clinical trial is intended to investigate or address.	Trial Indication	Indication	Entity	Indication
C112038	C66738		Trial Summary Parameter Test Code	INDIC	Trial Disease/Condition Indication; Trial Disease/Condition Indication Description	The textual representation of the condition, disease or disorder that the clinical trial is intended to investigate or address.	Trial Indication	Indication	Attribute	indicationDescription
C142175	C66738		Trial Summary Parameter Test Code	STYPE	Study Type; Study Type Classification	The nature of the investigation for which study information is being collected. (After clinicaltrials.gov)	Study Type	Study	Attribute	studyType
C48281	C66738		Trial Summary Parameter Test Code	TPHASE	Trial Phase; Trial Phase Classification	A step in the clinical research and development of a therapy from initial clinical trials to post-approval studies. NOTE: Clinical trials are generally categorized into four (sometimes five) phases. A therapeutic intervention may be evaluated in two or more phases simultaneously in different trials, and some trials may overlap two different phases. [21 CFR section 312.21; After ICH Topic E8 NOTE FOR GUIDANCE ON GENERAL CONSIDERATIONS FOR CLINICAL TRIALS, CPMP/ICH/291/95 March 1998]	Trial Phase	Study	Attribute	studyPhase
C49652	C66738		Trial Summary Parameter Test Code	TINDTP	Trial Intent Type	The planned purpose of the therapy, device, or agent under study in the clinical trial.	Clinical Study by Intent	StudyDesign	Attribute	trialIntentType
C49658	C66738		Trial Summary Parameter Test Code	TBLIND	Study Blinding Design; Study Blinding Schema; Study Masking Design; Trial Blinding Design; Trial Blinding Schema; Trial Masking Design	The type of experimental design used to describe the level of awareness of the study subjects and/ or study personnel as it relates to the respective intervention(s) or assessments being observed, received or administered.	Trial Blinding Schema	StudyDesign	Attribute	studyDesignBlindingScheme
C49660	C66738		Trial Summary Parameter Test Code	TTYPE	Trial Scope; Trial Type	The nature of the interventional study for which information is being collected.	Trial Type	StudyDesign	Attribute	trialType
C49692	C66738		Trial Summary Parameter Test Code	PLANSUB	Anticipated Enrollment; Planned Enrollment; Planned	The planned number of subjects to be entered in a clinical trial. (NCI)	Planned Subject Number	StudyDesignPopulation	Attribute	plannedNumberOfParticipants

Code	Codelist Code	Codelist Extensible (Yes/No)	Codelist Name	CDISC Submission Value	CDISC Synonym(s)	CDISC Definition	NCI Preferred Term	USDM Entity Name	USDM Role	USDM Item Name
					Number of Subjects; Target Enrollment					
C49693	C66738		Trial Summary Parameter Test Code	AGEMIN	Planned Minimum Age of Subjects	The anticipated minimum age of the subjects to be entered in a clinical trial. (NCI)	Planned Minimum Age of Subjects	StudyDesignPopulation	Attribute	plannedMinimumAgeOfParticipants
C49694	C66738		Trial Summary Parameter Test Code	AGEMAX	Planned Maximum Age of Subjects	The anticipated maximum age of the subjects to be entered in a clinical trial. (NCI)	Planned Maximum Age of Subjects	StudyDesignPopulation	Attribute	plannedMaximumAgeOfParticipants
C49696	C66738		Trial Summary Parameter Test Code	SEXPOP	Sex of Participants	The specific sex, either male, female, or mixed of the subject group being studied. (NCI)	Sex of Study Group	StudyDesignPopulation	Attribute	plannedSexOfParticipants
C49802	C66738		Trial Summary Parameter Test Code	TITLE	Official Study Title; Study Title; Trial Title	The sponsor-defined name of the clinical study.	Trial Title	Study	Attribute	studyTitle
C98746	C66738		Trial Summary Parameter Test Code	INTMODEL	Intervention Model	The general design of the strategy for assigning interventions to participants in a clinical study. (clinicaltrials.gov)	Intervention Model	StudyDesign	Attribute	interventionModel
C70793	C66738		Trial Summary Parameter Test Code	SPONSOR	Clinical Study Sponsor; Sponsor; Study Sponsor	An individual, company, institution, or organization that takes responsibility for the initiation, management, and/or financing of a clinical study. [After ICH E6, WHO, 21 CFR 50.3 (e), and after IDMP]	Clinical Study Sponsor	Organization	Valid Value	Valid Value Set for Attribute organizationType
C85826	C66738		Trial Summary Parameter Test Code	OBJPRIM	Study Primary Objective; Trial Primary Objective	A principle objective of the study.	Trial Primary Objective	Objective	Valid Value	Valid Value Set for AttributeobjectiveLevel
C85827	C66738		Trial Summary Parameter Test Code	OBJSEC	Study Secondary Objective; Trial Secondary Objective	An auxiliary objective of the study.	Trial Secondary Objective	Objective	Valid Value	Valid Value Set for AttributeobjectiveLevel

3.1.4 Controlled Terminology

CDISC, in collaboration with the National Cancer Institute's Enterprise Vocabulary Services (EVS), supports the Controlled Terminology (CT) needs of the CDISC Standards. Controlled terminology is the set of codelists, definitions, and valid values used with CDISC model elements. Within CDISC there are many volunteer teams that evaluate and manage CDISC CT. For instance, the Protocol Entities team has been developing and publishing the semantics for those concepts found in clinical research protocols and the CDISC Glossary team harmonizes the semantics and definitions for concepts commonly found in CDISC standards documents. The DDF Terminology subset of CDISC CT is one of the main deliverables supporting the USDM, and development of CDISC CT for the USDM has been harmonized with existing, published CDISC CT (including SDTM, Protocol, and CDISC Glossary) in order to ensure maximum reuse of terms and definitions. Any new CT that has been developed for the USDM has undergone review from the Protocol Entities and CDISC Glossary teams. USDM related CT is developed and published using the same process as all other CDISC CT, in order to ensure a consensus based, fit for use, and harmonized set of terms.

3.1.5 CTR

Clinical Trial Registry (CTR)-XML lets technology vendors implement tools that support a “write once, use many times” solution based on a single XML file that holds the information needed to generate submissions for multiple clinical trials for clinical trial registry submissions primarily to the World Health Organization (WHO),European Medicines Agency (EMA) EudraCT Registry and United States ClinicalTrials.gov.

Although not a current use case for USDM v2.0, automated submissions for multiple clinical trials for clinical trial registry submissions may in the future be possible using data structed in USDM format. CTR was released in 2016 and although the standard includes some overlap with the USDM it is anticipated that the USDM will develop to be more content rich and implementable as a model and will therefore supersede the CTR model.

3.1.6 ODM

ODM-XML is a vendor-neutral, platform-independent format for exchanging and archiving clinical and translational research data, along with their associated metadata, administrative data, reference data, and audit information. ODM-XML facilitates the regulatory-compliant acquisition, archival and exchange of metadata and data. It has become the language of choice for representing case report form content in many electronic data capture (EDC) tools.

The current version of ODM (ODM-XML v1.3.2) was released in 2013. ODM-XML v2.0 is currently in development and adds significant functionality to the ODM standard. These additions include:

- Multi-lingual Support
- Data Query Support
- Traceability (Trace-XML Features) Support
- HL7 FHIR Interoperability
- Study Design Model (SDM-XML) integration and enhancement
- CDISC 360 Support
- Data Capture

Although USDM is a reference model and ODM is a transport model there is overlap between the USDM and ODM v2.0 standard in terms of the elements related to study design in terms of biomedical concepts and those elements related to EDC build (visits, forms, variables etc). Therefore during the development of the USDM, areas of development for ODM v2.0 were investigated and where possible aligned with USDM.

3.1.7 SDM

Study/Trial Design Model in XML (SDM-XML) is an extension of ODM-XML and allows organizations to provide rigorous, machine-readable, interchangeable descriptions of the designs of their clinical studies, including treatment plans, eligibility and times and events. SDM-XML defines three key sub-modules – Structure, Workflow, and Timing – permitting various levels of detail in any representation of a clinical study's design.
The current version of SDM (v1.0) was released in 2011. Note that SDM has now been incorporated into ODM-XML v2.0 (still in development). SDM was used as an input reference model during the development of the USDM.

3.2 Relationship to Other Standards

3.2.1 ICH M11 guideline, clinical study protocol template and technical specifications

The purpose of this new harmonised guideline is to introduce the clinical protocol template (Clinical Electronic Structured Harmonised Protocol (CeSHarP)) and the technical specification to ensure that protocols are prepared in a consistent fashion and provided in a harmonised data exchange format acceptable to the regulatory authorities. At the time of scoping of for USDM v2.0, the content of the guideline was not public ally available and therefore cold not be included as a scoping input for this version.
The guideline, clinical study protocol template and technical specifications were released in October 2022 for public review and was used a reference input during the development phase where possible. Due to the highly aligned domain of the USDM and M11 it is anticipated that there will be additional alignment activates in future versions of the USDM.

3.2.2 HL7 FHIR SOA

The Vulcan Schedule of Activities (SOA) Project entails a research protocol's Schedule of Activities using FHIR Resources that can be consumed by the EHR and corresponding Electronic Data Capture (EDC) system to support the research workflow and data exchange. When a Patient is enrolled in a study, the research personnel will attach the Patient to the ResearchSubject and ResearchStudy thus connecting the CarePlan with the Schedule of Activities (the research visits and corresponding tests (activities).
For USDM V2.0 elements related a schedule of activities was further developed in order to provide structured information that would enable creation of a visual representation of the this information in an SOA format. Therefore there are important connections between the USDM and the Vulcan SOA project. As a result, there is an ongoing collaboration between the DDF and the Vulcan SOA project to ensure alignment where possible.

3.3 Use of USDM for Populating Protocol Content

A secondary aim of USDM v2.0 is to demonstrate that protocol-related content can be pulled from a reference implementation of the USDM and populated programmatically into the corresponding fields of a structured document, such as TransCelerate's Common Protocol Template (CPT). A successful demonstration is anticipated to facilitate expanding future versions of the USDM for this purpose. The TransCelerate CPT was selected to conduct this proof of concept since it is a publicly available resource proposed to harmonize clinical trial protocol content in a streamlined format (https://www.transceleratebiopharmainc.com/assets/clinical-content-reuse-solutions/). The POC exercise relies on a prioritized set of structured fields within the CPT for content already existing in USDM v1.0 and extended in USDM v2.0. For example, the table below shows a selection of structured CPT field names mapped to the USDM v2.0 model which are used in the proof of concept.

CPT Section	CPT Variable Display Name	CPT Variable Name (compacted)	CPT Var Type	Mapping Type	USDM Field	USDM Field Type	Logic
Synopsis	Number of Participants	CPT:NumberofParticipants	Text	ManyToOne	plannedNumberOfParticipants	Integer	If multiple populations available in studyDesign, add all the numeric values.
Study Rationale	Study Rationale	CPT:StudyRationale	Rich Text	OneToOne	studyRationale	Text	Retrieve studyRationale value
Objectives and Endpoints	Objectives Endpoints and Estimands	CPT:ObjectivesEndpointsAndEstimands	RichText	OneToMany	objectiveDesc, endpointDesc	Text	See below fields.
Objectives and Endpoints	Primary Endpoints	CPT:EndpointsPrimary	RichText	OneToMany	endpointDesc | endpointPurposeDesc	Text	Take respective ObjectiveEndpoints from primary objective
Objectives and Endpoints	Primary Objectives	CPT:ObjectivesPrimary	RichText	OneToMany	objectiveLevel	Code	Take the objective whch has 'Study Primary Objective' in the objectiveLevel as in screenshot below. Refer CDISC codes mentioned in Data Mapping spreadsheet.
Objectives and Endpoints	Secondary Endpoints	CPT:EndpointsSecondary	RichText	OneToMany	endpointDesc | endpointPurposeDesc	Text	Take respective ObjectiveEndpoints from secondary objective
Objectives and Endpoints	Secondary Objectives	CPT:ObjectivesSecondary	RichText	OneToMany	objectiveLevel	Code	Take the objective whch has 'Study Secondary Objective' in the objectiveLevel as in screenshot below. Refer CDISC codes mentioned in Data Mapping spreadsheet.
Scientific Rationale for Study Design	Scientific Rationale for Study Design	CPT:ScientificRationaleforStudyDesign	RichText	OneToOne	studyDesignRationale	Text	Retrieve studyDesignRationale value
Inclusion Criteria	Planned Maximum Age of Subjects	CPT:PlannedMaximumAgeofSubjects	Text	ManyToOne	plannedMaximumAgeOfParticipants	Text	1. If all values are integers, then pick Maximum value from the list. 2. If multiple values available, atleast one non-integer value is present, then display blank in the output. 3. If only one value available, irrespective of Integer/Non-Integer, display the value as is in the output.
Inclusion Criteria	Planned Minimum Age of Subjects	CPT:PlannedMinimumAgeofSubjects	Text	ManyToOne	plannedMinimumAgeOfParticipants	Text	1. If all values are integers, then pick Minimum value from the list. 2. If multiple values available, atleast one non-integer value is present, then display blank in the output. 3. If only one value available, irrespective of Integer/Non-Integer, display the value as is in the output.
Inclusion Criteria	Sex of participants	CPT:Sexofparticipants	Choice	vs.CodeList<>	plannedSexOfParticipants	Code[]	1. Refer to CDISC code list for Sex and corresponding eCPT mapping values in Data mapping sheet 2. If multiple values available, consider distinct values from the valid codes and display Male/Female/Male or Female in the output 3. If only value is available and not a valid C-DISC code, display decode value as is in the CPT output. 4. If multiple and all of the codes are invalid, then display blank.
Study Interventions Administered	Arm Description	CPT:ArmDescription	RichText	OneToOne ManyToOne	studyArmDesc	Text	studyArmDescription, ArmName and Decode Value of ArmType to be sent as an arrayList in response.
Study Interventions Administered	Arm Name	CPT:ArmName	RichText	OneToOne	studyArmName	Text	studyArmDescription, ArmName and Decode Value of ArmType to be sent as an arrayList in response.

CPT Section	CPT Variable Display Name	CPT Variable Name (compacted)	CPT Var Type	Mapping Type	USDM Field	USDM Field Type	Logic
Study Interventions Administered	Arm Type	CPT:ArmType	RichText	OneToOne	studyArmType	Code	studyArmDescription, ArmName and Decode Value of ArmType to be sent as an arrayList in response.
Study Interventions Administered	Intervention Description	CPT:InterventionDescription	RichText	OneToOne	interventionDesc	Code[]	Create model as mentioned in screenshot and just populate interventionDescription for now. There are other fields like intervention name, type, dosage etc.. Which are not available in USDM.
Populations for Analyses	Populations for Analyses	CPT:PopulationsForAnalyses	RichText	ManyToOne	populationDesc (analysisPopulation)	Text	Retrieve all analysisPopulationDescription as comma seperated (e.g Desc1, Desc2 and Desc3)
Page Header	Version Number	CPT:VersionNumber	Text	OneToMany	protocolVersion	Text, text	protocolVersion sort by EffectiveDate and Version
Protocol and Brief Title	Condition or Disease	CPT:ConditionDisease	Text	Proxy	indicationDesc	Text	Retrieve all indicationDescriptions as comma seperated (e.g Desc1, Desc2 and Desc3)
Title Page	Acronym	CPT:Acronym	Text	OneToOne	studyAcronym	Text	Retireve studyAcronym value
Title Page	Amendment Number	CPT:AmendmentNumber	Text	Proxy	protocolAmendment	Text	protocolAmendment sort by EffectiveDate and Version
Title Page	Protocol Short Title	CPT:ProtocolShortTitle	RichText	OneToOne	briefTitle	Text	briefTitle sort by EffectiveDate and Version
Title Page	Protocol Title	CPT:ProtocolTitle	RichText	OneToMany	studyTitle (else scientificTitle)	Text	studyTitle if available else pick from scientificTitle
Title Page	Regulatory Agency ID	CPT:RegulatoryAgencyID	Choice	vs.CodeList<>	organisationIdentifierScheme	Code	Retrieve organisationIdentifierScheme where Type = 'Regulatory Agency' (First element to be considered if multiple array elements)
Title Page	Regulatory Agency Number	CPT:RegulatoryAgencyNumber	Text	OneToMany	studyIdentifier	Text, text	Retrieve studyIdentifier where Type = 'Regulatory Agency' (First element to be considered if multiple array elements)
Title Page	Sponsor Legal Address	CPT:SponsorLegalAddress	Text	OneToOne	organizationLegalAddress	Text	To be retrieved from Organization class (attribute name of organizationLegalAddress, where Organization Type=Clinical Study Sponsor) and concatenate all Address properties Take First value if there are more than one.
Title Page	Sponsor Name	CPT:SponsorName	Text	OneToOne	organizationName	Text	To be retrieved from Organization class (attribute name of OrganizationName, where Organization Type=Clinical Study Sponsor)
Title Page	Study Phase	CPT:StudyPhase	Choice	vs.CodeList<>	studyPhase	aliasCode	Retrive decode Value from standardCode element. Transform into CPT master code value

4 USDM Features

4.1 Overview

The USDM provides the ability to define a version of a clinical study that includes:

1. The main study details including:
    - The version of the external protocol that the study relates to
    - The various identifiers allocated to the study
2. One or more study designs within the study with each study design detailing:
    - the Arms and Epochs within the design and the relationships between them
    - The encounters planned for the study and the relationship with the epochs of the study
    - A detailed data specification for the data to be captured as part of the study
    - The procedures to be performed as part of the study design
    - The timing around the collection of data and the performance of procedures
    - The subject populations defined within the study design
    - The objectives & endpoints defined within the study design
    - The study estimands defined within the study design
    - The interventions defined as part of the study design
    - The relevant indication

The USDM is designed to hold a single version of a study but the model can be used to implement systems that hold multiple versions of multiple studies.

4.2 Internal Identifiers Within the Model

The USDM normative form is a UML model. Each class defined within the UML has as identification attribute that can be used by an implementation to provide a unique identifier for an instance of the class. The identifier should be unique and self-consistent within the scope of a version of a study. No attempt is made to define the form, type or structure of these identifiers, the attributes are defined as strings.

The identifiers are important in that one of the main uses of the USDM has been to define the API for the Study Definitions Repository (SDR) implementation. This API is designed to transport a single study in its entirety. An issue arises as, within this large structure, the same instance may have relationships from several other instances. As such the content could be included (duplicated) at several places within the API (formatted as JSON) structure. So as not to repeat the same information within the JSON structure, the API has been designed to include an instance once

and only once and allow for zero, one or more references to it as dictated by the USDM model and the relationships within. This mechanism relies on the unique identifiers.
The location of where instances will be included within the API structure and where they will be referenced is specified within the UML. The location where instances will be included is indicated by an attribute's type being the type of the class. Where an instance is referenced is indicated by the type of the attribute being "string" and the attribute name suffixed with "Id".
For example, for the Encounter class, all instances are included from the StudyDesign class using the attribute
`encounters List<Encounter>`
whereas the StudyEpoch references the instances using the attribute
`encounterIds List<string>`
The only exception is the identifier at the head of the model within the Study class. Implementations are free to allocate the value to this field using, for example, a UUID, to ensure uniqueness within the implementation.

4.3 Controlled Terminology

Controlled terminology is referenced in multiple places across the USDM. So as to provide a mechanism to refer to controlled terms in a consistent manner, the USDM employs the Code class. The Code class uses four attributes to define the term being used, the terminology from which the term is taken and the version of that terminology. This allows for a any controlled term, be it a SNOMED, LOINC, CDSIC or some other term, to be referred to in a consistent manner.
Certain attributes within the USDM class have been constrained to using terms from a given code list from specified terminologies and these are specified within the Controlled Terminology spreadsheet. Most of terms referenced are CDISC Controlled Terms but some other controlled vocabularies are referenced.
Where a CDISC code is demanded by the model but flexibility is needed, flexibility is given to users to include other terms (aliases), using the AliasCode class. Here one standard term is required but zero, one or more aliases can be provided.

4.4 Study, Study Versions and Identifiers

The Study class is the root of the USDM collecting together the definition of the study as a whole. It provides a few basic study details such as the study title, type, phase, rationale and acronym, links the study with its constituent parts that include one or more study designs, the identifiers for the study and the relationship with external protocol documents.
The Study class also allows for the Business Therapeutic Area to be stated. This is provided for downstream processes and for sponsor organisations to define the business area within the enterprise handling the study. It should be noted that this is not the same as the therapeutic areas defined within the StudyDesign class.
The Study class links to the StudyProtocolVersion class to define which versions of an external protocol document this study definition relates. Because the traditional paper / PDF protocol document has been split into two parts, the document and an electronic design using the USDM, there is a need to link which electronic definition is valid with which version of the document.
The Study class allows for links to the one or more identifiers related to the study. Multiple identifiers are permitted but they must be of one of three types: Sponsor, Registry or Regulatory Authority. The Study definition should have one, and only one, sponsor identifier but multiple other identifiers are permitted. Note the use of the ISO 3166-1 terminology for the country codes within the address field.
The Study class allows for one or more study designs to be included. This provides a single mechanism for master and umbrella studies. Typically, there would be a one to one relationship between study and study design.

4.5 Study Design

The StudyDesign class in the container for a single design within a study definition. It provides the slots for key parameters such as the trail type, trial intent type, blinding scheme and intervention model. The class also provides a place to store one or more codes defining the therapeutic area to which the study design relates. No controlled terminology is provided for the population of this field but the following controlled vocabularies are available for users to populate these fields. A sponsor's own controlled terms can also be used.

Dictionary / Terminology Name	URL
EUDRACT	https://eudract.ema.europa.eu/docs/technical/EUDRACT_Eutct_Pick_Lists_and_coded_values_v1_0.xls
ICD-10	https://www.icd10data.com/ICD10CM/Codes
MEDDRA	https://www.meddra.org/
MeSH	https://www.ncbi.nlm.nih.gov/mesh/
NCI Thesaurus	https://ncit.nci.nih.gov/ncitbrowser/
SNOMEDCT	https://www.nlm.nih.gov/healthit/snomedct/index.html
US FDA	https://www.fda.gov/drugs/development-resources/spectrum-diseasesconditions

4.6 Arms and Epochs

The high level study design consisting of the Arms and Epochs is defined using the StudyArm, StudyEpoch, StudyCell and StudyElement classes. The manner in which the classes are used follows the CDISC SDTM model. Epochs are also related to the Study Encounters (a more generic term for visits).
StudyElements and Encounters have entry and exit rules that are defined using the TransitionRule class. It should be noted that while theStudyElements and Encounter classes share the use of the TransitionRule class it is not expected that the instances within any study design will overlap, they are, most likely, distinct sets.
Given that the use of the classes is based on the SDTM model, the information within these classes can be used to populate parts of the SDTM Trial Design domains.

4.7 Activities

Activities are the means by which the procedures to be performed and the data to be captured are specified at a detailed level. The Activity class is used to group together data capture and procedures. The composition of these groupings is left to those designing studies. The Activity class can be linked to one or more procedures, one or more Biomedical Concepts, one or more groups of Biomedical Concepts and / or one of more surrogate Biomedical Concepts. Activities can be reused across multiple points in a study timeline.

4.8 Procedures

The procedures linked to the Activity class are simple text descriptions of the procedure required. Procedures can be optional with a text representation for the condition being provided.

4.9 Biomedical Concepts

A Biomedical Concept (BC) is a small model that defines a clinical concept in a standardised and reusable manner; it is a specification focused on the data and not how the data are captured or processed. As such, BCs are atomic entities and should not be split apart; to do so causes a loss of meaning. A BC is identifiable (has an identifier) and is complete (contains everything needed to use it).

A BC defines an observation but it requires context, the context of a clinical study. This is why, in the USDM, the BCs are linked to activities and thus the remainder of a study design.

Within the USDM the CDISC BC model has been represented in a manner consistent with the rest of the USDM itself. For example, the controlled terminology references use the Code object to be compatible with all of the CT references across the USDM. Additional attributes have also been added to allow for configuration as part of a study to enable or disable certain qualifiers or to constrain terminology responses to match the needs of a study, for example, constraining units to metric values.

One item to note that is important. Constraints can be applied to the content placed into the USDM but when those constraints are applied is not specified. A protocol may leave everything in the BCs unconstrained and only when deployment in capture systems happens will those constraints be applied.

The USDM allows for the inclusion of a single BC (e.g. Heart Rate), a collection of BCs (e.g. Vital Signs pre-configured to include say Height, Weight, Heart Rate and other tests) or surrogate BCs. Surrogate BCs are a placeholder mechanism for when a BC definition is not available. This allows the name of a test to be specified but no further detail need be provided. Surrogates can contain a name and description pair for the concept required. A reference field is also provided to allow for links to any reference materials, eg a URL to an external resource could be used.

A single BC uses the BiomedicalConcept class as its root instance connected to one or more BiomedicalConceptProperty instances to define the various properties of the BC, either the result value, units, qualifiers etc. Some of the property nodes will require controlled terminology references and these are placed within ResponseCode instances which then onward refer to a Code instance holding the actual term reference.

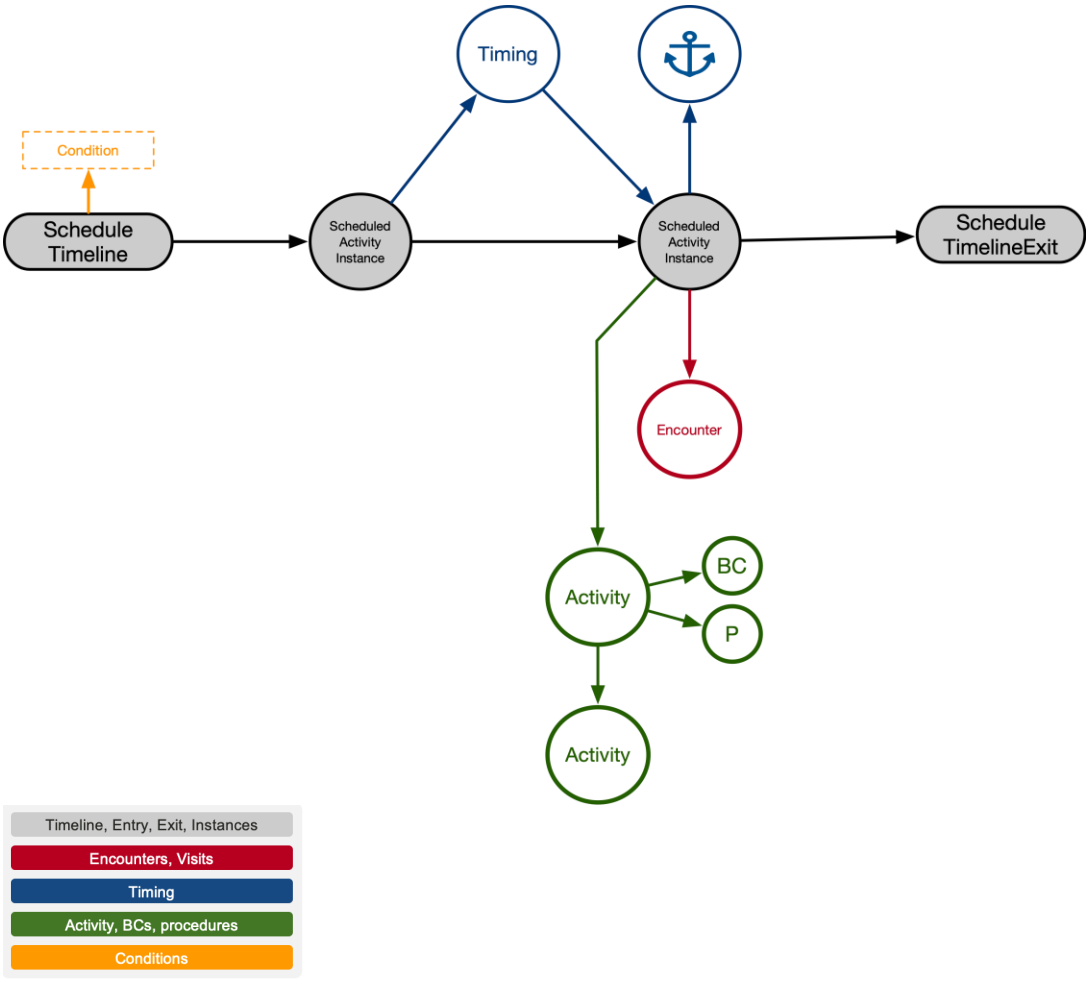
Such BCs can then be grouped using a BiomedicalConceptCategory instance which groups together 1 or more BCs. It is assumed that, to be useful, more than a single BC should be added to a grouping such as vital signs described above. These groupings are expected to be sponsor defined but, in the future, we would expect some to be industry defined.

4.10 Study Timing

One of the key aspects of a study design is the timing of encounters (visits) and the activities to be performed within those encounters. USDM 2 has included an update that replaces the workflow mechanism used in USDM 1 that linked encounters with activities with a mechanism for building timelines that can be reused within a study and, given external library management, across studies.

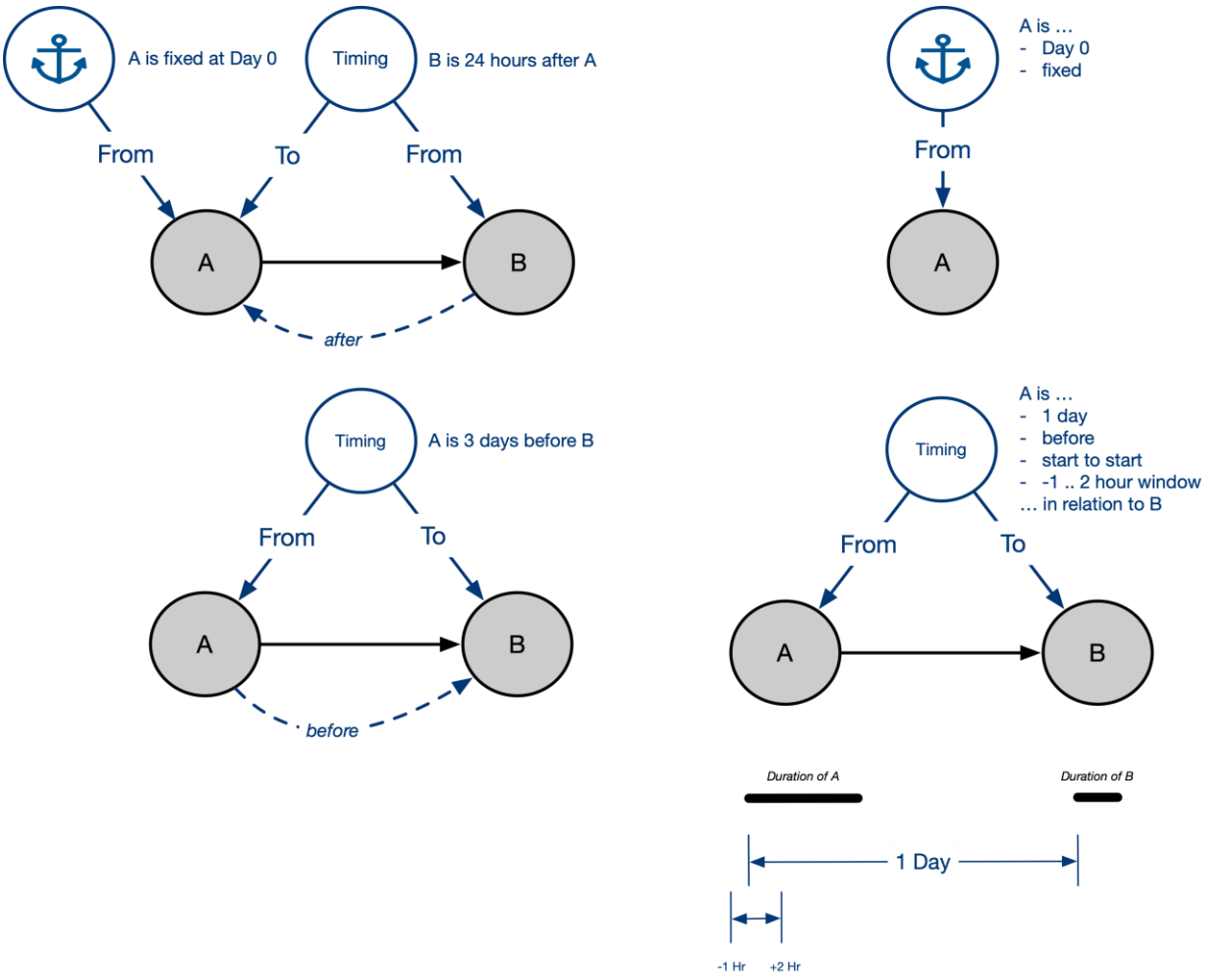
Timelines

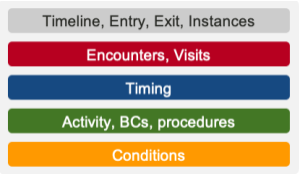
The study timing mechanism, shown in the figure below, is based around the notion of a timeline. A timeline is composed of an entry point with an associated condition (ScheduledTimeline class), a sequence of steps (the ScheduledActivityInstance class), possible branches to allow for multiple paths and cycles (not shown in the figure), Timing relating the steps (the Timing class), and one or more exits (the ScheduleTimelineExit class). A timeline is named and can be referenced / reused within other timelines. The steps within a timeline link the encounters with the activities required for each step and thus defines the timing for the encounters. The ScheduledActivityInstance class is the link between the high level study design defined by the StudyArms and StudyEpochs classes, the Encounter classes and the detailed study design defined by the Activity class.



Timing

The timing between steps comes in two flavours, a relative time of before or after, and an anchor time that is fixed. The figure below illustrates the timing capabilities.

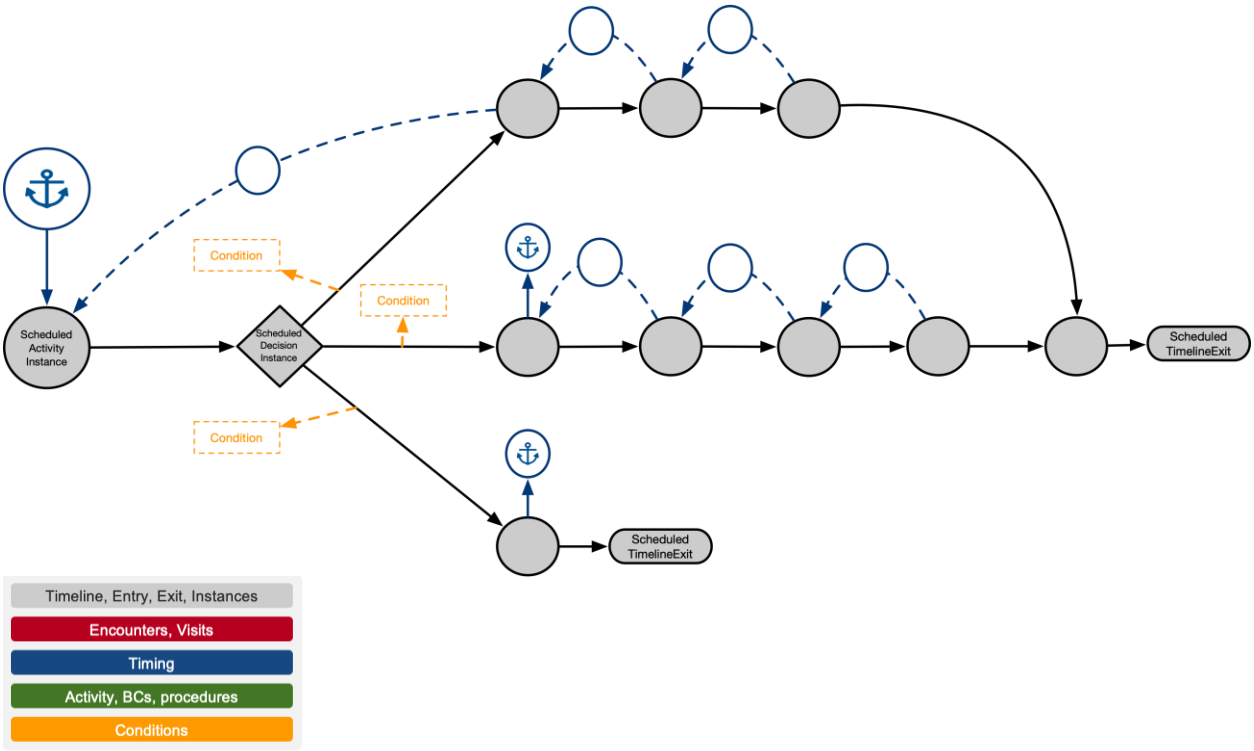




The timing class allows for explicit timing to be built into a timeline using a combination of anchors (fixed timing) and relative timing.

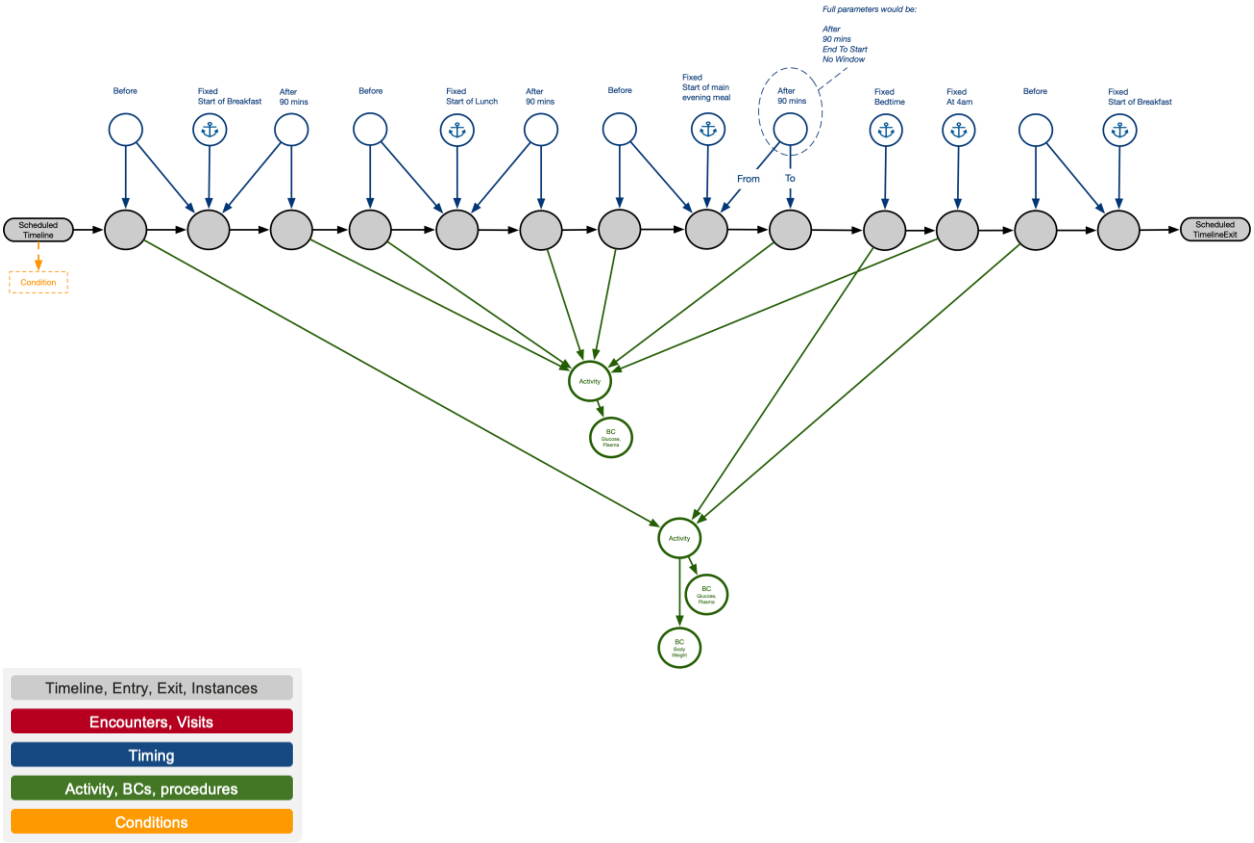
Decisions and Branching

Decisions and branching is handled using the ScheduledDecisionInstance class and using instances of the class within a timeline as shown in the figure below. Each decision point can handle multiple conditions such that simple yes / no decisions can be handled as well as a complex switch with multiple paths. Each possible route is setup with an associated destination. For switches there should be a "default" condition specified for the case when none of the other conditions are satisfied. The decision can also be used to create cycles.



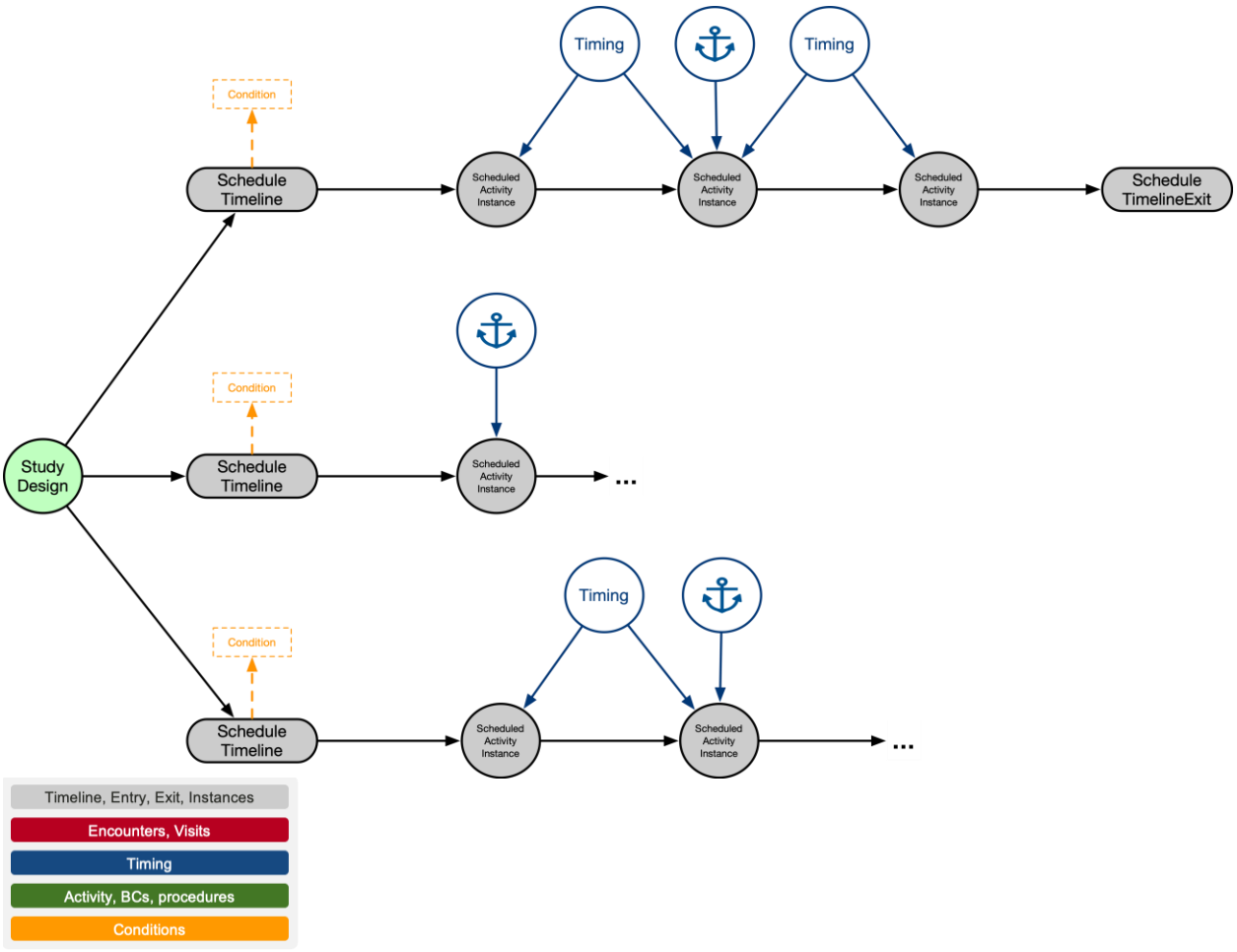
Profiles

Profiles can be created using the various classes as depicted within the following figure. A profile is another use of the timeline pattern. A condition for entry can be defined but need not be. Within the example depicted below, anchors are used to fix the meal times over a single day and the associated observations scheduled in relation to the fixed mealtimes. The activities are shared across the steps within the profile. The profile can be "attached" to an Activity such that it is executed as part of that Activity or can be attached to a ScheduledActivityInstance for execution from another timeline.



Unscheduled

Unscheduled visits within a study are handled by creating separate timelines for each unscheduled "event" that needs to be handled within the study design. A study design would typically have one "main" timeline with a condition such as "subject identified". Further timelines can be created and linked to the StudyDesign instance with the timeline having an appropriate condition, for example, "Adverse event" or "Lost contact with subject". Each timeline is then free to detail the steps taken under the respective circumstances.



4.11 Indications

The indication for a study design can be placed into the Indication class. Each indication has a textual description plus the ability to define one or more codes from external code systems (including a sponsor's own terminology) that define the indication.

4.12 Study Estimands

Study Estimands and the definition of the treatments to be investigated, the population, the variable and the handling of Intercurrent Events (ICEs) are handled within the Estimand, IntercurrentEvent and AnalysisPopulation classes along with the relationships to Endpoints (for the variable of interest) and InvestigationalIndications (for the treatment)

4.13 Investigational Interventions

The interventions for a study design can be placed into the InvestigationalIntervention class. Each intervention has a textual description plus the ability to define one or more codes from external code systems (including a sponsor's own terminology) that define the intervention.

4.14 Study Objectives and Endpoints

The study design objectives and endpoints can be defined within the Objective class and the Endpoint class. The objective class allows for the textual description of the objective and its level (primary, secondary etc) and a link to one or more associated endpoints containing a the endpoint definition in textual form.

4.15 Study Populations

The USDM currently implements a mechanism to define the subject populations for a study design using the StudyDesignPopulation class. The population definition consists of a text description plus a set of properties related to the age and sex of the population.

5 USDM Data Dictionary

Any property without a description in the table below, are either relationships or instance identifiers and were deemed to be out of scope for terminology development. Please see Internal Identifiers Within the Model for additional information on the use of identifier variables in the model.

Class Name	Property Name	Type	Description	Codelist Ref
Activity			An action, undertaking, or event, which is anticipated to be performed or observed, or was performed or observed, according to the study protocol during the execution of the study.	
	activityIsConditionalReason	String	The explanation for why the study activity is subject to or dependent upon something else.	
	activityId	String		
	activityIsConditional	boolean	An indication as to whether the study activity is subject to or dependent upon something else.	
	bcCategoryIds	List<String>		
	definedProcedures	List<Procedure>		
	activityName	String	The literal identifier (i.e., distinctive designation) of the clinical study activity.	
	previousActivityId	String		
	biomedicalConceptIds	List<String>		
	activityDescription	String	The textual representation of the study activity.	
	bcSurrogateIds	List<String>		
	activityTimelineId	String		
	nextActivityId	String		
Address			A standardized representation of the location of a person, business, building, or organization. (NCI)	
	country	Code	A sovereign nation occupying a distinct territory and ruled by an autonomous government.	[(Point out to ISO 3166-1 Alpha-3 Country code)]
	city	String	A relatively large and/or densely populated area of human habitation with administrative or legal status that may be specified as a component of a postal address.	
	line	String	The street name and number, building number, apartment or unit number, or post office box number where an entity is physically located.	
	district	String	An administrative or territorial division of a city, town, county, parish, state, country, or other locality based on a shared characteristic.	
	postalCode	String	An alphanumeric code assigned to a mail delivery area.	
	state	String	A sub-division of a country that forms part of a federal union. States are usually, but not always, more autonomous than provinces and may have different laws from the central government.	
	text	String	A standardized representation of the complete set of components denoting the physical address of the person, business, building, or organization.	
AliasCode			An alternative symbol or combination of symbols which is assigned to the members of a collection.	
	standardCodeAliases	List<Code>		
	aliasCodeId	String		
	standardCode	Code		
AnalysisPopulation			A target study population on which an analysis is performed. These may be represented by the entire study population, a subgroup defined by a particular characteristic measured at	

Class Name	Property Name	Type	Description	Codelist Ref
			baseline, or a principal stratum defined by the occurrence (or non-occurrence, depending on context) of a specific intercurrent event. (ICH E9 R1 Addendum)	
	analysisPopulationId	String		
	populationDescription	String	The textual representation of the study population for analysis.	
BiomedicalConcept			A unit of biomedical knowledge created from a unique combination of characteristics that include implementation details like variables and terminologies, used as building blocks for standardized, hierarchically structured clinical research information.	
	bcConceptCode	AliasCode	A concept unique identifier assigned to a biomedical concept that points to the meaning of that biomedical concept.	
	bcProperties	List<BiomedicalConceptProperty>		
	bcSynonyms	List<String>	A word or an expression that serves as a figurative, symbolic, or exact substitute for a biomedical concept, and which has the same meaning.	
	bcReference	String	A citation to an authoritative source for a biomedical concept.	
	biomedicalConceptId	String		
	bcName	String	The literal identifier (i.e., distinctive designation) of the biomedical concept.	
BiomedicalConceptCategory			A grouping of biomedical concepts based on some commonality or by user defined characteristics.	
	bcCategoryParentIds	List<String>		
	bcCategoryDescription	String	The textual representation of the biomedical concept category.	
	bcCategoryMemberIds	List<String>		
	bcCategoryName	String	The literal identifier (i.e., distinctive designation) of the biomedical concept category.	
	bcCategoryChildrenIds	List<String>		
	biomedicalConceptCategoryId	String		
BiomedicalConceptProperty			A characteristic from a set of characteristics used to define a biomedical concept.	
	bcPropertyDatatype	String	The structural format of the biomedical concept property response value. The datatype is carried in the attribute and influences the set of allowable values the attribute may assume. (After HL7)	
	bcPropertyName	String	The literal identifier (i.e., distinctive designation) of the biomedical concept property.	
	bcPropertyId	String		
	bcPropertyConceptCode	AliasCode	A concept unique identifier assigned to a biomedical concept property that points to the meaning of that biomedical concept property.	
	bcPropertyResponseCodes	List<ResponseCode>		
	bcPropertyEnabled	boolean	An indication as to whether the biomedical concept property is activated for use within a given usage context for a biomedical concept.	
	bcPropertyRequired	boolean	An indication as to whether the biomedical concept property is required.	
BiomedicalConceptSurrogate				
	bcSurrogateName	String	The literal identifier (i.e., distinctive designation) of the biomedical concept surrogate.	
	bcSurrogateId	String		
	bcSurrogateDescription	String	The textual representation of the biomedical concept surrogate.	
	bcSurrogateReference	String	A citation to an authoritative source for a biomedical concept surrogate.	
Code			A symbol or combination of symbols which is assigned to the members of a collection.	

Class Name	Property Name	Type	Description	Codelist Ref
	codeId	String		
	code	String	The literal value of a code.	
	codeSystem	String	The literal identifier (i.e., distinctive designation) of the system used to assign and/or manage codes.	
	codeSystemVersion	String	The version of the code system.	
	decode	String	Standardized or dictionary-derived human readable text associated with a code.	
Encounter			Contact between subject/patient and healthcare practitioner/researcher, during which an assessment or activity is performed. Contact may be physical or virtual.	
	previousEncounterId	String	A system identifier assigned to a clinical encounter that occurs immediately prior to the current clinical encounter.	
	encounterName	String	The literal identifier (i.e., distinctive designation) for a protocol-defined clinical encounter.	
	encounterScheduledAtTimingId	String		
	transitionStartRule	TransitionRule		
	encounterEnvironmentalSetting	Code	The environment/setting where the event, intervention, or finding occurred.	[(C127262)]
	nextEncounterId	String	A system identifier assigned to a clinical encounter that occurs immediately after the current clinical encounter.	
	encounterDescription	String	The textual representation of the protocol-defined clinical encounter.	
	encounterContactModes	List<Code>	The means by which an interaction occurs between the subject/participant and person or entity (e.g., a device).	[(C171445)]
	encounterId	String		
	transitionEndRule	TransitionRule		
	encounterType	Code	A characterization or classification of contact between subject/patient and healthcare practitioner/researcher, during which an assessment or activity is performed.	[(C188728)]
Endpoint			A defined variable intended to reflect an outcome of interest that is statistically analyzed to address a particular research question. NOTE: A precise definition of an endpoint typically specifies the type of assessments made, the timing of those assessments, the assessment tools used, and possibly other details, as applicable, such as how multiple assessments within an individual are to be combined. [After BEST Resource] (CDISC Glossary)	
	endpointId	String		
	endpointLevel	Code	A characterization or classification of the study endpoint that determines its category of importance relative to other study endpoints.	[(C188726)]
	endpointDescription	String	The textual representation of the study endpoint.	
	endpointPurposeDescription	String	The textual representation of the study endpoint purpose.	
Estimand			A precise description of the treatment effect reflecting the clinical question posed by a given clinical trial objective. It summarises at a population level what the outcomes would be in the same patients under different treatment conditions being compared. (ICH E9 R1 Addendum)	
	estimandId	String		
	summaryMeasure	String	A synopsis of the clinical endpoint of interest within the analysis target study population.	
	analysisPopulation	AnalysisPopulation		
	treatment	InvestigationalIntervention		
	variableOfInterest	Endpoint		
	intercurrentEvents	List<IntercurrentEvent>		

Class Name	Property Name	Type	Description	Codelist Ref
Indication			The condition, disease or disorder that the clinical trial is intended to investigate or address.	
	codes	List<Code>	A short sequence of characters that represents the disease indication.	[(point out to multiple Biomedical coding dictionaries such as SNOMEDCT (for FDA), MedDRA, NCIt, ICD's, etc.)]
	indicationId	String		
	indicationDescription	String	The condition, disease or disorder that the clinical trial is intended to investigate or address.	[(point out to multiple Biomedical coding dictionaries such as SNOMEDCT (for FDA), MedDRA, NCIt, ICD's, etc.)]
IntercurrentEvent			An event(s) occurring after treatment initiation that affects either the interpretation or the existence of the measurements associated with the clinical question of interest. (ICH E9 Addendum on Estimands)	
	intercurrentEventStrategy	String	A textual description of the planned strategy to manage and/or mitigate intercurrent events.	
	intercurrentEventId	String		
	intercurrentEventName	String	The literal identifier (i.e., distinctive designation) of the intercurrent event.	
	intercurrentEventDescription	String	The textual representation of the intercurrent event.	
InvestigationalIntervention				
	codes	List<Code>		
	investigationalInterventionId	String		
	interventionDescription	String		
Objective			The reason for performing a study in terms of the scientific questions to be answered by the analysis of data collected during the study.	
	objectiveEndpoints	List<Endpoint>		
	objectiveId	String		
	objectiveDescription	String	The textual representation of the study objective. (BRIDG)	
	objectiveLevel	Code	A characterization or classification of the study endpoint that determines its category of importance relative to other study objectives.	[(C188725)]
Organization			A formalized group of persons or other organizations collected together for a common purpose (such as administrative, legal, political) and the infrastructure to carry out that purpose. (BRIDG)	
	organizationId	String		
	organizationIdentifierScheme	String	The name of the organization that provides the identifier for the entity.	
	organizationType	Code	A characterization or classification of the formalized group of persons or other organizations collected together for a common purpose (such as administrative, legal, political) and the infrastructure to carry out that purpose.	[(C188724)]
	organizationName	String	A non-unique textual identifier for the organization. (BRIDG)	
	organizationIdentifier	String	A unique symbol that establishes identity of the organization. (BRIDG)	
	organizationLegalAddress	Address		
Procedure			Any activity performed by manual and/or instrumental means for the purpose of diagnosis, assessment, therapy, prevention, or palliative care.	
	procedureIsConditionalReason	String	The explanation for why the study procedure is subject to or dependent upon something else.	
	procedureType	String	A characterization or classification of the study procedure.	

Class Name	Property Name	Type	Description	Codelist Ref
	procedureCode	Code	A symbol or combination of symbols which is assigned to medical procedure.	[(Point out to external dictionary like CPT, MedDRA, SNOMEDCT, etc.)]
	procedureIsConditional	boolean	An indication as to whether the study procedure is subject to or dependent upon something else.	
	procedureId	String		
ResponseCode			A symbol or combination of symbols representing the response to the question.	
	code	Code	A symbol or combination of symbols which is assigned to the members of a collection.	
	responseCodeId	String		
	responseCodeEnabled	boolean	An indication as to whether the response code is activated for use within a given usage context.	
ScheduleTimeline			A chronological schedule of planned temporal events.	
	scheduleTimelineName	String	The literal identifier (i.e., distinctive designation) of the schedule timeline.	
	scheduleTimelineId	String		
	entryCondition	String	A logical evaluation on which rests the validity of entry into a schedule timeline.	
	scheduleTimelineEntryId	String		
	scheduleTimelineDescription	String	The textual representation of the schedule timeline.	
	scheduleTimelineExits	List<ScheduleTimelineExit>		
	scheduleTimelineInstances	List<ScheduledInstance>		
ScheduleTimelineExit			To go out of or leave the schedule timeline.	
	scheduleTimelineExitId	String		
ScheduledActivityInstance			A scheduled occurrence of an activity event.	
	activityIds	List<String>		
ScheduledDecisionInstance			A scheduled occurrence of a decision event.	
	conditionAssignments	Map<String, String>		
ScheduledInstance			A scheduled occurrence of a temporal event.	
	scheduledInstanceTimings	List<Timing>		
	scheduledInstanceId	String		
	scheduledInstanceType	ScheduledInstanceType	A characterization or classification of the scheduled instance.	[(CNEW)]
	scheduledInstanceEncounterId	String		
	scheduleTimelineExitId	String		
	scheduleSequenceNumber	Integer	A numeral or string of numerals expressing a relative sequence of scheduled temporal events.	
	scheduledInstanceTimelineId	String		
Study			A clinical study involves research using human volunteers (also called participants) that is intended to add to medical knowledge. There are two main types of clinical studies: clinical trials (also called interventional studies) and observational studies. [[http://ClinicalTrials.gov]](CDISC Glossary)	
	studyDesigns	List<StudyDesign>		
	studyRationale	String	A statement describing the overall rationale of the study. This field describes the contribution of this study to product development, i.e., what knowledge is being contributed from the conduct of this study.	
	studyType	Code	The nature of the investigation for which study information is being collected. (After clinicaltrials.gov)	[(C99077)]
	studyProtocolVersions	List<StudyProtocolVersion>		
	studyPhase	AliasCode	A step in the clinical research and development of a therapy from initial clinical trials to post-approval studies.	[(C66737)]

Class Name	Property Name	Type	Description	Codelist Ref
			NOTE: Clinical trials are generally categorized into four (sometimes five) phases. A therapeutic intervention may be evaluated in two or more phases simultaneously in different trials, and some trials may overlap two different phases. [21 CFR section 312.21; After ICH Topic E8 NOTE FOR GUIDANCE ON GENERAL CONSIDERATIONS FOR CLINICAL TRIALS, CPMP/ICH/291/95 March 1998]	
	studyVersion	String	A plan at a particular point in time for a study.	
	studyId	UUID		
	studyTitle	String	The sponsor-defined name of the clinical study.	
	businessTherapeuticAreas	List<Code>	A therapeutic area classification based on the structure and operations of the business unit.	[(point out to external dictionaries)]
	studyAcronym	String	A word or words formed from the beginning letters or a combination of syllables and letters of a compound term, which identifies a clinical study.	
	studyIdentifiers	List<StudyIdentifier>		
StudyArm			A planned pathway assigned to the subject as they progress through the study, usually referred to by a name that reflects one or more treatments, exposures, and/or controls included in the path.	
	studyArmId	String		
	studyArmType	Code	The literal identifier (i.e., distinctive designation) of the study arm type.	
	studyArmName	String	The literal identifier (i.e., distinctive designation) of the study arm.	
	studyArmDataOriginType	Code	A characterization or classification of the study arm with respect to where the study arm data originates.	[(C188727)]
	studyArmDescription	String	The textual representation of the study arm.	
	studyArmDataOriginDescription	String	The textual representation of the study arm data origin.	
StudyCell			A partitioning of a study arm into individual pieces, which are associated with an epoch and any number of sequential elements within that epoch.	
	studyElements	List<StudyElement>		
	studyEpoch	StudyEpoch		
	studyArm	StudyArm		
	studyCellId	String		
StudyDesign			A plan detailing how a study will be performed in order to represent the phenomenon under examination, to answer the research questions that have been asked, and informing the statistical approach.	
	studyDesignId	String		
	studyObjectives	List<Objective>		
	trialIntentType	List<Code>	The planned purpose of the therapy, device, or agent under study in the clinical trial.	[(C66736)]
	bcSurrogates	List<BiomedicalConceptSurrogate>		
	trialType	List<Code>	The nature of the interventional study for which information is being collected.	[(C66739)]
	studyDesignBlindingScheme	AliasCode	The type of experimental design used to describe the level of awareness of the study subjects and/ or study personnel as it relates to the respective intervention(s) or assessments being observed, received or administered.	[(C66735)]
	studyPopulations	List<StudyDesignPopulation>		
	studyInvestigationalInterventions	List<InvestigationalIntervention>		
	biomedicalConcepts	List<BiomedicalConcept>		
	studyDesignName	String	The literal identifier (i.e., distinctive designation) of the study design.	

Class Name	Property Name	Type	Description	Codelist Ref
	studyDesignDescription	String	The textual representation of the study design.	
	studyScheduleTimelines	List<ScheduleTimeline>		
	studyDesignRationale	String	Reason(s) for choosing the study design. This may include reasons for the choice of control or comparator, as well as the scientific rationale for the study design.	
	interventionModel	Code	The general design of the strategy for assigning interventions to participants in a clinical study. (clinicaltrials.gov)	[(C99076)]
	encounters	List<Encounter>		
	activities	List<Activity>		
	bcCategories	List<BiomedicalConceptCategory>		
	studyCells	List<StudyCell>		
	studyIndications	List<Indication>		
	therapeuticAreas	List<Code>	A categorization of a disease, disorder, or other condition based on common characteristics and often associated with a medical specialty focusing on research and development of specific therapeutic interventions for the purpose of treatment and prevention.	[(point out to external dictionaries)]
	studyEstimands	List<Estimand>		
StudyDesignPopulation			The population within the general population to which the study results can be generalized.	
	plannedSexOfParticipants	List<Code>	The specific sex, either male, female, or mixed of the subject group being studied. (NCI)	[(C66732)]
	plannedNumberOfParticipants	int	The planned number of subjects to be entered in a clinical trial. (NCI)	
	plannedMaximumAgeOfParticipants	String	The anticipated maximum age of the subjects to be entered in a clinical trial. (NCI)	
	studyDesignPopulationId	String		
	populationDescription	String	The textual representation of the study population.	
	plannedMinimumAgeOfParticipants	String	The anticipated minimum age of the subjects to be entered in a clinical trial. (NCI)	
StudyElement			A basic building block for time within a clinical study comprising the following characteristics: a description of what happens to the subject during the element; a definition of the start of the element; a rule for ending the element.	
	studyElementDescription	String	The textual representation of the study design element.	
	transitionStartRule	TransitionRule		
	studyElementId	String		
	studyElementName	String	The literal identifier (i.e., distinctive designation) of the study design element.	
	transitionEndRule	TransitionRule		
StudyEpoch			A named time period defined in the protocol, wherein a study activity is specified and unchanging throughout the interval, to support a study-specific purpose.	
	nextStudyEpochId	String		
	previousStudyEpochId	String		
	studyEpochDescription	String	The textual representation of the study epoch.	
	encounters	List<Encounter>		
	studyEpochId	String		
	studyEpochType	Code	A characterization or classification of the study epoch, i.e., the named time period defined in the protocol, wherein a study activity is specified and unchanging throughout the interval, to support a study-specific purpose.	[(C99079)]
	studyEpochName	String	The literal identifier (i.e., distinctive designation) of the study epoch, i.e., the named time period defined in the protocol, wherein a study activity is specified and unchanging throughout the	

Class Name	Property Name	Type	Description	Codelist Ref
			interval, to support a study-specific purpose.	
StudyIdentifier			A sequence of characters used to identify, name, or characterize the study.	
	studyIdentifierId	String		
	studyIdentifier	String	A sequence of characters used to identify, name, or characterize the study.	
	studyIdentifierScope	Organization		
StudyProtocolVersion			A plan at a particular point in time for a formal investigation to assess the utility, impact, pharmacological, physiological, and/or psychological effects of a particular treatment, procedure, drug, device, biologic, food product, cosmetic, care plan, or subject characteristic. (BRIDG)	
	publicTitle	String	The descriptive name of the protocol that is intended for the lay public, written in easily understood language.	
	scientificTitle	String	A more extensive descriptive name of the protocol that is intended for medical professionals, written using medical and scientific language.	
	studyProtocolVersionId	String		
	protocolStatus	Code	A condition of the protocol at a point in time with respect to its state of readiness for implementation.	[(C188723)]
	briefTitle	String	The short descriptive name for the protocol.	
	protocolAmendment	String	A written description of a change(s) to, or formal clarification of, a protocol. (ICH E6)	
	protocolVersion	String	A plan at a particular point in time for a formal investigation to assess the utility, impact, pharmacological, physiological, and/or psychological effects of a particular treatment, procedure, drug, device, biologic, food product, cosmetic, care plan, or subject characteristic. (BRIDG)	
	protocolEffectiveDate	Date	The date and time specifying when the protocol amendment takes effect or becomes operative.	
	officialTitle	String	The formal descriptive name for the protocol.	
Timing			The chronological relationship between temporal events.	
	timingRelativeToFrom	Code	The name of the reference event used to define the temporal relationship with another event.	
	timingId	String		
	timingType	Code	A characterization or classification of the chronological relationship between temporal events.	[(C66728/STENRF)]
	timingWindow	String	A time period, or other type of interval, during which a temporal event may be achieved, obtained, or observed.	
	relativeFromScheduledInstanceId	String		
	relativeToScheduledInstanceId	String		
	timingValue	String	The temporal value of the chronological relationship between temporal events.	
TransitionRule			A guide that governs the allocation of subjects to operational options at a discrete decision point or branch (e.g., assignment to a particular arm, discontinuation) within a clinical trial plan.	
	transitionRuleId	String		
	transitionRuleDescription	String	The textual representation of the transition rule.	

6 USDM API

The Reference Architecture API is designed as a mechanism for bulk transfer. The API has been designed to allow for bulk creation of a study within an Study Definitions Repository, the reading of such a study and the update of a study. No other API features are defined nor is a granular API at this time.
The API has been defined using the OpenApi Specification Version 3. As such the various routes, rules and constraints for the use of the API are contained within the API specification itself. If further rules etc are required then these will be added to the machine readable specification.
Note should be made of the need for cross referencing in the API. As the JSON transported is large there is a need not to repeat content. Therefore therefore the API has been designed to include an instance once and allow for zero, one or more references to it as dictated by the USDM design and the relationships within. This mechanism relies on the identifiers. Within the USDM the UML indicates the place where an instance is included by specifying an attribute and the reference the type of the class. References are all of the type string with the attribute name suffixed with "Id". One exception is the identifier at the head of the model within the Study class. Implementations are free to allocate the value to this field using, for example, a UUID, to ensure uniqueness within the implementation

7 Examples

For the purposes of internal review, examples of use of the model will be provided in a recorded presentation. A link to the presentation can be found on the Internal Review Instructions Page.

8 Appendices

8.1 USDM Team

The USDM has been developed in partnership with both TransCelerate Biopharma and Accenture. CDISC would like to acknowledge the support and input from the following groups:

- TransCelerate DDF Core Team
- TransCelerate Member Company SMEs
- The Accenture DDF development team
- The CDISC DDF volunteer team

The table below represents the CDISC USDM Development team.

Name	Institution/Organization
John Owen	Project Manager, CDISC
Dave Iberson-Hurst	USDM Product Owner, CDISC
Erin Muhlbradt	Controlled Terminology Expert, NCI-EVS
Craig Zwickl	Controlled Terminology Expert, CDISC
Jared Schreibman	Software Engineer, CDISC
Chris Upkes	Principal Consultant, Neo4J
Gaston Guitart	Consulting Engineer, Neo4J

8.2 Glossary and Abbreviations

ADaM	Analysis Data Model
ADaMIG	ADaM Implementation Guide
BRIDG	Biomedical Research Integrated Domain Group
Biomedical Concept	A high-level building block of clinical research and/or healthcare information that encapsulates lower level implementation details like variables and terminologies.
CDASH	Clinical Data Acquisition Standards Harmonization Project
CDISC	Clinical Data Interchange Standards Consortium
CFAST	Coalition for Accelerating Standards and Therapies
Collected	“Collected” refers to information that is recorded and/or transmitted to the sponsor. This includes data entered by the site on CRFs/eCRFs as well as vendor data such as core lab data. This term is a synonym for “captured”.
Controlled Terminology	A finite set of values that represent the only allowed values for a data item. These values may be codes, text, or numeric. A code list is one type of controlled terminology.
CRF	Case report form (sometimes called a case record form). A printed, optical, or electronic document designed to record all required information to be reported to the sponsor for each trial subject.

Domain	A collection of observations with a topic-specific commonality about a subject.
eCRF	Electronic case report form
Foundational Standards	Used to refer to the suite of CDISC standards that describe the clinical study protocol (Protocol), design (Study Design), data collection (CDASH), laboratory work (Lab), analysis (ADaM), and data tabulation (SDTM and SEND). See http://www.cdisc.org/ for more information on each of these clinical data standards.
MedDRA	Medical Dictionary for Regulatory Activities. A global standard medical terminology designed to supersede other terminologies (such as COSUSDMRT and ICD9) used in the medical product development process.
NCI EVS	National Cancer Institute (NCI) Enterprise Vocabulary Services
NIH	National Institutes of Health
Patient	A recipient of medical treatment.
PRO	Patient-reported outcome
SDS	Submission Data Standards. Also the name of the team that maintains the SDTM and SDTMIG.
SDTM	Study Data Tabulation Model
SDTMIG	SDTM Implementation Guide (for Human Clinical Trials)
SHARE	Shared Health and Clinical Research Electronic Library. CDISC's metadata repository.
Subject	A participant in a study.
UML	Unified Modeling Language

8.3 Revision History

8.4 USDM Implementation Guide

The USDM v1.0 was released as part of the DDF Reference Architecture in August 2022. Version v1.0 of the USDM has no associated implementation guide therefore there is no revision history for the Implementation Guide.

8.5 Amendments between USDM v1.0 and USDM v2.0

The table below lists at a high level the major changes that occurred between USDM v1.0 and USDM v2.0

#	Sprint #	Overview	Notes
1	1	Bugfixes and review comments from DDF Phase I	1. StudyEpoch Class: Add encounters relationship, 1 -> 0..* 2. IntercurrentEvent Class: strategy attribute rename to "intercurrentEventStrategy" and is of type String 3. PointInTime Class: remove from the model 4. Encounter Class Attributes "startRule" and "endRule" should be renames and prefixed with "transition", so "transitionStartRule", "transitionEndRule" 5. Workflow Class Attribute "workflowId" renamed to "uuid" 6. Estimand Class Attribute "variableOfInterest" type should be Endpoint not Encounter
2	1	Addition of Therapeutic Area	1. Class: Study Attribute businessTherapeuticArea 2. Class: StudyDesign Attribute therapeuticAreas
3	1	Allow for multiple trial types entries on the StudyDesign class	1. Class StudyDesign Attribute trialType amended to a list
4	2	Terminology Flexibility	1. Code and CodeAlias classes added to the model
5	2	Addition of name and description for StudyDesign class	1. Class: StudyDesign Attribute studyDesignName 2. Class: StudyDesign Attribute studyDesignDescription
7	3	Attribute name changes	1. Class: Study Attribute: studyIdentifier amended to studyIdentifiers 2. Class: Study Attribute: studyProtocolVersion amended to studyProtocolVersions 3. Class: Study Attribute: studyDesign amended to studyDesigns
9	3	Visit Contact Mode	1. Not sure what has changed here
10	4	Allow Study Phase to use the Code Alias	1. Class: Study Attribute studyPhase amended from Code to AliasCode
10	4	Add flaq for Activity and Procedures being optional	1. Class: Activity Attribute activityIsOptional added 2. Class: Procedure Attribute procedureIsOptional added 3. Also see additional change to 16 below
12	5	Additional elements added in to support eCPT population	1. Class: Study Attribute; studyRationale added 2. Class: Study Attribute: studyAcronym added 3. Class: StudyDesignPopulation Attribute: plannedNumberOfParticipants added

#	Sprint #	Overview	Notes
			4. Class: StudyDesignPopulation Attribute: plannedMaximumAgeOfParticipants added 5. Class: StudyDesignPopulation Attribute: plannedMinimumAgeOfParticipants added 6. Class: StudyDesignPopulation Attribute: sexOfParticipants added 7. Class: StudyDesign Attribute: studyDesignRationale added 8. Class: Organization Attribute: organizationLegalAddress added
15	6	New class for Address	Class: Address added with the following attributes • Text • Line • City • District • State • Postal Code • Country
16	6	Amend activityIsOptional and procedureIsOptional to conditional	1. Class: Activity Attribute activityIsOptional amended to activityIsConditional 2. Class: Procedure Attribute procedureIsOptional amended to procedureIsConditional
17	6	Addition of TBLIND/Trial Blinding Schema (valid values in codelist C66735) code to studyDesignBlindingScheme	1. Class: StudyDesign Attribute studyDesignBlindingScheme codelist TBLIND added
19	7	Biomedical Concepts sub model added	Reference Biomedical Concepts for additional information. Addition of the following Classes (note that class StudyData was removed and replaced with the Biomedical Concept classes • BiomedicalConcept • BioemdcialConceptProperty • ResponseCode • BiomedicalConceptCategory • BiomedicalConceptSurrogate
20	9	Study Timing and "Timepoints" added to the model	Reference Study Timing for additional information. Addition of the following Classes (note that class StudyData was removed and replaced with the Biomedical Concept classes • ScheduleTimeline • Timing • ScheduledInstance • ScheduledDecisionInstance • ScheduledActivityInstance • ScheduleTimelineExit

As part of the v2.0 updates the elements of the RA (USDM, CT, API and IG) are stored within a Github repository (see https://github.com/cdisc-org/DDF-RA) and version managed as a series of releases corresponding to the sprints, a subsequent release for internal review, a release for public review and a release for the final publication as v2.0.

- **Controlled Terminology** - for a complete list of controlled terminology changes between USDM v1.0 and and the latest version of the API (For internal Review this will be the API from release 1.09). Please refer to the USDM_CT_Changes.xlsx file in the controlled terminology deliverable folder
- **UML** - A list of changes to the UML model can be found above.

- **API** - For a complete list of API changes between USDM v1.0 and USDM v2.0 it is recommended to use a file comparison tool to compare the API from USDM v1.0 and the latest version of the API (For internal Review this will be the API from release 1.09). Please refer to the USDM API.yaml files in the API deliverable folder.

8.6 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: https://www.cdisc.org/sites/default/files/2020-09/cdisc_policy_003_intellectual_property_v2019.pdf