IntuitionLabs

INTUITIONLABS / RESEARCH & PRACTICAL GUIDANCE

CDISC USDM 4.0: Protocol-to-SDTM Implementation Guide

> Build practical AI for pharma and biotech with IntuitionLabs. We help life-science teams turn complex information and workflows into useful software, governed knowledge systems and AI tools.

Published by IntuitionLabs · 2026-09-23 · cdisc usdm 4.0 · usdm implementation · protocol to sdtm · digital protocol · sdtm trial design · ta te tv ti ts · clinical data standards · study definition repository

Original article: https://intuitionlabs.ai/articles/cdisc-usdm-4-protocol-to-sdtm-guide

In this article

Executive Summary

CDISC Unified Study Definitions Model (USDM) 4.0 is a structured representation of study definitions released on **3 June 2025**, while **Implementation Handbook 1**, published in **June 2026**, explains a narrower and immediately useful workflow: derive the five foundational Study Data Tabulation Model (SDTM) Trial Design domains, **TA, TE, TV, TI and TS**, from a governed digital protocol. (www.cdisc.org) (www.cdisc.org) The model supplies shared study concepts, terminology, an application programming interface (API) specification, conformance rules and an implementation guide; it does not prescribe a sponsor's database or remove the need for a standards programmer to judge the resulting datasets. (www.cdisc.org) (www.cdisc.org) The decision is therefore whether a sponsor can maintain a trustworthy versioned study definition and exchange it with downstream systems, not whether installing a converter will make a submission acceptable.

The practical mapping starts with **StudyArm**, **StudyEpoch**, **StudyCell** and **StudyElement** for the planned pathway in TA and TE; **Encounter** for planned visits in TV; **EligibilityCriterion** and **EligibilityCriterionItem** for TI; and multiple study classes, including objectives and interventions, for TS. (evs.nci.nih.gov) (evs.nci.nih.gov) The handbook flags several derivations that require review: TA branch and transition text, TE duration when schedule timing is incomplete, and TS rows when arrays or repeated administrations appear. (www.cdisc.org) It also states that each USDM API payload carries one study version, whereas final trial design datasets must consider all official amendments. (www.cdisc.org) A controlled pilot should therefore preserve source IDs, approved version IDs, mapping rules, reviewer decisions and the exact output produced at each build.

A feasible architecture has a protocol authoring system publish a validated USDM 4.0 payload into a study definition repository, then give EDC, clinical trial management, trial master file and SDTM build processes contract-specific projections. TransCelerate describes the repository as the exchange point between study builders and downstream **EDC** and **CTMS** consumers. (transcelerate.github.io) (transcelerate.github.io) One version trap is concrete: CDISC's OpenAPI file exposes `/v4/studyDefinitions`, while TransCelerate's reference repository associates USDM 4.0 with its own **V5** API and a required `usdmVersion` header. These are

implementation version labels, not interchangeable URLs. (github.com) (github.com) Current vendor feature pages should be treated as claims about their own integrations, not proof of native USDM 4.0 import. (www.openstudybuilder.com) (www.veeva.com)

The adoption recommendation is a bounded pilot with one approved protocol version, one amendment, the five trial design domains and a small number of downstream consumers. Test payload structure, USDM semantics, mapping business rules, SDTM conformance and traceability separately. CDISC's CORE engine can test conformance, and FDA's **June 2026** technical guide expects Trial Design Model datasets with SDTM study submissions and recommends checking standards, technical and FDA rules before filing. (www.cdisc.org) (www.fda.gov) CDISC Handbook 1 gives illustrative estimates for the workflow: the existing process typically takes 8–32 hours of defining and programming plus 2–3 weeks for alignment, while the envisioned approach may reduce programmer effort to 1–2 hours for a final check. These are CDISC's indicative workflow estimates, not independently measured end-to-end savings; report local mapping coverage and unresolved decisions rather than treating them as a guaranteed productivity gain.

Introduction and Background

A protocol is read by people, but its arm structure, visit schedule, eligibility logic and study metadata are repeatedly interpreted by systems. USDM addresses that exchange problem by representing a study definition as structured data. CDISC's Digital Data Flow materials position the logical model, controlled terminology, API and implementation guide as complementary artifacts, with the logical model serving implementers rather than specifying a physical database. (www.cdisc.org) (transcelerate.github.io) TransCelerate describes a study definition repository as a central store for standardized information originating in digital protocols and related sources. (transcelerate.github.io)

The audience here is the clinical standards lead who must turn an approved definition into valid study metadata, the study-build team responsible for EDC and operations systems, and the architect responsible for versioned API exchange. The relevant decision is whether USDM 4.0 is ready for a controlled protocol-to-SDTM pilot. It is: CDISC has a current release, conformance materials and a focused handbook, and the published API and reference repository offer concrete integration starting points. (www.transceleratebiopharmainc.com) (transcelerate.github.io) (transcelerate.github.io) That does not mean every trial design variable is automatically resolved or every vendor has a production USDM 4.0 interface.

The term **Trial Design Model (TDM)** refers to SDTM datasets describing planned trial structure. **TA** is Trial Arms, **TE** Trial Elements, **TV** Trial Visits, **TI** Trial Inclusion/Exclusion Criteria, and **TS** Trial Summary. CDISC's SDTM implementation guide defines TA at the level of a planned element within an arm, TE as the definitions of those elements, TV as planned visits, TI as criteria, and TS as structured study parameters. (www.cdisc.org) (www.pmda.go.jp) (phuse-org.github.io) (evs.nci.nih.gov) (www.pmda.go.jp) The planned design must be distinguished from subject-level records of what actually happened. (www.cdisc.org)

IntuitionLabs approaches the question as an adjacent integration and submission-engineering adviser, not as a USDM repository vendor. Its published service description covers SDTM/ADaM architecture and data flows from EDC through submission packaging; its CTMS practice describes API integration and validation across clinical operations. (intuitionlabs.ai) (intuitionlabs.ai) That perspective makes the integration contract and the human approval boundary central to the adoption decision.

Key Changes

From narrative protocol to versioned study objects

USDM 4.0 turns protocol design into linked objects that can be referenced consistently across downstream projections. CDISC released version 4.0 in **June 2025** with a logical model, controlled terminology, API specifications, conformance rules and an implementation guide. (github.com) (www.transceleratebiopharmainc.com) The implementation handbook added in **June 2026** takes a specific use case, constructing TA, TE, TV, TI and TS, instead of attempting to define every downstream artifact. (clinline.org) Separate **standardized source content** from product authoring and storage. A sponsor may store content in a graph, relational store or document service as long as the exchanged definition and its semantics meet the chosen contract. (transcelerate.github.io) (transcelerate.github.io)

A structured study definition is also not identical to the final clinical protocol document. The final ICH M11 guideline defines a harmonized protocol template and technical specification, but says these do not prescribe the process of developing a protocol. (database.ich.org) HL7 describes ICH M11 elements as a subset of USDM, while its protocol implementation guide treats FHIR as an exchange representation and says deeper USDM integration is still evolving. (hl7.org) (hl7.org) Teams should agree which approved protocol content is authoritative, which values are represented in USDM, and which narrative passages remain governed outside the model.

From isolated domain programming to traceable derivation

Handbook 1 documents source-to-output relationships that can be expressed as repeatable transforms. For example, TA combines **StudyArm**, **StudyEpoch** and **StudyElement** through **StudyCell**; TE is anchored in **StudyElement**; and TV is largely based on **Encounter**. (raw.githubusercontent.com) NCI's DDF terminology defines a study design cell as a piece of an arm associated with its design context and an encounter as physical or virtual contact linked to assessments or activities. (evs.nci.nih.gov) These are useful boundary definitions for cross-system contracts: EDC needs schedule detail, CTMS needs operational visit or milestone alignment, and the SDTM program needs planned design rows. (transcelerate.github.io) (www.veeva.com)

The transform is not pure field copying. USDM's previous and next references determine some planned ordering, so generated **TV.VISITNUM** and **TA.TAEORD** require graph traversal and deterministic tie handling. (www.cdisc.org) Branch and transition descriptions may require suggested text and standards review. (www.cdisc.org) The implementation can prepopulate candidate records, but it should log the rule, source IDs and approval decision for every derived value. FDA emphasizes traceability from source data through tabulation and analysis, which makes lineage a practical design criterion. (www.fda.gov)

From one snapshot to amendment-aware output

The handbook notes that a USDM API payload contains **one study version** and that all official versions must be considered when producing final trial design domains. (www.cdisc.org) CDISC's SDTM guidance likewise says TI must include each complete set of criteria when criteria change, distinguished by **TIVERS**, and a

changed criterion is treated as a new criterion. (www.cdisc.org) (www.cdisc.org) A repository that overwrites its old JSON would lose the evidence needed to explain which planned visits or eligibility text applied at a given time.

A pilot should therefore treat each approved amendment as a new immutable source snapshot. Record approval date, effective scope, source-document version, USDM version identifier, payload hash, mapping code version and output build ID. The reference SDR API documents revision history, version creation and comparison operations, although a sponsor still needs to validate how its own deployment implements them. (github.com) These controls support a clear answer to a standards reviewer: why this TA, TV or TI row exists, and which protocol version created it.

Artifact Map and Ownership

The implementation question starts with artifact selection. USDM is a family of linked specifications, not one downloadable JSON file. CDISC's DDF release page lists the model, terminology, API, conformance rules and guide, while the current repository documents separately licensed code and documentation. (transcelerate.github.io) (github.com) The errata page should be included in every version-selection record because CDISC documented differences between written conformance specifications and executable CORE rules. (www.cdisc.org)

Table 1 allocates a practical owner and update trigger to each artifact. The owner column is a recommended governance assignment, not a claim that CDISC mandates that role.

Artifact	Purpose and current evidence	Recommended owner and trigger
USDM 4.0 logical model and guide	Defines study concepts and implementation guidance, released with v4.0 on 3 June 2025. (transcelerate.github.io) (github.com)	Clinical standards owner; review when CDISC publishes a model or guide revision.
Controlled terminology	Supplies coded values for USDM concepts; NCI publishes dated terminology versions and archives. (www.cancer.gov)	Terminology steward; pin each import to a named terminology release.
OpenAPI and JSON contract	CDISC repository publishes an OpenAPI study-definition interface and a whole-study create operation. (github.com)	Enterprise architect; pin specification commit or release tag with consumer tests.
Conformance specifications, executable rules, errata	CDISC provides USDM rules through CORE and records rule differences in errata. (www.fda.gov) (github.com)	Validation lead; compare written and executable rule versions before release.
Handbook 1 and reference implementation	Describes construction of TA, TE, TV, TI and TS; CDISC describes the example implementation as illustrative and incomplete. (www.cdisc.org)	SDTM standards lead; approve mapping rules and exceptions for each study.

The matrix prevents a frequent version error: a model version, an API endpoint version, a terminology release and an SDTMIG version are different identifiers. In the CDISC specification, the interface path is `/v4/studyDefinitions`; in TransCelerate's SDR example, the endpoint family is **V5** for a payload identifying **USDM 4.0**. (github.com) (github.com) Pin both sides of the contract in deployment records. For licensing, the

DDF-RA repository states MIT for code and scripts and CC BY 4.0 for documentation; confirm the exact license of every component and terminology source included in a product distribution. (github.com) (www.cancer.gov)

Reference Architecture and Exchange Contracts

A vendor-neutral pipeline has four boundaries. First, a protocol-authoring tool captures approved design objects. Second, a study definition repository receives the canonical USDM payload and retains each approved version. Third, consumer adapters create purpose-specific projections for EDC, CTMS, electronic trial master file (**eTMF**) and study-data programming. Fourth, validation services record conformance and reviewer decisions before an output is promoted. TransCelerate explicitly describes study builders upstream and EDC and CTMS downstream, with room for more consumers. (www.transceleratebiopharmainc.com) (transcelerate.github.io) Its SDR example is a reference implementation for connectivity, not a finished commercial product. (transcelerate.github.io)

The source system should publish one versioned **StudyDefinition/StudyVersion** exchange unit with stable identifiers. Ingestion should reject duplicate IDs, dangling references, invalid terminology and unsupported schema revisions before mapping begins. The repository should expose a read contract that returns the source payload and its version metadata, and a history contract for amendments. CDISC's OpenAPI includes whole-study POST and GET operations plus study history; TransCelerate's reference API documents GET, POST, PUT, validation and compare endpoints. (transcelerate.github.io) (transcelerate.github.io) An integration should not infer compatibility merely from a URL path: test payload shape, required headers, status codes, idempotency and error behavior against the deployed implementation.

The consumers need different projections:

- **EDC study build:** planned encounters, activities and timing inform visit and form configuration, but an EDC-specific build still requires form metadata, edit checks, role configuration and testing. An OpenStudyBuilder page describes ODM export for EDC setup, which is evidence of a separate exchange path rather than proof that all EDCs ingest USDM directly. (www.openstudybuilder.com)
- **CTMS operations:** approved study IDs, arms, visit milestones and dates can seed operational planning. A documented Veeva connection syncs selected CTMS and EDC study data after configuration, but that product documentation does not establish a native USDM 4.0 import. (www.veeva.com) (cdmshelp.veeva.com)
- **eTMF document control:** publish protocol version and amendment identifiers, document expectations and links to approved artifacts. A clinical operations connection can transfer final case report form material to an eTMF, but an explicit USDM-to-eTMF content mapping must be designed and tested locally. (www.veeva.com)
- **SDTM trial design:** read approved design snapshots, run a deterministic transform into TA, TE, TV, TI and TS, then route exceptions through standards review. Handbook 1 is the controlling mapping starting point. (clinline.org) (www.fda.gov)
- **FHIR interoperability:** treat FHIR as an additional exchange contract. HL7's protocol guide uses ResearchStudy and Composition, while developing schedule guidance scopes USDM extraction to timelines, encounters and activities. (hl7.org) (build.fhir.org)

A one-way projection from canonical USDM into each consumer reduces ambiguity about ownership. HL7's developing USDM-to-FHIR schedule guidance explicitly recommends reading the USDM source rather than previously generated FHIR output. (build.fhir.org) That principle also helps the SDTM build: a corrected TV row should be explained through a corrected source, mapping rule or documented reviewer override, rather than silently edited in a downstream spreadsheet.

Worked Protocol-to-SDTM Mapping

Consider a small illustrative design with a **Placebo** arm, a **Screening** epoch, one planned element, a **Screening 1** encounter and one inclusion criterion. These identifiers appear in CDISC's pinned USDM 4.0 pilot sample; only a reduced slice is used here. (raw.githubusercontent.com) The sample's StudyCell explicitly joins the arm, epoch and element IDs:

```
{
  "id": "StudyCell_1",
  "extensionAttributes": [],
  "armId": "StudyArm_1",
  "epochId": "StudyEpoch_1",
  "elementIds": ["StudyElement_1"],
  "instanceType": "StudyCell"
}
```

This is a valid JSON object copied from the pinned sample, not a complete standalone USDM study payload. The sample also places `arms`, `epochs`, `elements`, `studyCells`, `encounters` and `eligibilityCriteria` under a study design, while reusable `eligibilityCriterionItems` sit under its study version. (raw.githubusercontent.com) A production transform resolves IDs against the entire versioned payload. If a referenced element or criterion item is missing, generation stops and records a source error rather than emitting a guessed SDTM value.

Table 2 is a minimal class-to-domain crosswalk. It follows Handbook 1's mapping direction and deliberately separates source object, generated value and review point. The exact SDTM variable set must be checked against the study's selected SDTMIG version. (www.fda.gov) (www.fda.gov)

USDM 4.0 source	SDTM target and minimal derivation	Human-governed check
`StudyArm` + `StudyEpoch` + `StudyCell` + `StudyElement`	**TA:** one planned element per arm pathway; traverse the linked cells to set arm, epoch and `TAEORD`. (www.cdisc.org) (www.cdisc.org)	Confirm planned branch and transition text; the handbook says `TABRANCH` is not directly available and `TATRANS` may need an editable suggestion. (www.cdisc.org)
`StudyElement` and related schedule/intervention timing	**TE:** create one element definition even if reused across arms; derive `TEDUR` only when sufficient timing exists. (www.cdisc.org) (www.cdisc.org)	Confirm start and end rules and any derived duration before freezing output.
`Encounter` with schedule links and arm context	**TV:** derive planned clinical encounters, visit order and arm-specific rows where visits diverge. (www.cdisc.org) (www.cdisc.org)	Resolve previous/next references deterministically and review `VISITNUM`, planned day and visit window. (www.cdisc.org)

USDM 4.0 source	SDTM target and minimal derivation	Human-governed check
`EligibilityCriterion` + `EligibilityCriterionItem` + `StudyVersion`	**TI:** create inclusion/exclusion rows from criterion identifiers, categories and text; map the study version into `TIVERS`. (www.cdisc.org) (www.cdisc.org)	Approve criterion text and preserve every complete amended set. (www.who.int)
Study identifiers, objectives, interventions and other study-level classes	**TS:** form parameter/value rows for protocol and study characteristics; split arrays into separate `TSSEQ` records. (www.cdisc.org) (www.cdisc.org)	Select sponsor-scoped `STUDYID`, deduplicate repeated intervention facts and resolve parameters not final until conduct ends. (www.cdisc.org)

For this slice, the **TA** output would contain the planned path through `StudyArm_1`, `StudyEpoch_1` and `StudyElement_1`, while the **TE** output would define `StudyElement_1` once. The **TV** projection would locate the screening encounter in the ordered schedule, and **TI** would point from the design-specific criterion to its reusable criterion item and the approved study version. **TS** would draw from study-level metadata and objectives that are outside the single StudyCell object. (raw.githubusercontent.com) (raw.githubusercontent.com) The example intentionally does not invent a complete submission row, because the sample slice lacks the full code lists, timing decisions and sponsor-specific review needed to populate one correctly.

Two implementation details carry disproportionate risk. First, the handbook instructs implementers to select the **sponsor-scoped study identifier** for `STUDYID`, not merely the first identifier in an array. (www.cdisc.org) Second, ordering is often represented through prior and next relationships, not stored ordinal values. A transform must detect cycles, disconnected nodes and ambiguous branches before assigning `TAEORD` or `VISITNUM`. (raw.githubusercontent.com) Those checks are testable with a small fixture set: a straight path, a reused element, a branching arm, a changed visit and an amended eligibility criterion.

Implementation Considerations and Process Changes

API, JSON, identifiers and terminology

The simplest transport pattern is to publish an approved USDM JSON payload and an immutable manifest containing its study ID, study version, schema version, terminology release and content hash. The CDISC OpenAPI contract supports whole-study operations, while the TransCelerate SDR reference API adds operations for validation and comparison. (transcelerate.github.io) (transcelerate.github.io) The API's path version is a product-interface identifier; the model version belongs in payload and contract metadata. In the SDR example, `/v5` is associated with USDM 4.0 and uses a required `usdmVersion` header. (github.com) Pin the precise OpenAPI file or release tag used in the pilot, such as CDISC's `v4.0.0` sample and API artifacts, instead of assuming a moving `main` branch is stable. (raw.githubusercontent.com) (github.com)

Use stable source IDs within a study version and preserve cross-references during export. The CDISC sample visibly uses object IDs such as `StudyCell_1`, `StudyArm_1` and `StudyElement_1`; the helper package documents workbook-unique IDs for cross-references. (raw.githubusercontent.com) (github.com) Where a

consumer has its own keys, maintain an explicit mapping table with source object ID, target system ID, effective version and mapping status. Do not derive durable identity from a display label such as “Screening 1,” because labels can be edited without changing the underlying concept.

Terminology should be treated as versioned input, not decoration. NCI publishes dated CDISC terminology archives and states that the terminology can be used without licensing restrictions. (www.cancer.gov) A validation service should record both the code and the code-system/version used to interpret it; tests should cover deprecated, missing and mismatched values. The CDISC CORE documentation says USDM rules are expressed in JSONata and that its engine can test study data against CDISC and custom rules. (github.com) (github.com) These tests complement schema validation: a syntactically legal payload can still have invalid references or inconsistent clinical meaning.

Amendments and downstream propagation

On approval of an amendment, freeze the previous source snapshot, ingest the new one and compute a semantic difference by stable object ID. The SDR reference API documents revision history and a comparison operation, but the sponsor should independently preserve the comparison result used for each build. (github.com) Classify a change by affected consumer and output variable rather than sending a blanket “protocol changed” alert. A revised eligibility criterion may change TI and screening configuration; a new encounter may change TV, EDC visit setup and CTMS milestones; an arm or epoch change may alter TA and TE. These are impact-assessment rules to implement and test, not claims that any current vendor performs them automatically. (transcelerate.github.io) (transcelerate.github.io)

- **Eligibility text or category:** compare `EligibilityCriterion` and `EligibilityCriterionItem`; regenerate TI, preserve prior complete sets, and review `IETESTCD` and `TIVERS`. (raw.githubusercontent.com) (www.who.int) (evs.nci.nih.gov)
- **Encounter timing or sequence:** regenerate TV; review EDC visit schedules, CTMS milestones and planned window logic. (phuse-org.github.io) (build.fhir.org) (www.veeva.com)
- **Arm, epoch, cell or element:** regenerate TA and possibly TE; examine transition logic, reused elements and treatment-path descriptions. (raw.githubusercontent.com) (www.pmda.go.jp)
- **Objective, intervention or trial metadata:** regenerate affected TS parameters; split array values and deduplicate as the handbook directs. (github.com) (www.pmda.go.jp)
- **Document-only wording:** retain the approved protocol document and assess whether any structured field changed; do not create a dataset delta without a source change.

For each change, record the source object IDs, before/after values, mapping-rule version, target artifacts, reviewer, decision and build output hash. TI's version handling is especially explicit: CDISC says each complete set of criteria must be included when criteria change. (phuse-org.github.io) The handbook also warns that final quality control of the trial design domains may need to wait until study end because some TS information reflects actual execution. (www.cdisc.org) “Automatically generated” should therefore describe the draft construction step, not final sign-off.

Validation and release gates

A controlled build uses layered checks with separate failure reports:

- **Transport and schema:** parse JSON, validate against the pinned USDM 4.0 schema/OpenAPI interpretation and reject missing required fields or unsupported revisions. CDISC CORE development documentation identifies schemas derived from DDF-RA OpenAPI specifications. (github.com)
- **Semantic conformance:** run the applicable USDM rules and terminology checks, then reconcile any differences between textual specifications and executable rules against CDISC's errata. (www.gov.uk) (github.com)
- **Mapping logic:** test joins, branch ordering, duplicate elimination, array-to-row expansion, sponsor `STUDYID` selection and version selection against approved fixtures. (raw.githubusercontent.com) (raw.githubusercontent.com)
- **SDTM conformance:** run the study's selected SDTMIG and conformance-rule versions against generated TA, TE, TV, TI and TS, then document justified exceptions. (www.cdisc.org) (www.fda.gov)
- **Submission context:** compare the outputs with regulatory technical rules and reviewer-guide expectations. FDA recommends presubmission evaluation against standards conformance, technical rejection criteria and FDA business rules; PMDA's guide also calls for TDM datasets and accompanying metadata. (www.fda.gov) (www.pmda.go.jp)
- **Human approval:** require protocol owner and SDTM standards lead sign-off for branch logic, criterion interpretation, incomplete schedule timing and TS parameters that depend on study conduct. (www.gov.uk) (www.fda.gov)

This separation matters because passing a JSON schema says nothing about whether a proposed `TATRANS` sentence is clinically correct. Likewise, an SDTM validator cannot prove that an arm was sourced from the right approved amendment. CDISC says both USDM and SDTM rules can apply to construction of the domains, and FDA emphasizes traceability across source, tabulation and analysis. (www.fda.gov) (www.fda.gov) The release record should capture every validator version, rule set, unresolved issue and reviewer disposition.

Data Analysis and Evidence

The most defensible quantitative claim about this workflow is its **scope**, not its speed. CDISC's June 2026 handbook covers **five** foundational trial design domains: TA, TE, TV, TI and TS. (clinline.org) CDISC's SDTMIG describes specific record grains for each: TA is one planned element per arm, TE defines each unique element, TI records each inclusion or exclusion criterion, and TV describes planned visits. (www.pmda.go.jp) (github.com) (www.pmda.go.jp) (www.pmda.go.jp) These grains let a team estimate its own build and review load from its actual protocol instead of importing an unverified productivity benchmark.

Table 3 is a reader-entered pilot sizing worksheet. Counts are illustrative input fields, not industry averages. The multiplication is a screening measure of review surface, not a forecast of person-hours or cost.

Measure	Reader-entered value	How to use it
Approved protocol elements, `P`	___ arms, epochs, elements, encounters, criteria and study parameters	Count distinct structured objects in the pilot snapshot and amended snapshot. (raw.githubusercontent.com)

Measure	Reader-entered value	How to use it
Downstream consumers, `C`	___ EDC, CTMS, eTMF, SDTM or other interfaces	Count only consumers with an actual contract and acceptance test. (transcelerate.github.io) (transcelerate.github.io)
Approved mappings, `M`	___ source-to-target rules	Count rule versions, including joins, ordering and terminology translations. (github.com) (transcelerate.github.io)
Validation rules, `V`	___ schema, USDM, mapping and SDTM checks	Record executed rules and exceptions, not just configured rule libraries. (ispe.org) (github.com)
Review surface	`P × C` object-consumer touchpoints and `M + V` controlled rules	Compare baseline and amendment runs; do not equate counts with effort without local measurements.

A second metric is **mapping coverage**: the number of required target fields with an approved source or explicit derivation divided by the number of required target fields in the pilot's selected SDTMIG version. A third is **unresolved decision rate**: open expert decisions divided by candidate output records. A fourth is **amendment impact precision**: changed target artifacts confirmed by review divided by all artifacts flagged by the change detector. These are proposed pilot measures. They become evidence only when the sponsor records counts, denominator definitions, protocol versions and reviewer outcomes.

As of **September 2026**, official source dates also indicate what must be pinned: USDM 4.0 was released **3 June 2025**; Handbook 1 followed in **June 2026**; CDISC's CORE page described executable rule coverage and the USDM errata identifies specification-to-rule differences. (github.com) (github.com) (www.fda.gov) FDA's **June 2026** technical guide generally expects TDM datasets with each SDTM study submission and asks sponsors to evaluate standards, technical and business rules. (www.fda.gov) A 2026 CDISC 360i demonstration described an automated pipeline from USDM JSON study design toward SDTM-ready outputs, with CORE checks on both USDM and SDTM; its authors also stated that USDM alone is insufficient for seamless end-to-end automation. (phuse.s3.eu-central-1.amazonaws.com) (phuse.s3.eu-central-1.amazonaws.com) None of these dates implies a regulator mandate to author protocols in USDM. They establish a usable standards stack and a separate obligation to produce conformant submission data. The evidence collected for this report did not establish a public, independently measured time or cost benchmark for automatic protocol-to-SDTM generation; procurement cases should request a local demonstration against the worksheet above.

Implications and Future Directions

A build, buy or partner choice should be evaluated against the same controlled use case. **Build** is plausible when an organization already owns a versioned study-definition service and SDTM mapping library. **Buy** is plausible when a product demonstrates USDM 4.0 import/export, amendment comparison, source lineage and testable TA/TE/TV/TI/TS output against the sponsor's protocol. **Partner** is plausible when the missing work is cross-system contract design, terminology governance and validation evidence. The scorecard should ask for a live sample payload, pinned schema, mapping-rule inventory, exception queue, version diff and reproducible output, not a generic “digital protocol” label. TransCelerate calls its SDR a reference interoperability

demonstration, while OpenStudyBuilder's published API guide documents **USDM 3.11** export and describes 4.0 as a plan; support must be checked against the actual deployed release. (transcelerate.github.io) (www.openstudybuilder.com)

A bounded pilot can use one interventional study with a straight arm path, one schedule change and one eligibility amendment. The five outputs should be built twice, once before and once after the amendment, and compared at record and variable level. The pilot succeeds when source IDs and version identifiers survive the exchange, generated records match an expert-approved gold set, exceptions are explainable, and downstream consumers can process the change without hidden manual edits. CDISC's handbook already identifies the domain-specific review points, including transition text, timing, TI versioning and TS completion. (build.fhir.org) HL7's developing schedule guidance scopes extraction to timelines, encounters and activities, which reinforces the need for explicit activity and timing contracts when extending the pilot beyond SDTM. (build.fhir.org)

Broader interoperability remains a moving target. HL7's protocol guide uses FHIR ResearchStudy and Composition and describes future expansion of USDM integration; its developing USDM schedule guidance is deliberately limited to schedule objects. (hl7.org) (build.fhir.org) WHO registry guidance separately expects a protocol file link with version and date and an audit trail for registry changes. (www.who.int) (www.who.int) Those are distinct interfaces, each with its own owner and validation scope. For an adjacent adviser such as **IntuitionLabs**, the practical contribution is to design the API and lineage contract, connect clinical systems and make review evidence reproducible, consistent with its published SDTM and CTMS integration services. (intuitionlabs.ai) (intuitionlabs.ai)

Frequently Asked Questions (FAQs)

Is USDM 4.0 mandatory for an SDTM submission?

The verified FDA guide expects TDM datasets with SDTM study submissions and recommends standards and business-rule checks; it does not in this evidence prescribe USDM as the protocol-authoring format. (www.fda.gov) USDM is a way to structure source design information. The sponsor still chooses its supported SDTMIG version, validates the generated datasets and documents material discrepancies. (www.fda.gov) (www.fda.gov)

Can a USDM JSON file generate TA, TE, TV, TI and TS automatically?

Handbook 1 describes automatic construction of those five domains from a structured definition, but it explicitly leaves some branch, transition, timing and final study-summary decisions for review. (clinline.org) The repository or transform should produce a draft with lineage and exceptions, followed by expert sign-off and SDTM conformance testing. CDISC says its accompanying reference implementation illustrates most required features but is incomplete. (www.cdisc.org)

How are protocol amendments represented in TI and other domains?

Keep each approved USDM study-version payload, compare it with the previous version and regenerate affected outputs. The handbook says all official versions matter; SDTMIG says TI includes each complete criterion set with `TIVERS`, and a changed criterion is a new criterion. (www.who.int) (www.cdisc.org) (www.who.int) Persist the diff, output build and reviewer decision so the final submission can be reconstructed.

Does `/v5` mean USDM 5.0?

No. In the published TransCelerate SDR reference repository, the V5 endpoint family supports **USDM 4.0** and expects a `usdmVersion` header. CDISC's separate OpenAPI artifact uses a `/v4/studyDefinitions` path. (github.com) (github.com) Confirm both the model version and each deployed API contract before integration.

Conclusion

USDM 4.0 is ready for a controlled protocol-to-SDTM pilot because CDISC has published the model and exchange artifacts and Handbook 1 supplies an explicit route to the five foundational Trial Design domains. (www.transceleratebiopharmainc.com) (github.com) The engineering task is to make a versioned study definition trustworthy: resolve object IDs, pin terminology and API contracts, derive planned design records deterministically, and retain lineage from approved protocol content to each generated value. The standards task is to approve the interpretations that the model alone cannot settle, especially branch and transition text, schedule-derived timing, amended criteria and late-finalized trial summary parameters. (www.who.int) (www.who.int)

The strongest pilot output is not a polished demo screen. It is a reproducible package containing the source payload and its version, the mapping code and rules, the generated TA/TE/TV/TI/TS datasets, a semantic amendment diff, USDM and SDTM validation reports, and a signed exception log. CDISC and FDA materials support that layered approach to conformance and traceability. (github.com) (www.fda.gov) If the package survives an amendment and is intelligible to the EDC, CTMS and submission teams, the organization has evidence for scale-up. If not, the pilot has identified the specific missing contract or governance decision while the scope is still small and accountable.

Source links

Links cited in this article, in order of first appearance. Full addresses are provided for readers using extracted text.

1. https://www.cdisc.org/ddf/usdm-errata-3-4
2. https://www.cdisc.org/sites/default/files/2026-06/USDM-HB1%20v1.0_FINAL.pdf
3. https://www.cdisc.org/ddf
4. https://evs.nci.nih.gov/ftp1/CDISC/DDF/DDF%20Terminology.html
5. https://transcelerate.github.io/ddf-home/overview.html
6. https://transcelerate.github.io/ddf-home/faq.html
7. https://github.com/cdisc-org/DDF-RA/blob/main/Deliverables/API/USDM_API.yaml
8. https://github.com/transcelerate/ddf-sdr-api

9. https://www.openstudybuilder.com/guide_api/
10. https://www.veeva.com/products/vault-platform/veeva-connections/clinical-operations/
11. https://www.cdisc.org/core
12. https://www.fda.gov/media/153632/download
13. https://www.transceleratebiopharmainc.com/initiatives/digital-data-flow/
14. https://www.cdisc.org/standards/foundational/sdtmig/sdtmig-v3-3/html
15. https://www.pmda.go.jp/files/000267937.pdf
16. https://phuse-org.github.io/fhir-schedule-of-activities-ig/artifacts.html
17. https://intuitionlabs.ai/services/sas-life-sciences
18. https://intuitionlabs.ai/services/veeva-vault-ctms
19. https://github.com/cdisc-org/cdisc-rules-engine/blob/main/docs/development.md
20. https://clinline.org/previous-presentations-and-recordings/
21. https://intuitionlabs.ai/articles/clinical-trial-protocol-guide
22. https://database.ich.org/sites/default/files/ICH_Step4_M11_Final_Guideline_2025_1119.pdf
23. https://hl7.org/fhir/uv/pharmaceutical-research-protocol/2025Sep/06-info-models.html
24. https://hl7.org/fhir/uv/pharmaceutical-research-protocol/2025Sep/index.html
25. https://raw.githubusercontent.com/cdisc-org/DDF-RA/v4.0.0/Documents/Examples/CDISC_Pilot/CDISC_Pilot_Study.json
26. https://github.com/cdisc-org/DDF-RA
27. https://www.cancer.gov/about-nci/organization/cbiit/vocabulary/cdisc
28. https://www.transceleratebiopharmainc.com/how-transcelerate-microsoft-accenture-and-cdisc-are-working-to-revolutionize-digital-data-flow/
29. https://www.openstudybuilder.com/info_ddf/
30. https://cdmshelp.veeva.com/gr/study-administrators/clinops-edc-connection/
31. https://build.fhir.org/ig/HL7/Vulcan-schedule-ig/en/usdm-alignment.html
32. https://www.who.int/tools/clinical-trials-registry-platform/network/who-data-set
33. https://github.com/cdisc-org/usdm
34. https://www.who.int/tools/clinical-trials-registry-platform/network/registry-criteria
35. https://github.com/ClinLine/USDM_SDTM_mapper
36. https://www.gov.uk/government/publications/guidance-on-gxp-data-integrity
37. https://www.cdisc.org/standards/foundational/sdtmig/sdtm-and-sdtmig-conformance-rules-v2-0
38. https://clinline.org/wp-content/uploads/2025/10/Poster_USDM2SDTM.pdf
39. https://ispe.org/topics/gamp
40. https://phuse.s3.eu-central-1.amazonaws.com/Archive/2026/Connect/US/Austin/PAP_DS02.pdf
41. https://www.fda.gov/industry/fda-data-standards-advisory-board/study-data-standards-resources
42. https://www.who.int/tools/clinical-trials-registry-platform/about/glossary

THE PUBLISHER

About IntuitionLabs

IntuitionLabs is an AI consulting, custom software development and data engineering firm serving pharmaceutical, biotechnology, medical-device and other life-science organizations. We work with clinical, regulatory, medical-affairs, commercial, quality and IT teams to connect technology decisions with the work people need to accomplish.

AI consulting and adoption

Our AI enablement services cover readiness assessments, use-case selection, governance and policies, team workshops, adoption measurement and ongoing advisory support. We help organizations structure the information layer behind AI: source material, context, permissions and maintained knowledge that make generated answers useful and reviewable. Private LLM inference and hosted AI options support teams evaluating how to operate AI with appropriate control over their data and infrastructure.

Software, data and life-science workflows

IntuitionLabs develops custom software for pharma and biotech, integrates enterprise systems, and builds data engineering and business intelligence solutions. Areas of focus include AI agents, regulatory research, medical writing, medical affairs, CMC information, competitive intelligence and clinical-document workflows. Our eTMF intelligence work includes cross-system reconciliation and inspection-readiness support.

Enterprise platforms and regulated delivery

We provide Veeva services, application support, managed services, integrations and custom applications, alongside enterprise content work involving platforms such as Egnyte. For regulated workflows, our services include GxP enablement, computer-system validation and software development addressing 21 CFR Part 11 requirements. The applicable controls, validation responsibilities and acceptance criteria are defined for each engagement.

Work with IntuitionLabs

Explore AI enablement, pharma and biotech software development, data engineering and BI, and Veeva services. Contact IntuitionLabs to discuss your workflow, information sources and implementation needs.

IntuitionLabs publishes educational research to help life-science teams make informed technology decisions. Coverage of a product or organization does not imply a client relationship, endorsement or partnership.

Website: https://intuitionlabs.ai