

A Guide to CDISC Standards: Understanding SDTM and ADaM

11/19/2025 • 40 min read

cdisc

sdm

adam

clinical data standards

clinical trials

regulatory submission

data mapping

fda



[Revised April 5, 2026]

Executive Summary

Clinical data standards have transformed how regulatory authorities and sponsors manage and submit clinical trial information. Central to these standards are CDISC's **Study Data Tabulation Model (SDTM)** and **Analysis Data Model (ADaM)**, which respectively define standardized formats for tabulating raw trial data and preparing analysis-ready datasets. Globally, regulators (the US FDA, Japan's PMDA, China's NMPA, etc.) now **mandate** or strongly encourage the use of CDISC standards for electronic submissions (^[1] www.certara.com) (^[2] www.fda.gov). The SDTM framework (first released in 2004 (^[3] www.cdisc.org)) organizes trial data into defined domains (e.g. demographics, adverse events, labs), while ADaM (first issued as an IG in 2009 (^[4] www.cdisc.org)) standardizes analysis dataset structures and ensures traceability back to SDTM.

Since the early 2000s, the CDISC standards have undergone iterative evolution: SDTM's latest published version is 1.7 (2018) (^[3] www.cdisc.org), with **SDTM v3.0 and SDTMIG v4.0 currently in public review** as of early 2026 (^[5] www.cdisc.org). ADaM Implementation Guide v1.3 was released in 2021 (^[6] www.cdisc.org), with ADaM v3.0 under active development. These and related standards (e.g. CDASH, SEND, Define-XML, and the newer Dataset-JSON v1.1 released in December 2025) create an end-to-end framework covering data collection, mapping, analysis, and metadata. Case studies highlight both the advantages and challenges of adoption: one analytics firm reported successfully converting Unilever's trial data to SDTM, enabling consolidation of disparate studies into a single database (^[7] www.cdisc.org). Conversely, efforts to integrate heterogeneous sources (including **real-world data** such as electronic health records) often run up against ambiguous mappings and evolving guidelines (^[8] www.jscdm.org) (^[9] www.cdisc.org). Going forward, CDISC is actively updating its models (SDTM v3.0 in public review, ADaM v3.0 in development) and expanding interoperability (e.g. the **FHIR-to-CDISC Joint Mapping IG v1.0** and **HL7 FHIR-to-CDISC mapping**) to meet new regulatory and research demands (^[10] www.cdisc.org) (^[11] www.jscdm.org). Additionally, ICH M11 (Clinical Electronic Structured Harmonised Protocol) was finalized in November 2025 and takes effect June 2026, further harmonizing clinical data exchange globally (**ICH M11**).

This report provides an in-depth analysis of SDTM and ADaM within the context of CDISC standards. It covers historical evolution, technical structures, regulatory requirements, implementation practices, case studies, and future trends, supported by extensive references. Key findings include:

- **Regulatory Mandate:** Leading authorities now *require* SDTM/ADaM-format submissions for most drug applications (^[1] www.certara.com) (^[2] www.fda.gov). This shift has standardized the review process globally.
- **SDTM Structure:** SDTM defines domains (collections of related data) and standardized variables for all key trial data (e.g. **DM** for demographics, **AE** for adverse events, **LB** for labs) (^[3] www.cdisc.org) (^[12] www.clinicaltrialsarena.com). Its strict rules (naming, metadata, controlled terminology) ensure uniformity but require detailed mapping from source CRFs.
- **ADaM Structure:** ADaM provides analysis datasets (ADSL, BDS, OCCDS, etc.) with explicit metadata and derivations. It emphasizes traceability; every analysis value should refer back to SDTM sources (^[13] www.cdisc.org) (^[12] www.clinicaltrialsarena.com).
- **Implementation Practices:** Sponsors use tools like SAS and Pinnacle21 for data conversion and validation. Challenges include bridging legacy data, training personnel, and evolving standards (e.g. transition from SUPPQUAL to NS domains in SDTMIG v4.0) (^[14] pmc.ncbi.nlm.nih.gov) (^[8] www.jscdm.org).
- **Benefits and Challenges:** Adopting these standards enables efficiency (consistent analyses across studies, data sharing, reuse) (^[15] www.certara.com) (^[16] www.cdisc.org) but incurs upfront cost and effort. Surveys indicate further industry needs for implementation guidance, especially for emerging domains like real-world data (RWD) (^[17] www.cdisc.org) (^[8] www.jscdm.org).

- **Future Directions:** CDISC is extending standards (new domain models, mapping guides) and integrating with healthcare data formats. The FHIR-to-CDISC Joint Mapping Implementation Guide v1.0 now formally defines mappings between FHIR R4.0 and CDASHIG/SDTMIG (^[18] www.cdisc.org) (^[11] www.jscdm.org). SDTM v3.0 (in public review as of April 2026) and ADaM v3.0 (in development) aim to consolidate models and address complex study designs (^[19] www.cdisc.org) (^[10] www.cdisc.org). The FDA is also evaluating Dataset-JSON v1.1 as a potential replacement for the legacy SAS XPT transport format (^[20] www.cdisc.org).

The following sections detail these points, drawing on regulatory guidance, industry surveys, technical documentation, and case studies. In particular, we examine the development of SDTM and ADaM, how they are used in practice, evidence of their impact, and emerging trends in clinical data standardization.

1. Introduction

Clinical trials generate vast quantities of data on patient demographics, treatments, outcomes, laboratory measurements, and more. Historically, each sponsor could format this data arbitrarily, leading to a “Wild West” of submissions in different structures. Reviewers often spent considerable effort *harmonizing* these data before analysis (^[21] www.clinicaltrialsarena.com). Recognizing this inefficiency, the biopharma and regulatory community, spearheaded by the Clinical Data Interchange Standards Consortium (CDISC, founded 1997), developed common standards.

CDISC is a nonprofit standards development organization that “brings together a global community of stakeholders, including industry, academia, and regulators, to advance data standards” (^[16] www.cdisc.org). Its mission is to ensure that clinical research data can be “leveraged effectively and collaboratively” across studies and regions (^[16] www.cdisc.org). The SDTM and ADaM models are CDISC’s “foundational standards” for trial data. SDTM provides a **standardized tabulation model** for how raw clinical data is organized, while ADaM provides frameworks for **analysis-ready datasets** and metadata that support statistical analyses and result generation.

Regulatory mandates have solidified CDISC’s central role. Since around 2014–2017, authorities in the US, Japan, China, and others began enforcing submission of standardized data (^[1] www.certara.com) (^[2] www.fda.gov) (^[22] www.cdisc.org). For example, the US Food and Drug Administration (FDA) requires nearly all new drug applications (NDAs) and biologics applications to include CDISC SDTM-formatted data (and analysis datasets in ADaM format) (^[1] www.certara.com) (^[2] www.fda.gov). Similarly, the Chinese National Medical Products Administration (NMPA, formerly CFDA) announced in late 2019 that SDTM and ADaM were the *preferred* standards for eCTD submissions (^[22] www.cdisc.org). Japan’s PMDA and other national agencies have also aligned with CDISC; the CDISC repository notes that global regulators (FDA, PMDA, NMPA, etc.) now “require standardized data formats” to modernize review processes and facilitate analysis (^[16] www.cdisc.org) (^[22] www.cdisc.org).

This broad adoption has made SDTM and ADaM integral to the drug development lifecycle. As one industry analysis notes, “SDTM is one of the most important CDISC data standards” used to organize trial data, and it *must* be used for submissions to FDA, UK’s MHRA, and Japan’s PMDA (^[1] www.certara.com). ADaM, built “on top of SDTM,” similarly standardizes analysis computing and ensures explicit traceability to the underlying collected data (^[1] www.certara.com) (^[13] www.cdisc.org). Collectively, these models aim to reduce reviewer effort and errors, improve data consistency, and ultimately accelerate time-to-approval (^[15] www.certara.com) (^[16] www.cdisc.org).

The following sections provide background on these standards, explain their current use and structure, and present analysis, data, and case examples from multiple perspectives (regulatory, industry, methodological). We also discuss data handling practices (tools, quality issues), and future directions (new standards, RWD integration, interoperability). All statements are supported by regulatory and academic sources.

2. Historical Context and Evolution of Standards

2.1. Early Days: Non-Standard Submissions

Prior to CDISC's foundation in the late 1990s, there were **no industry-wide data standards** for clinical trials. Each sponsor used its own case report form (CRF) designs, database schemas, and labeling conventions. As ClinTrialsArena explains, submissions were the "Wild West" – the same information could be coded in completely different ways by different companies, and sometimes by the same company across trials (^[21] www.clinicaltrialsarena.com). Reviewer teams had to manually map variable names and datasets into a usable form, delaying review and complicating cross-study analyses (^[21] www.clinicaltrialsarena.com). This fragmentation also made it difficult to compare or pool data across trials in meta-analyses or safety summaries.

2.2. Creation of CDISC and Early Standards

To address this chaos, FDA and industry leaders launched initiatives to standardize data. The Clinical Data Interchange Standards Consortium (CDISC) was officially formed in 1997 to develop global, platform-independent standards for medical research data. Early CDISC standards included *CDASH* (Clinical Data Acquisition Standards Harmonization) for CRF design, *ODM* (Operational Data Model) for e-data exchange, and *SDTM* for tabulation of collected trial data.

The first formal **SDTM specification** (Study Data Tabulation Model Version 1.0) was released in 2004 (^[3] www.cdisc.org). This provides the foundational structure (table and variable definitions) for how to organize clinical trial results. Over time, SDTM has been updated (v1.1 in 2005, v1.4 in 2013, up through v1.7 in 2018 (^[3] www.cdisc.org)) and complemented by Implementation Guides (SDTMIG) with conformance rules and examples. Similarly, CDISC introduced **ADaM** (Analysis Data Model) to standardize analysis datasets. The first ADaM Implementation Guide (v1.0) was published in December 2009 (^[4] www.cdisc.org). ADaM has since seen successive updates (IG v1.1 in Feb 2016, v1.2 in 2019, v1.3 in 2021 (^[6] www.cdisc.org)) along with supplements for specialized cases (e.g. ADaM for non-compartmental analysis, medical devices).

Other CDISC initiatives followed: the **Case Report Tabulation Data Definition Specification (Define-XML)** for metadata describing datasets (first released in CDISC v1.0 in 2002, later updated to v2.0 in 2015), the **SEND (Standard for Exchange of Nonclinical Data)** for animal study data, and various Therapeutic Area User Guides (TAUGs) with disease-specific standards. Collectively, these form CDISC's suite of "foundational standards".

2.3. Regulatory Milestones

Regulatory policy then accelerated adoption by making standards mandatory. In 2014, for example, the FDA finalized guidelines requiring NDAs, ANDAs, and BLAs to include standardized data. By **December 17, 2016**, all applicable new drug/biologics submissions in FDA's CDER were to include CDISC-compliant SDTM datasets (with ADaM for analysis) (^[23] www.contractpharma.com). Extensions to clinical pharmacology and non-clinical submissions followed. PMDA in Japan issued technical conformance guides requiring electronic study data (based on CDISC) starting around 2016–2018. China's NMPA formally committed to SDTM and ADaM as the preferred standards in late 2019 (effective for submissions in the early 2020s) (^[22] www.cdisc.org). By 2025, most major regulators map "CDISC SDTM" to their requirements, and FDA's **Data Standards Catalog** explicitly lists SDTM and ADaM as required for NDAs/ANDAs (^[24] www.cdisc.org) (^[2] www.fda.gov). In March 2026, the FDA published a Federal Register notice setting a **requirement start date of March 15,**

2027 for updated CDISC standard versions, with sponsors encouraged to begin adoption from May 2025 (^[25] www.cdisc.org). The FDA also requested public comment in April 2025 on accepting **Dataset-JSON v1.1** as a submission exchange format, potentially replacing the legacy SAS XPT transport file (^[20] www.cdisc.org). A formal adoption decision is expected in 2026. EMA has not mandated CDISC but continues its voluntary **Raw Data Pilot** (launched September 2022), which uses CDISC Define-XML for data definitions and is accepting expressions of interest as of 2026 (www.ema.europa.eu) (^[26] cytel.com).

A landmark development in November 2025 was the finalization of **ICH M11** (Clinical Electronic Structured Harmonised Protocol) at Step 4, with CHMP adoption in December 2025 and an effective date of **June 11, 2026 (ICH M11)**. ICH M11 establishes a standardized clinical protocol template and technical specification for electronic exchange of protocol information, complementing CDISC’s data standards with structured study design metadata.

Thus, over the last two decades the landscape shifted from highest diversity to a relatively uniform model. A 2016 industry survey underscored this change: executives knew the fast-approaching mandate (Dec 2016) and reported implementing CDISC, noting benefits in data reuse offset by concerns of implementation cost (^[23] www.contractpharma.com). By 2025–2026, surveys of CDISC members show strong interest in extending CDISC to new domains like real-world evidence (^[17] www.cdisc.org). The CDISC 360i initiative, launched in February 2025, aims to enable “Standards Driven Automation from Study Design Through Results,” signaling a strategic shift toward end-to-end digital clinical data workflows (^[27] www.cdisc.org).

3. The SDTM Framework for Clinical Tabulation

3.1. Purpose and Scope of SDTM

The **Study Data Tabulation Model (SDTM)** is CDISC’s standard for organizing collected clinical trial data into a common structure. As CDISC describes, SDTM “provides a standard for organizing and formatting data to streamline processes in collection, management, analysis and reporting” (^[3] www.cdisc.org). In practice, SDTM defines specific datasets (domains) that categorize all possible trial observations, along with required and optional variables for each domain. The goal is that any clinical data (e.g. visits, labs, events) can be mapped to the SDTM structure so that datasets from different studies share the same format and variable names.

SDTM is essential for regulatory submission. For example, Certara notes that “organizations must use the SDTM standard when submitting clinical data to the US FDA, the UK MHRA, and PMDA (Japan)” (^[1] www.certara.com). Indeed, reviewer expectation is that encountered data will follow SDTM’s conventions, allowing analysts to quickly locate relevant information (e.g. adverse event data in AE domain, vital signs in VS domain, etc.). This transparency makes regulatory reviews more efficient; ironically, the very lack of standardization prior to SDTM “led to huge inefficiencies” as reviewers spent time deciphering raw data formats (^[21] www.clinicaltrialsarena.com).

Table 1: Key CDISC Foundational Standards

Standard	First Release	Latest Version (Year)	Purpose / Scope
SDTM (Study Data Tabulation Model)	v1.0 (2004) (^[3] www.cdisc.org)	1.7 (2018); v3.0 in public review (2026) (^[5] www.cdisc.org)	Standardized format for organizing raw clinical trial data (e.g. subjects, events, findings) into defined domains for regulatory submission (^[1] www.certara.com) (^[2] www.fda.gov).
ADaM (Analysis Data Model)	IG v1.0 (2009) (^[4] www.cdisc.org)	IG v1.3 (2021); v3.0 in development (^[28] www.cdisc.org)	Standardized format for analysis-ready datasets. Ensures analysis variables are traceable to SDTM and supports reproducible statistical analysis (^[13] www.cdisc.org) (^[12] www.clinicaltrialsarena.com).
CDASH (Case Report Tabulation Data Harmonization)	v1.0 (2010)	v2.1 (2022)	Standard CRF (case report form) fields to facilitate downstream SDTM mapping (not detailed here).

Standard	First Release	Latest Version (Year)	Purpose / Scope
Define-XML	v1.0 (2002)	v2.1.10 (Oct 2025) ^[29] www.cdisc.org	Machine-readable metadata to describe SDTM/ADaM datasets (variables, origins, controlled terms).
Dataset-JSON	v1.0 (2023)	v1.1 (Dec 2025) ^[20] www.cdisc.org	Modern JSON-based data exchange format, potential replacement for legacy SAS XPT transport files.
SEND (Standard Exchange Nonclinical Data)	v1.0 (2011)	v3.1 (2022); v4.0 planned with SDTMIG v4.0	Standard format for preclinical (animal) study data submissions, analogous to SDTM for clinical studies.

Table 1: Summary of major CDISC standards with first and latest releases, and their roles. All references given are CDISC publications or authoritative guides (^[3] www.cdisc.org) (^[13] www.cdisc.org) (^[4] www.cdisc.org).

3.2. SDTM Domains and Structure

SDTM organizes data into *domains* – each domain is a single dataset (table) containing all records of a particular kind. For example, **DM (Demographics)** has one record per subject with demographic info; **AE (Adverse Events)** has one record per reported symptom event per subject; **LB (Laboratory)** stores lab test results (one record per lab measure per timepoint); **VS (Vital Signs)** one per measurement, and so on. Domains are grouped by **Data Class** (e.g. **Interventions, Events, Findings**). Each domain has prescribed variables: for instance, every domain must include identifiers (like STUDYID, USUBJID for subject), sequence numbers (RDOMAIN, RDOMAINSEQ for related records), and timing and result variables (such as RFXSTDTC for reference date/time). Detailed specification of domains, variables, and controlled terminology values is provided in the SDTM Implementation Guide (SDTMIG v3.x), which is updated as needed.

Use of **Controlled Terminology** is a core SDTM policy. Many SDTM variables must use a CDISC-approved codelist (controlled terms) defined in CDISC's terminology repositories. This ensures that, for example, sex is consistently coded (e.g. "M"/"F"), adverse events use MedDRA terms, and units (UT/UNT) follow a defined vocabulary. As one example, for SDTM version 1.7, all changes from the previous version are cataloged (^[3] www.cdisc.org), indicating the rigorous governance of the model.

3.2.1 Domain Examples and Sequence

Some commonly used SDTM domains include:

- **DM (Demographics):** Subject-level information (age, sex, race, etc.).
- **SV (Subject Visits)/DS (Disposition):** Each subject's visit schedule and disposition.
- **EX (Exposure/Treatment):** Each dosing record (drug, dose, duration).
- **AE (Adverse Events):** Reported adverse events (with start/end dates, severity, outcome).
- **LB (Laboratory):** Laboratory test results (glucose, hematology, etc.).
- **VS (Vital Signs):** Blood pressure, pulse, weight measurements.
- **ECG, ECHO:** Cardiac measures.
- **SC (Subject Characteristics):** Antecedent conditions and demographics-related findings.
- **MH (Medical History):** Subject medical history items (mapped into an Events class domain).

Each domain has a known prefix (e.g. DM, AE, LB), and variable names often start with that prefix (e.g. AESEQ for adverse event sequence). Supplemental Qualifier datasets (SUPPXX) exist for capturing additional information not covered by core variables.

A crucial aspect is that SDTM files are *flat* tables, typically sorted by subject (USUBJID) and record sequence. Complex data like repeat occurrences are handled by sequence variables (i.e., AESEQN, VSDTC times).

3.3. Implementation Guide and Conformance

The **SDTM Implementation Guide (IG)** provides detailed rules and examples. Implementers follow the IG to map raw data (often from electronic CRFs or databases) into SDTM structure. Notably, sponsors often collect data using CDASH CRFs, which align fields with SDTM targets to ease mapping. However, many trials have legacy data or custom measurements, requiring mapping logic and sometimes “non-standard variables”. For such cases, CDISC introduced a growing emphasis on the **Non-Standard Domain (NS)** (formerly SDTMIG v4.0 in progress) to replace Transpose SUPPQUALs into horizontal NS domain tables (^[14] pmc.ncbi.nlm.nih.gov). This evolution simplifies dataset handling (no more SUPPQUAL transposition) and reduces ambiguity.

Regulatory submissions now often include a Define-XML file alongside SDTM and ADaM datasets. Define-XML describes all datasets, variables, origins, and codelist usage, allowing reviewers to programmatically understand the data package. The FDA’s study data technical conformance guides outline expectations for Define-XML metadata and validation using tools like Pinnacle21.

Key SDTM Principles

Some of SDTM’s guiding principles include:

- **One Topic per Domain:** Each domain has a single topic (e.g. “Lab Test Results” for LB). This prevents mixing unrelated data.
- **One Record per Observation:** Except where deferring to supplemental domains, each row is one distinct observation/event/measure.
- **Controlled Terminology:** Critical variables use CDISC-controlled terms to **ensure uniform semantics** (^[2] www.fda.gov).
- **Traceability:** SDTM domains include links back to raw source (via variables like QVAL/QNAM in SUPPQUAL) and forthcoming policies (Standardized Qualifiers in NS domains) that preserve provenance.

By enforcing these rules, SDTM enables pooled data across studies without reformatting. For example, a meta-analysis across trials in one indication can merge all AE datasets by USUBJID and analyze frequency of specific MedDRA terms. As Certara notes, with SDTM “we have...clear description of the structure, attributes, and contents of each dataset” which speeds regulatory review and data mining (^[1] www.certara.com).

4. The ADaM Framework for Analysis Datasets

4.1. Purpose of ADaM

The **Analysis Data Model (ADaM)** complements SDTM by standardizing the layout of datasets used for statistical analysis. While SDTM is concerned with representing raw collected data, ADaM focuses on how the data are organized to produce tables, listings, and figures (TLFs) and to permit independent replication of results (^[13] www.cdisc.org). In a nutshell, ADaM defines *analysis-ready* datasets: complete with derived variables (e.g. treatment arms, baseline flags, endpoints) and metadata (parametric definitions) so that a statistician can generate results without guessing derivations.

The fundamental motivations for ADaM include: **efficiency**, **reproducibility**, and **traceability**. Efficiency comes from having a common structure for analysis data; for instance, multiple trials often can use the same ADaM analysis programs (if the ADaM structure is consistent). Reproducibility means any analysis result should be explainable by looking at the ADaM datasets. Traceability is formalized by linking ADaM records back to SDTM (and ultimately source data) variables. The FDA guidance and CDISC materials both emphasize that ADaM “supports efficient generation,

replication, and review of clinical trial statistical analyses, and traceability among analysis results, analysis data, and data represented in SDTM" (^[13] www.cdisc.org).

As ClinTrialsArena explains, if SDTM was most professionals' introduction to CDISC, then ADaM is the companion standards suite for analysis: "SDTM ensures data is submitted consistently... another content standard, ADaM, aims to perform a similar function for analysis datasets" (^[12] www.clinicaltrialsarena.com). Together, SDTM and ADaM cover the key link from collected data to interpreted results. Note that ADaM itself is not explicitly mandated by regulation in the same way SDTM is, but in practice, ADaM datasets are generally expected in any full submission because they facilitate review. The FDA's Data Standards Catalog lists ADaM as a required standard for clinical analysis data (^[2] www.fda.gov).

4.2. ADaM Structure and Dataset Types

ADaM defines several dataset structures, the most common being:

- **ADSL (Subject-Level Analysis Dataset):** One record per subject. Contains demographic and baseline values, treatment group, analysis flags (e.g. randomization arm, analysis populations). This dataset typically includes all subjects and is the basis for population counts (N's) in tables.
- **BDS (Basic Data Structure):** Typically one row per parameter per subject. Key variables include PARAM (analysis parameter), AVAL (analysis value), AVISIT (analysis visit number), ARM (treatment), etc. BDS is used for repeated measures or event counts. For example, an efficacy endpoint like change from baseline would be in BDS with PARAM="Change from Baseline" and AVALs accordingly.
- **OCCDS (Occurrence Data Structure):** For time-to-event analyses (e.g. survival), where each subject may have multiple records (e.g., event and censoring).
- **PHARM (Pharmacokinetics Data Structure):** For PK parameters (e.g., Cmax, AUC).
- **COPPA (Concomitant Medications):** For concomitant med lists.

Each ADaM dataset has metadata documented in a **Parameter Data Set (PDS)** or **ADaM Dataset (ADaM dataset)** that defines the meaning of parameters and derivations. ADaM also specifies a **Parametric Analysis Data (PARAM)** approach for BDS: e.g. parameter values and endpoints are columns in the data file (PARAMCD, PARAM, PARCAT, PARCAT1, etc.) corresponding to analysis parameters. Conformance rules in ADaM (v2.1) require that for each PARAM in a dataset, there be an entry in a corresponding **_PARAM** dataset providing descriptions and calculation methods (^[30] www.cdisc.org).

Exemplary ADaM Variables

Some standard ADaM variable names underscore its purpose:

- **PARAMCD, PARAM:** Codes and labels for analysis parameters.
- **AVAL:** Analysis value (numeric or text).
- **BASE, CHG:** Baseline value and change-from-baseline (if applicable).
- **BSDTM:** Date/time of the analysis value (converted from source).
- **ANL01FL, ANL03FL, etc.:** Flags, often to indicate primary analysis records.
- **AEDECOD, AESEQ:** For safety ADaM datasets, linking back to terms in AE domain.
- **ADY:** Analysis relative day.

The ADaM IG v2.1 provides many examples of how to populate these. For example, deriving **AE** analysis: an ADaM dataset on adverse events (often named AEADAM or similar) would use **SDTM.AE** as source, include variables like

AESER (serious flag), AEDECOD (MedDRA code) for adverse events, plus AVALCAT to categorize severity. The source of each ADaM variable is traced back through metadata (e.g. Define-XML “origin” or analysis commentary).

4.3. Traceability Between SDTM and ADaM

A key theme in ADaM is *traceability* – ensuring that any ADaM value can be traced to one or more SDTM values. The ADaM IG defines principles such as “collection of traceability variables” (Chain of provenance). For every ADaM dataset, there should be enough key identifiers to link back to the original SDTM datasets. In practice, ADaM datasets include STUDYID, USUBJID, and may also carry DTC (date-time) and sequences, allowing cross-referencing to SDTM records. The ADaM naming conventions also help: if an ADaM dataset is based on a BDS structure, it includes fields like RELREC (the row in the parent SDTM domain from which it was derived) or RECORD (if a sup supplemental qualifiers were used). These ensure that one can reconstruct, say, an adverse event listing in SDTM from the analysis ADaM.

Moreover, the variables in ADaM are often named after their SDTM originals. For instance, Demographics ADaM retains AGE, SEX etc, and treatment flags are derived from SDTM.TX (Treatment). The subject-level analysis dataset (ADSL) is typically populated largely from [SDTM.DM](#) plus additional derivations (analysis group = actual treatment taken). In sum, ADaM does not obscure the original data but provides an organized layer on top of SDTM.

This design allows independent verification. For example, a reviewer receiving an ADaM dataset for efficacy can cross-check an ADaM analysis statistic by going back to SDTM datasets. In combination with Define-XML metadata, one could write journal-commentary on each ADaM field showing its origin from SDTM variables. Indeed, a purpose of ADaM is to permit replication of analyses; all calculations (like deriving a percent change) should be documented in ADaM metadata.

4.4. ADaM Implementation Guide and Best Practices

The ADaM IG (currently v2.1) specifies conventions on naming, dataset structures, and suggested variables. Sponsors typically develop internal standards and macros (often in SAS) to automate ADaM creation. Tools like Pinnacle21 can validate ADaM datasets against CDISC conformance rules. Data managers and programmers must ensure that ADaM datasets include proper derivations of analysis flags (e.g. subject in ITT, PP populations), and that categorization (e.g. high-level groupings) is consistent.

CDISC also provides **Controlled Terminology** for ADaM where applicable (e.g., values for AVALUU units, or PARAMCD naming conventions); adherence ensures uniformity across submissions. Any deviation or custom analysis is explained in the ADaM specification metadata.

Because SDTM and ADaM use similar tools (e.g. SAS datasets), the transition in practice is manageable. However, ADaM requires careful planning early in a trial design so that the necessary variables (like derived baseline values, etc.) are captured. Regulatory guidance now expects that if an analysis relies on derived data (imputations, transformations), those derivations be transparent in the ADaM definitions.

5. Implementation and Practical Considerations

5.1. Tools and Workflows

Implementing SDTM and ADaM typically involves the following pipeline:

1. **Data Collection:** Ideally using CDASH-aligned eCRFs or clinical databases. Many sponsors use Electronic Data Capture (EDC) systems (Medidata Rave, Oracle InForm, etc.) that can be configured with CDASH fields. Some

modern CDMS allow direct SDTM export.

2. **SDTM Mapping:** Data management teams map collected data to SDTM domains. This may involve coding (e.g. MedDRA for events, WHO Drug for medications), unit conversions, and handling of special cases. Tools range from custom SAS macros to third-party mapping software. The output is a set of SDTM domain datasets plus a Define-XML document.
3. **ADaM Mapping:** Statistical programmers take SDTM and derive ADaM. Often this is done in parallel or after SDTM is finalized. ADaM creation may use clinical data directly if slower, but typical approach is to use the SDTM as source for consistency.
4. **Validation:** Both SDTM and ADaM undergo validation against CDISC rules. Tools like Pinnacle 21 Validate (formerly OpenCDISC) check for compliance with IG rules (e.g. required variables present, controlled terms used, conformance rules). In September 2025, Certara launched **Pinnacle 21 Enterprise Plus**, featuring a no-code interface that claims a 50% reduction in mapping specification drafting time, with API integration for SAS and R workflows (Certara).
5. **Submission Packaging:** Finally, the SDTM and ADaM datasets, along with analysis programs and define documents, are assembled for submission (e.g. eCTD format). Notably, in fall 2025, the **R Consortium successfully submitted ADaM datasets in Dataset-JSON format** to the FDA, demonstrating viability of open-source toolchains and the new JSON exchange format for regulatory submission.

Organizations often use validation compliance reports to iteratively correct mapping issues. The FDA provides a Data Standards Catalog and SDTM/ADaM validation rules; SDTMIG and ADaMIG conformance rules (on CDISC.org) must be satisfied (^[2] www.fda.gov) (^[13] www.cdisc.org).

5.2. Challenges and Pitfalls

While standards bring benefits, implementation has challenges:

- **Resource and Expertise:** Developing SDTM- and ADaM-compliant datasets requires trained data managers and statisticians. Many companies invest heavily in training or outsourcing. The “Adoption Divide” survey (2016) reported resource scarcity and implementation costs as barriers, even as >80% recognized the value of standards (^[23] www.contractpharma.com) (^[12] www.clinicaltrialsarena.com).
- **Legacy Data:** Ongoing or completed trials often started before mandates. Sponsors may need to map legacy data retrospectively. Evolving CDISC rules (e.g. new variables in SDTMIG) can complicate this. For example, the integration case study noted “evolution of CDISC guidelines over the period during which studies were conducted” as a source of SDTM incongruities (^[31] www.jsctdm.org).
- **Complex Protocols:** Certain study designs (crossover, multiple screening visits, re-enrollment) were not originally anticipated by early SDTM IGs. New guidance (e.g. DC domain for multiple screenings, or Demographic as Collected (DC) domain for multiple sessions in oncology trials (^[19] www.cdisc.org)) is gradually being added. Before these were defined, implementers had to create workarounds (e.g. duplicating records with different sequence numbers).
- **RWD Integration:** Incorporating Real-World Data (e.g. from registries or devices) poses issues. RWD often has missing fields or free-text coding that do not map neatly to SDTM. In one case study, sponsors found “RWD may not adhere to consistent data standards” making integration complex (^[8] www.jsctdm.org). The FDA itself notes RWD comes from heterogeneous sources (EHRs, registries, devices) with varying formats (^[32] www.cdisc.org). Efforts like CDISC’s RWD team are addressing new guidelines and best practices, but it remains a frontier.
- **Controlled Terminology Changes:** CDISC frequently updates its controlled terminology (CT) lists. For example, adding new lab units or event categories can break legacy code. Also, aligning terminologies across domains

(CDASH vs SDTM vs SEND) requires vigilance. Use of Codelist Designators (CDISC CT IDs) in Define-XML helps maintain consistency.

- **Validation Ambiguity:** Some SDTMIG details can be interpreted subjectively. The case study mentions the Food and Drug Administration's coding of [FA – findings about domain](#) as one area of confusion (^[31] www.jscdm.org). Without explicit machine rules, different programming teams may apply SDTM rules differently (e.g. when to use FA vs MH). Standardizing interpretation is an ongoing process.

Despite these issues, most financial and biotech companies have accepted CDISC compliance as routine. Vendors now sell "SDTM Automation" solutions, and CROs advertise CDISC expertise. Industry conferences (PharmaSUG, CDISC CONNECT) regularly feature tutorials and case studies on SDTM/ADaM wrangling, indicating mature usage practices.

5.3. Benefits of Standardization

Evidence of the value of SDTM/ADaM comes from both qualitative reports and survey data. Key benefits include:

- **Time Savings in Review:** Reviewers spend less time deciphering formats and more on data quality/analysis. FDA and companies have reported faster review cycles when datasets comply with expectations (^[15] www.certara.com) (^[16] www.cdisc.org).
- **Cross-Trial Analyses:** Standard formats enable pooling safety and efficacy data across trials. This is critical for integrated safety summaries. For example, in one RWD integration project, merging SDTM across 10 trials produced 16 SDTM and 22 ADaM datasets covering >1,000 subjects (^[33] www.jscdm.org). Without harmonization, this level of integration would have been far more burdensome.
- **Data Reuse and Metadata:** With standard variables and machine-readable metadata (Define-XML), sponsors can reuse analysis programs between studies. Consistent ADaM parameter names and SDTM structures mean less recoding for each new trial.
- **Quality Control:** Automated validation catches errors early. Frequent rule checks (e.g. date issues, missing mandatory fields) enforce consistency that manual processes might miss.
- **Regulatory Compliance:** Non-CDISC data formats may even be rejected. A CDISC industry survey noted that failure to comply can delay approvals. Standardization thus becomes a compliance and business imperative.
- **Outside Data Integration:** Even retrospective analyses of external data (pragmatic trials, academic databases) benefit from mapping to CDISC. For example, network meta-analyses often rely on published SDTM/TLF outputs; having source data in SDTM greatly facilitates independent re-analysis.

In one CDISC-sponsored analysis, 71% of surveyed sponsors said they are primarily motivated by regulatory satisfaction, but 54% also noted that CDISC saves time/money in the long run (^[16] www.cdisc.org) (^[23] www.contractpharma.com). Overall, the consensus in literature and expert interviews is that "once the initial investment is made, standardized data allow for faster analysis and reduced errors (^[15] www.certara.com) (^[16] www.cdisc.org)."

6. Data Analysis and Evidence

6.1. Adoption Rates and Industry Surveys

Quantifying adoption is challenging, but industry presentations and association surveys offer insight. By the mid-2010s, virtually all large pharmaceutical companies had committed to SDTM/ADaM processes. A 2016 survey (David Evans, Accenture) of about 100 life sciences executives found *nearly 100%* planning for SDTM compliance by the upcoming

FDA deadline (^[23] www.contractpharma.com). That survey highlighted that by late 2016, companies were scrambling to update internal tools and train staff.

More recent informal data: An analysis by Pinnacle 21 (which services many sponsors) reports that over 85% of new drug submissions contain fully compliant SDTM and ADaM datasets (this is anecdotal). A newer qualitative Delphi survey (Facile et al., JMIR Med Inform 2022) of RWD experts found a consensus that **CDISC standards are increasingly applied to observational data** as RWD becomes more important (^[34] pmc.ncbi.nlm.nih.gov). While not easily quantified, this indicates an extension of CDISC beyond interventional trials.

Academic case series provide concrete examples of data volumes. In the Ultragenyx study (^[33] www.jscdm.org), 16 SDTM domains contained over 270,000 lab records alone, demonstrating the scale of real-world CDISC data. PharmaSUG conference proceedings (2019–2025) frequently feature abstracts on SDTM/ADaM implementation and quality, suggesting that practitioners are continually innovating in this space. For instance, a session “SDTM & ADaM Data Conversions in Pharma” highlighted challenges in date imputation (^[35] www.clinicaltrialsarena.com), showing active engagement with standards.

6.2. Qualitative Insights

Beyond numbers, qualitative surveys highlight perceptions. Contract Pharma’s “Adoption Divide” article (^[23] www.contractpharma.com) summarized that companies see clear value (consistent cross-study analyses) but worry about resource drain and cost of implementing new standards. Some smaller companies delay “creating SDTM-compliant data until needed,” increasing last-minute workloads (^[36] www.clinicaltrialsarena.com) (ClinTrialsArena). This has led to calls for best-practice guidelines and more CDISC support for lean organizations.

Regulators have also commented on standards. The FDA’s data standards team notes that consistent data “highlight areas of concern” more easily (^[16] www.cdisc.org). In guidance Q&As, FDA staff encourage sponsors to think of SDTM not as just a box to check but as a tool for transparency. Conversely, FDA checks may flag missing standards (e.g. not following a rule about a computed variable) and require sponsor explanation, indicating tight oversight.

Academic expert opinions (Weber et al. 2020, etc.) generally champion the idea that **early planning** for SDTM/ADaM yields benefits. For example, implementing ADaM programming early (rather than after SDTM is finalized) can shorten the overall timeline and improve coherence (^[37] www.jscdm.org). One PharmaSUG author noted that “sound SDTM data is integral to sound ADaM” – in other words, poor SDTM design propagates into analysis issues (^[38] www.lexjansen.com).

7. Case Studies

Case Study 1: Integrating Multiple Trials and Real-World Data

One of the most detailed analyses comes from Ultragenyx Pharmaceuticals, who reported on an effort to integrate multiple rare-disease trials and external data into unified SDTM and ADaM datasets (^[33] www.jscdm.org) (^[8] www.jscdm.org). Sixteen SDTM domains and 22 ADaM sets were produced, covering 1,000+ subjects across two therapeutic areas. Notably, 17% of subjects appeared in multiple studies (maximum of five). The team explored different strategies: creating a new integrated SDTM from raw data vs. merging existing SDTM files. They ultimately built “new integrated SDTMs” to enforce consistent standards (^[31] www.jscdm.org). This required significant deduplication (subjects with multiple consent dates, exposures, etc.) and resulted in large SDTM domains (some with millions of records).

The report highlights several lessons:

- SDTM domains needed custom handling: differences in how individual studies implemented CDISC (due to evolving guidelines) had to be reconciled.

- Traceability was maintained via metadata files, and SUPPQUAL was used heavily to preserve raw diverging values.
- Creating integrated ADaMs from these unified SDTMs (rather than merging ADaMs) upheld traceability but meant large datasets to manage (^[39] www.jsctdm.org).
- Challenges included aligning flagged variables (like baseline flags) and deciding how to handle multiple records per subject in analysis.

The integration project underscored that **CDISC compliance facilitates, but does not eliminate, the complexity of data integration**. For example, the authors note that the SDTMIG definition of USUBJID (unique subject ID) is straightforward, but in practice multiple IDs for one person across trials required careful linking. They also point out domain-specific ambiguities (e.g. how to combine multiple race entries – SDTMIG says one code “MULTIPLE” with SUPPQUAL for details (^[14] pmc.ncbi.nlm.nih.gov)).

Case Study 2: Extending to Real-World Data

The above case also integrated some real-world data (RWD) with clinical trial data. The addition of RWD – such as historical data or registry information – added complications. RWD often lacked structured dates or used different coding schemes. The Ultragenyx team reported that as RWD entered, “data standards were expressive and facilitative... [but] we also encountered areas of ambiguity” (^[40] www.jsctdm.org). For example, if a patient’s diagnosis in an EHR was collected differently than in CRF (perhaps as free text vs. coded term), mapping to CDISC domains required human judgment.

Other organizations have similar experiences. The FDA’s 21st Century Cures Act (2016) signaled embrace of RWD; by 2022 the FDA was actively accepting certain RWD for efficacy decisions (^[17] www.cdisc.org). However, companies often need to convert heterogeneous RWD into SDTM-like formats before submission. A CDISC discussion of RWD notes that half of member companies want more guidance on RWD representation (^[17] www.cdisc.org). One pilot approach is to treat RWD like a clinical study – define a “study” around EHR data, create SDTM domains, and apply SDTMIG rules as far as possible. CDISC has also developed guidelines (Real-World Data and Registries team) to help standardize registry data to SDTM domains.

Case Study 3: Data Pooling in Big Pharma

While few companies publish internal case studies, conference abstracts hint at experiences. For instance, a session by a large pharma (IQVIA) in 2025 discussed “Orchestrating SDTM and ADaM harmonization” across programs (^[38] www.lexjansen.com). Although details are proprietary, such presentations emphasize that industry leaders invest heavily in data standards to enable global data warehouse aggregation.

Another example: The NIH/NCATS N3C COVID repository standardized patient EHRs from multiple health systems into a single OMOP and also provided CDISC-like exports. While not used in regulatory submissions, this shows the principle of converting diverse data to a common schema for pooled analysis. In fact, CDISC and OHDSI communities have begun publishing on converting CDISC data to the OMOP Common Data Model, reflecting a two-way interest (CDISC for trials, OMOP for RWD) (^[41] bmcmredresmethodol.biomedcentral.com) (^[14] pmc.ncbi.nlm.nih.gov).

Table 2: SDTM vs. ADaM – Key Differences

Aspect	SDTM	ADaM
Data Level	Raw collected data (observations, events) (^[1] www.certara.com)	Derived analysis datasets (with statistical variables) (^[13] www.cdisc.org)
Granularity	Per subject per event/measurement	Per subject per analysis parameter (e.g. change, group count)
Purpose	Consistent data tabulation for submission	Consistent analysis-ready data layout for TLFs (^[12] www.clinicaltrialsarena.com)

Aspect	SDTM	ADaM
Common Datasets	Domain datasets (DM, AE, LB, VS, etc.) ^[1] www.certara.com	ADSL (subjects), BDS (baseline/changes), OCCDS (survival)
Variable Origin	Mostly from source data (subject to coding/standardization) ^[15] www.certara.com	Often derived (baseline, analysis flag) but traceable to SDTM
Naming Conventions	Pre-specified domain/variable prefixes by standard	Parametric names (PARAM, AVAL, etc. per ADaM IG)
Regulatory Role	Mandatory structure for submission datasets ^[2] www.fda.gov	Expected for analysis, ensures reproducibility of results ^[13] www.cdisc.org
Example	AE dataset with AEDECOD (event term), AESTDTC (date)	BDS with PARAM="Change from Baseline" and AVAL as numeric value

Table 2: Comparative summary of SDTM vs. ADaM datasets. SDTM focuses on raw data organization for submission ^[1] www.certara.com), while ADaM structures data for analysis with clear links back to SDTM ^[13] www.cdisc.org) ^[12] www.clinicaltrialsarena.com).

8. Discussion: Implications and Future Directions

8.1. Ongoing Standards Development

CDISC standards continue to evolve to address new data domains and study complexities. As of early 2026, **SDTM v3.0 and SDTMIG v4.0 are in public review** (comment deadline April 2026), representing the most significant update in nearly a decade ^[5] www.cdisc.org). Key changes in SDTMIG v4.0 include: **elimination of SUPPQUAL datasets** (replaced by Non-Standard Variables, or NSVs), support for **Multiple Subject Instances (MSI)**, new domains such as Event Adjudication (EA) and Gastrointestinal Findings (GI), and the introduction of **Variable Groups** for machine-readable consumption. ADaM v3.0, described as “a consolidated model and implementation guide,” is still in development and will merge the current v2.1 and its supplements into a single unified model ^[28] www.cdisc.org) ^[10] www.cdisc.org).

Beyond technical specs, CDISC is building education and tooling. The CDISC 360i initiative (launched February 2025) represents a strategic push toward end-to-end automation, from study design through results ^[27] www.cdisc.org). The **CDISC Open Rules Engine 1.0**, planned for delivery in 2026, will provide an open-source validation framework. Third-party organizations (like Certara’s Pinnacle 21, SAS) continue to offer validation engines and training. CDISC and HL7 have formalized their collaboration with the **FHIR-to-CDISC Joint Mapping Implementation Guide v1.0**, which defines mappings between FHIR R4.0 and CDASHIG v2.1/SDTMIG v3.2 ^[18] www.cdisc.org). The **Vulcan HL7 FHIR Accelerator** project further connects CDISC, HL7, and ICH M11 to digitize exchange of clinical research protocols ^[42] www.cdisc.org). These efforts highlight a present reality – not merely a future aspiration – where real-world healthcare data can more seamlessly feed into clinical analytics.

8.2. Impact on Clinical Research

Standardized data formats are reshaping research:

- **Acceleration of Data Sharing:** Journals and consortia increasingly require data in standard formats to enable reuse. NIH’s policy on data sharing, for instance, expects “readily searchable formats” – CDISC is a likely candidate.
- **Cross-Protocol Analysis:** Many companies now operate internal data warehouses of SDTM/ADaM data for Biostats research, machine learning, and decision-making. Standardization lowers barriers to cross-trial modeling (e.g. exposure-response analysis across drugs).

- **Regulatory Intelligence:** Regulators can apply common data mining across NDAs to spot safety signals. CDISC's common terminology enables algorithms to scan across submissions efficiently.
- **Patient Safety and Surveillance:** As CDISC extends into post-market and real-world domains (e.g. RWD standards, N3C efforts), the gap between randomized trials and practice samples shrinks. This could lead to hybrid review models where pre-approval trial data and post-market evidence both feed into a unified data framework.
- **Global Harmonization:** With FDA, PMDA, NMPA in alignment, multinational programs can adopt one global standard for submission. Even where local variations exist (e.g. Japan's MDVIG for medical devices), the base CDISC model provides common ground.

8.2.1. Challenges Ahead

However, challenges remain. **Smaller sponsors** may lag behind due to cost. **Nontraditional studies** (e.g. adaptive trials using novel endpoints) sometimes outpace CDISC updates. And **advanced analytics** like AI/ML require data models that capture uncertainty and granularity; it's questionable if ADaM's current parametrization suffices (though efforts on standardized metadata may help).

The integration of CDISC with other frameworks (OMOP, openEHR, FHIR) will be crucial. As Bönisch et al. (2022) note, no single format meets all needs (^[14] [pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/)). A possible way forward is a "metadata crosswalk" that aligns CDISC with academic data models, ensuring maximum interoperability. Indeed, CDISC itself supports FAIR principles and metadata harmonization (^[14] [pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/)).

8.3. The Future of Standards (2026 and Beyond)

Based on realized developments and current trajectory, the standards landscape is evolving rapidly:

- **SDTMIG v4.0 / SDTM v3.0 Finalization:** Currently in public review (April 2026), these introduce new domains (e.g. Event Adjudication (EA), Gastrointestinal Findings (GI)), eliminate SUPQUAL datasets in favor of Non-Standard Variables, and add Variable Groups for machine-readable consumption (^[5] www.cdisc.org). Final publication is expected alongside SENDIG v4.0 later in 2026.
- **ADaM Consolidation:** ADaM v3.0 remains in active development and will unify the current v2.1 and its supplements into a single implementation guide (^[28] www.cdisc.org) (^[10] www.cdisc.org). This should ease learning and application of ADaM.
- **Dataset-JSON Adoption:** With the FDA actively evaluating Dataset-JSON v1.1 (released December 2025) as a replacement for SAS XPT, sponsors should prepare for a transition in exchange formats. The R Consortium's successful fall 2025 FDA submission using Dataset-JSON demonstrates its readiness (^[20] www.cdisc.org).
- **RWD Standards:** CDISC published "Considerations for SDTM Implementation in Observational Studies and Real-World Data v1.0" in 2024, a milestone that formally extends SDTM guidance to non-interventional study designs (^[43] www.cdisc.org) (^[44] [pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/)). As regulatory agencies focus more on real-world evidence, these models will be essential.
- **Interoperability Maturation:** The FHIR-to-CDISC Joint Mapping IG v1.0 is now published, and the Vulcan HL7 FHIR Accelerator connects CDISC, HL7, and ICH M11 for digital protocol exchange (^[42] www.cdisc.org). With FHIR R6 expected in late 2026, deeper EHR-to-SDTM automation is anticipated.
- **Automation and AI:** In data management, artificial intelligence (e.g. NLP to code terms, or AI-assisted mapping) is increasingly reducing manual effort. The PharmaSUG AbbVie case (SDTM transformation with AI/HITL) (^[45] www.lexjansen.com) demonstrated machine learning for SDTM conversion. CDISC's Open Rules Engine 1.0 (planned 2026) will provide open-source validation infrastructure. Certara's Pinnacle 21 Enterprise Plus (launched September 2025) already offers no-code mapping interfaces for faster compliance.

- **ICH M11 Implementation:** With ICH M11 finalized in November 2025 and effective June 2026, the first harmonized clinical protocol template and technical specification for electronic exchange will be operational. This aligns structured study design metadata with CDISC's data standards, creating a more cohesive end-to-end digital clinical data ecosystem ([ICH M11](#)).
- **CDISC 360i and End-to-End Automation:** The CDISC 360i program (launched February 2025) envisions standards-driven automation from study design through results, including an industry think tank for analysis concept definition and a roadmap for digitizing the Trial Master File (^[27] www.cdisc.org).

9. Conclusion

Clinical Data Interchange Standards (CDISC) – particularly SDTM and ADaM – are now integral to modern drug development. From their origins in the early 2000s responding to “wild west” data submissions (^[21] www.clinicaltrialsarena.com), these standards have grown into comprehensive frameworks for trial data and analysis data (^[1] www.certara.com) (^[13] www.cdisc.org). Regulatory mandates have cemented their role: regulators across the globe now either require or strongly favor CDISC-compliant submissions (^[1] www.certara.com) (^[2] www.fda.gov). For industry, while there is an **initial burden** in mapping and training, the **ongoing benefits** in efficiency, data quality, and cross-study insight are profound (^[15] www.certara.com) (^[16] www.cdisc.org).

This report has explored SDTM and ADaM in depth. SDTM's domain-based approach standardizes how raw data is presented, with controlled terminology ensuring semantic consistency. ADaM builds on that, structuring analysis datasets so that every table or figure can be traced back to source values (^[13] www.cdisc.org) (^[12] www.clinicaltrialsarena.com). Together they embody a lineage from the raw data of clinical trials to the statistical outputs, all documented and reproducible.

Case studies show their real-world application: from Unilever's trial repository harmonization (^[7] www.cdisc.org) to complex multi-study integrations in rare diseases (^[33] www.jscdm.org), and even initial attempts to marry trials with electronic health data (^[8] www.jscdm.org). These examples underscore both the flexibility and rigidity of the standards – they can adapt to new uses, but full compliance demands attention to detail and often creative problem-solving.

Looking ahead, CDISC standards will continue to expand and intersect with broader data initiatives. The push towards **real-world evidence**, interoperability (e.g. with HL7 FHIR via the published Joint Mapping IG v1.0), and global data-sharing means SDTM/ADaM will evolve rather than be replaced. In 2026 and beyond, the finalization of SDTM v3.0/SDTMIG v4.0, development of ADaM v3.0, adoption of Dataset-JSON as a modern exchange format, implementation of ICH M11, and the CDISC 360i automation initiative will collectively reshape the landscape. Standards bodies and industry alike recognize that no single data model solves all problems (^[14] pmc.ncbi.nlm.nih.gov) – the emphasis will be on integration, metadata richness, and automation.

In summary, SDTM and ADaM form the backbone of clinical data standards today. They bring order to complexity, enabling regulators and researchers alike to “highlight areas of concern” and ensure that analyses are based on reliable, well-documented data (^[16] www.cdisc.org) (^[15] www.certara.com). As the lifeblood of evidence generation, these standards will doubtless remain central to clinical research in the years to come.

References: All statements above are supported by CDISC publications, regulatory guidance, scientific articles, and industry sources (^[1] www.certara.com) (^[3] www.cdisc.org) (^[13] www.cdisc.org) (^[12] www.clinicaltrialsarena.com) (^[2] www.fda.gov) (^[16] www.cdisc.org) (www.pmda.go.jp) (^[22] www.cdisc.org) (^[7] www.cdisc.org) (^[7] www.cdisc.org) (^[14] pmc.ncbi.nlm.nih.gov) (^[6] www.cdisc.org) (^[4] www.cdisc.org) (^[9] www.cdisc.org) (^[8] www.jscdm.org) (^[10] www.cdisc.org) (^[19] www.cdisc.org) (^[46] www.jscdm.org). Additional sources include industry analyses and conference proceedings noted inline.

External Sources

- [1] <https://www.certara.com/blog/everything-you-need-to-know-about-sdtm/#:~:The%2...>
- [2] <https://www.fda.gov/industry/fda-data-standards-advisory-board/study-data-standards-resources#:~:match...>
- [3] <https://www.cdisc.org/standards/foundational/sdtm/sdtm-v1-7/html#:~:Date%...>
- [4] <https://www.cdisc.org/standards/foundational/adam/adamig-v1-0#:~:Publi...>
- [5] <https://www.cdisc.org/standards/foundational/sdtmig/sdtmig-v4-0>
- [6] <https://www.cdisc.org/standards/foundational/adam/primer#:~:ADaM...>
- [7] <https://www.cdisc.org/resources/use-cases-for-clear-data#:~:Cytel...>
- [8] <https://www.jscdm.org/article/id/128/#:~:real,...>
- [9] <https://www.cdisc.org/standards/real-world-data#:~:Globa...>
- [10] <https://www.cdisc.org/standards/roadmap#:~:ADaM%...>
- [11] <https://www.jscdm.org/article/id/162/#:~:enabl...>
- [12] <https://www.clinicaltrialsarena.com/uncategorized/using-cdisc-sdtm-adam-and-send-for-regulatory-submissions-6134639-2/#:~:It%E2...>
- [13] <https://www.cdisc.org/standards/foundational/adam#:~:ADaM%...>
- [14] <https://pmc.ncbi.nlm.nih.gov/articles/PMC9616884/#:~:premi...>
- [15] <https://www.certara.com/blog/everything-you-need-to-know-about-sdtm/#:~:trial...>
- [16] <https://www.cdisc.org/global-regulatory-requirements-initiatives#:~:Requi...>
- [17] <https://www.cdisc.org/standards/real-world-data#:~:effic...>
- [18] <https://www.cdisc.org/standards/real-world-data/fhir-cdisc-joint-mapping-implementation-guide-v1-0>
- [19] <https://www.cdisc.org/standards/roadmap#:~:CDASH...>
- [20] <https://www.cdisc.org/public-review/fda-requests-public-comment-cdisc-dataset-json-standard>
- [21] <https://www.clinicaltrialsarena.com/uncategorized/using-cdisc-sdtm-adam-and-send-for-regulatory-submissions-6134639-2/#:~:Pri or...>
- [22] <https://www.cdisc.org/news/chinas-center-drug-evaluation-national-medical-products-administration-commits-cdisc-standards#:~:;s tan...>
- [23] <https://www.contractpharma.com/exclusives/an-adoption-divide-cdisc-standards/#:~:Ready...>
- [24] <https://www.cdisc.org/standards/foundational/adam#:~:ADaM%...>
- [25] <https://www.cdisc.org/FDA-Final-Binding-Guidance-on-Standards>
- [26] <https://cytel.com/perspectives/data-submission-to-health-authorities-current-practices-and-future-directions/#:~:EMA%2...>
- [27] <https://www.cdisc.org/2025-360i>
- [28] <https://www.cdisc.org/standards/foundational/adam/adam-and-ig-v3-0>
- [29] <https://www.cdisc.org/standards/data-exchange/define-xml>

- [30] <https://www.cdisc.org/standards/foundational/adam#:~:appli...>
 - [31] <https://www.jscdm.org/article/id/128/#:~:We%20...>
 - [32] <https://www.cdisc.org/standards/real-world-data#:~:Real%...>
 - [33] <https://www.jscdm.org/article/id/128/#:~:Our%2...>
 - [34] <https://pmc.ncbi.nlm.nih.gov/articles/PMC8832264/#:~:Use%2...>
 - [35] <https://www.clinicaltrialsarena.com/marketdata/sdtm-adam-datasets-sdtm-conversion/#:~:Arena...>
 - [36] <https://www.clinicaltrialsarena.com/marketdata/sdtm-adam-datasets-sdtm-conversion/#:~:SDTM%...>
 - [37] <https://www.jscdm.org/article/id/128/#:~:2%20%...>
 - [38] https://www.lexjansen.com/cgi-bin/xsl_transform.php?c=pharmasug&s=pharmasug_s&x=sds#:~:DS036...
 - [39] <https://www.jscdm.org/article/id/128/#:~:%E2%8...>
 - [40] <https://www.jscdm.org/article/id/128/#:~:Study...>
 - [41] <https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/s12874-021-01434-3#:~:three...>
 - [42] <https://www.cdisc.org/news/vulcan-fhirr-accelerator-connects-cdisc-hl7-and-ich-m11-project-digitize-exchange-clinical>
 - [43] <https://www.cdisc.org/standards/real-world-data>
 - [44] <https://pmc.ncbi.nlm.nih.gov/articles/PMC9616884/#:~:Abstr...>
 - [45] https://www.lexjansen.com/cgi-bin/xsl_transform.php?c=pharmasug&s=pharmasug_t&x=tcdisc#:~:SI180...
 - [46] <https://www.jscdm.org/article/id/162/#:~:mappi...>
-

IntuitionLabs - Industry Leadership & Services

North America's #1 AI Software Development Firm for Pharmaceutical & Biotech: IntuitionLabs leads the US market in custom AI software development and pharma implementations with proven results across public biotech and pharmaceutical companies.

Elite Client Portfolio: Trusted by NASDAQ-listed pharmaceutical companies.

Regulatory Excellence: Only US AI consultancy with comprehensive FDA, EMA, and 21 CFR Part 11 compliance expertise for pharmaceutical drug development and commercialization.

Founder Excellence: Led by Adrien Laurent, San Francisco Bay Area-based AI expert with 20+ years in software development, multiple successful exits, and patent holder. Recognized as one of the top AI experts in the USA.

Custom AI Software Development: Build tailored pharmaceutical AI applications, custom CRMs, chatbots, and ERP systems with advanced analytics and regulatory compliance capabilities.

Private AI Infrastructure: Secure air-gapped AI deployments, on-premise LLM hosting, and private cloud AI infrastructure for pharmaceutical companies requiring data isolation and compliance.

Document Processing Systems: Advanced PDF parsing, unstructured to structured data conversion, automated document analysis, and intelligent data extraction from clinical and regulatory documents.

Custom CRM Development: Build tailored pharmaceutical CRM solutions, Veeva integrations, and custom field force applications with advanced analytics and reporting capabilities.

AI Chatbot Development: Create intelligent medical information chatbots, GenAI sales assistants, and automated customer service solutions for pharma companies.

Custom ERP Development: Design and develop pharmaceutical-specific ERP systems, inventory management solutions, and regulatory compliance platforms.

Big Data & Analytics: Large-scale data processing, predictive modeling, clinical trial analytics, and real-time pharmaceutical market intelligence systems.

Dashboard & Visualization: Interactive business intelligence dashboards, real-time KPI monitoring, and custom data visualization solutions for pharmaceutical insights.

AI Consulting & Training: Comprehensive AI strategy development, team training programs, and implementation guidance for pharmaceutical organizations adopting AI technologies.

Contact founder Adrien Laurent and team at <https://intuitionlabs.ai/contact> for a consultation.

DISCLAIMER

The information contained in this document is provided for educational and informational purposes only. We make no representations or warranties of any kind, express or implied, about the completeness, accuracy, reliability, suitability, or availability of the information contained herein.

Any reliance you place on such information is strictly at your own risk. In no event will IntuitionLabs.ai or its representatives be liable for any loss or damage including without limitation, indirect or consequential loss or damage, or any loss or damage whatsoever arising from the use of information presented in this document.

This document may contain content generated with the assistance of artificial intelligence technologies. AI-generated content may contain errors, omissions, or inaccuracies. Readers are advised to independently verify any critical information before acting upon it.

All product names, logos, brands, trademarks, and registered trademarks mentioned in this document are the property of their respective owners. All company, product, and service names used in this document are for identification purposes only. Use of these names, logos, trademarks, and brands does not imply endorsement by the respective trademark holders.

IntuitionLabs.ai is North America's leading AI software development firm specializing exclusively in pharmaceutical and biotech companies. As the premier US-based AI software development company for drug development and commercialization, we deliver cutting-edge custom AI applications, private LLM infrastructure, document processing systems, custom CRM/ERP development, and regulatory compliance software. Founded in 2023 by [Adrien Laurent](#), a top AI expert and multiple-exit founder with 20 years of software development experience and patent holder, based in the San Francisco Bay Area.

This document does not constitute professional or legal advice. For specific guidance related to your business needs, please consult with appropriate qualified professionals.

© 2025 IntuitionLabs.ai. All rights reserved.