

## Understanding CDISC Basics

Jane Ma



**About Author:** Jane Ma holds MS of Health Informatics from University of Minnesota. She has 3 years software industry and 4 years pharmaceutical industry experience. She has been working as a clinical database analyst.

### Abstract

Data standards can make data and its associated program more portable. The CDISC (Clinical Data Interchange Standards Consortium) is a standard for collection, analysis and submission of clinical data to regulatory authority in support of marketing applications. This paper presents an overview of the CDISC basics and the associated contexts, including ODM, SDTM, LAB and AdAM. Included will be the evolution of the standard, the current status, and strategic reasons to considering implementation of the CDISC.

### CDISC Overview

It is widely recognized that standards improve process efficiency, regardless of the industry. Currently, clinical investigators, clinical study personnel, data managers and FDA reviewers must cope with a plethora of data formats and conventions. Some clinical investigators report the presence of many different computer systems for data entry at their sites for various trials, each of which uses different data conventions. Obviously, lack of standardization is not only inefficient and unnecessary expensive in the clinical trial data collection, analysis and submission; it multiplies the potential for error. Standardization of the look and feel of the data collection could reduce these inefficiencies and also help accelerate progress toward electronic data capture and submission. Therefore, the establishment of CDISC will be helpful to develop and support global, platform-independent data standards that enable information system interoperability to improve medical research and related areas of healthcare. <sup>1</sup>

CDISC plays a major role in today's international clinical research data standards. It has been committed to the development of industry standards to support the processing of clinical trials data over the past 8 years. Figure 1 shows the standardization of global clinical research model which was set up by ICH guideline.

There are actually four CDISC models, each intended for different purposes and addressing different categories of data <sup>1</sup>. These models have recently been acknowledged by the FDA. Also, many companies are reengineering their internal processes to adopt them.

### Components of CDISC

- Operational Data Model (ODM)

The Operational Data Model (ODM) is designed first to enable archiving and reconstitution of clinical trial data and second to help transfer clinical trial data between sponsors and CROs. It is also possible to use ODM to exchange data between applicants and authority. This provides a format for representing the study metadata, study data and administrative data associated with a clinical trial. It represents only the data that would

be transferred among different software systems during a trial, or archived after a trial. It need not represent any information internal to a single system, for example, information about how the data would be stored in a particular database.

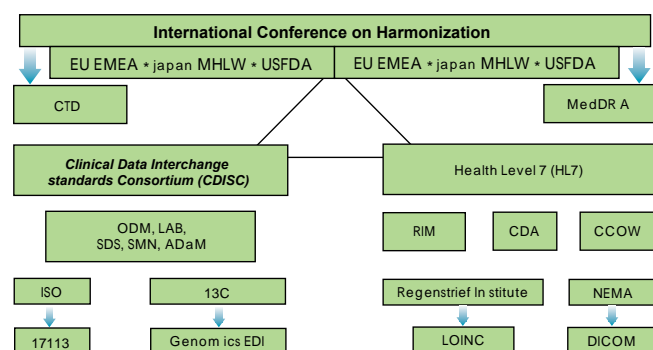


Figure 1. The World of Standards for Clinical Research 2

ODM is divided into data describing the study, administration data, reference data and clinical data. The following table shows the nested levels of clinical data (Tab 2).

Table 2. Clinical Data Level of ODM

Subject	A patient participation in the study.
Study Event	A patient visit or some other data-collection event (e.g., AE report). Each study event belongs to some patient in the study.
Form	Analogous to a page in a CRA. A form generally collects a set of logically and temporally related information. A series of forms is collected as part of a study event.
Item Group	A collection of items. An item group is a closely related set of items that are generally analyzed together. Item groups are aggregated into forms.
Item	An individual clinical data item, such as one systolic blood pressure reading.

- LAB Data Model (LAB)

This model contains data fields, code lists, and date and time format conventions. It is intended to ease the work required to transfer data from central labs or CROs to sponsors.

- Submission Data Models (SDS = Submission Domain Standards)

Study Data Model (SDM) represents the underlying conceptual model behind the SDS standards, and SDTM Implementation Guide (SDTM-IG) includes the detailed domain descriptions, assumptions, and examples. Support SDTM is likely to be the highest priority for most sponsors because of interest by regulatory authorities. A SDM standard recently published in 2004 covers 80% of the data used in CRFs. But it does not include specifications for efficacy data, which are so variable in nature that they preclude standardization (Tab 3).

Table 3 SDM Coverage

Tabulation Datasets	Each record includes a single observation for a subject
Analysis Datasets	Records arranged to support a specific analysis not covered by other data presentations (i.e., not covered in tabulations, listings, or profiles).
Data Listings Datasets	Each record includes a series of observations collected for a subject during a study or for each visit during the study, organized by domain.
Metadata Files	Needed to enable FDA tools to create datasets, profiles, and summary tables and graphs.
Subject Profiles	PDF files that include all study data collected for an individual subject, organized by time.
Summary Tables and Graphics	

Traditionally, clinical summary data are submitted in several formats, such as listings and tabulations. There is much replication of data involved. The FDA or other regulatory authorities, in particular, is anxious to switch to having the summary data submitted in a normalized fashion that their reviewers can apply standard tools for the reviewing.

- Analysis Dataset Model (ADaM)

ADaM builds a library of implementation models that represent Analysis Datasets that would be created to support specific statistical methodology used within clinical trials. The difference of ADaM from SDTM is relying on the distinction

between raw data and derived data. SDTM covers delivering the raw clinical data to regulatory authorities for browsing. ADaM is about delivering results from statistical analysis to the authorities. In the other words, ADaM is to not deliver analysis datasets per se to the authorities, but to deliver descriptions of the standard analysis methods used and allow the authorities to re-create the analysis based on data received via SDTM.

Figure 2 shows how the CDISC standards relate in clinical trials.

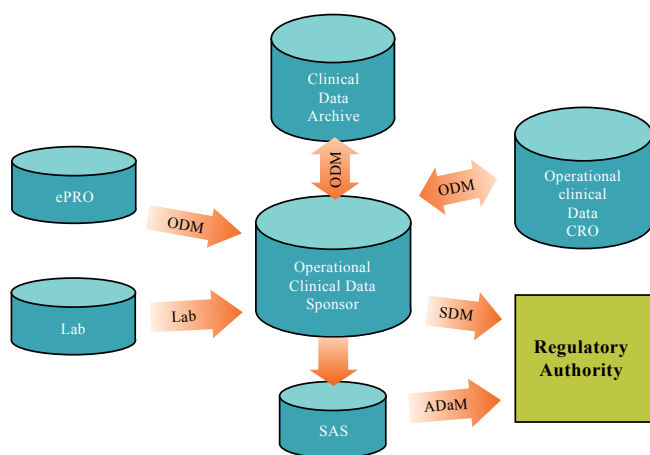


Figure 2. Relationship of CDISC components

The most important of other new standards are currently being published for review or are being developed are:

- The CRT-DDS (better known as define.xml) is a standard for providing data descriptions for Case Report Tabulations in an XML format for submission to a regulatory authority such as the U.S. Food and Drug Administration (FDA). The promise of this standard is that submissions done using the CRT-DDS can be processed considerable faster by the regulatory authorities.
- The Protocol standard is in early development. It will however become an important standard as it will enable to describe a full clinical research protocol in machine-readable format. This will allow for automation of a number of now cumbersome tasks, like automated database setup, computer-based workflows, creation of dynamic eCRFs, etc.

Table 1 summarizes the type of data each standard is intended to help transmit during the clinical research. While ODM is designed to help transfer CRF and audit trail data, it can also

be used to exchange data between IVRS and e-diary products and EDC solutions.

DATA Coverage by CDISC Standard				
Data Type	ODM	LAB	SDTM	ADaM
CRF	x		x	x+
Audit	x			
ePRO	x-			
LAB		x		

From the table, we can see that SDTM does not cover audit trail data, which is designed to transmit patient data to reviewers at the FDA for ad hoc analysis (e.g., identifying unexpected trends). ADaM is shown as containing more than the data included in CRFs, which is designed to transmit derived variables and descriptions of the programmed SAS analysis.

### Evolution of the CDISC

Since its inception in the US in 1997, CDISC gained international traction and popularity. This is reflected in the mission statement: “Clinical Data Interchange Standards Consortium (CDISC) is an open, multidisciplinary, non-profit organization that has established worldwide industry standards to support the electronic acquisition, exchange, submission and archiving of clinical trials data and metadata for medical and biopharmaceutical product development.”

- 1997 – Started as ‘grass roots’ group with 25 attendees at first meeting<sup>3</sup>
- 1998 – Invited to form DIA SIAC<sup>3</sup>
- 2000 – Independent, board-governed, non-profit organization. <sup>3</sup>
- 2001 – Joined as HL-7 associate<sup>3</sup>
- 2002 – First introduction in Japan
- 2002 – EU representative added to CDISC board
- 2004 – First Annual CDISC European **Interchange**

### Current Impacts of CDISC

CDISC standards have a large impact on the clinical development process. Data standards are a critical component in the quest to improve global public health. CDISC provides more efficient and effective use of medical information by all members of the healthcare and life sciences ecosystem.

Before CDSIC era, in the absence of industry standards, each biopharmaceutical company and CRO has developed their own. This has led to a plethora of standards within

the industry. In the absence of standards, each sponsor and project pair can require custom interface (Fig 3).

In CDISC era, it is possible to collect, process, and analyze patient and health related information quickly, more cost- effectively and with greater accuracy under a standard. Each sponsor or project only needs to support one import format for clinical data. Conversely, each vendor can only develop one export format that can be used by all sponsors and projects. This example, finally, illustrates the value of most data exchange standards – reduce the effort to transfer data by agreeing on a common transfer format (Fig 4).

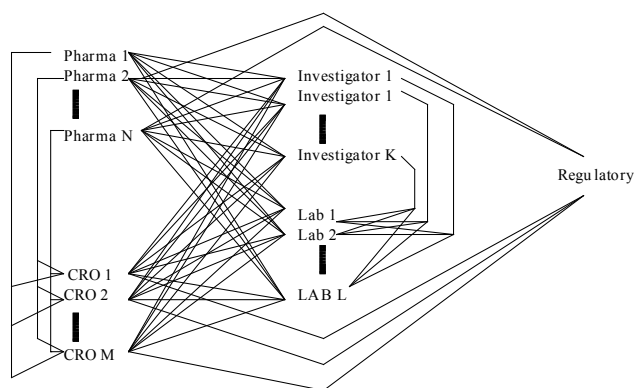


Diagram 3: Before CDISC – In the Absence of Standard 2

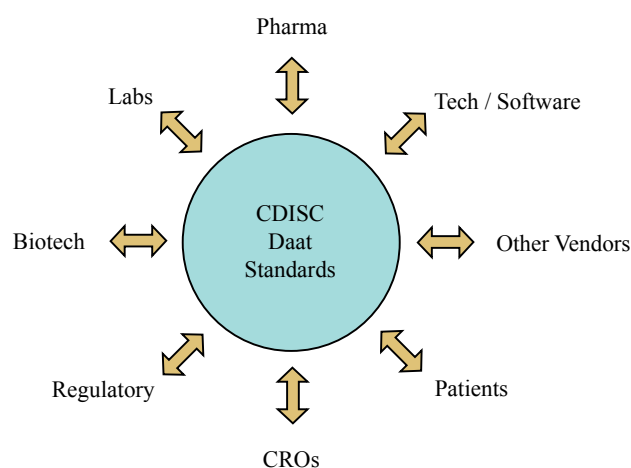


Diagram 4: After CDISC – One Standard for All 4

Thus, CDISC improves the cost and quality of healthcare delivery for patients and consumers. In the other words, establishment of data standards and application of the standards across all studies and projects can be resource intensive. One of the key benefits to

implement the standards is that the programs associated with this data become more portable. They can be moved from one study to the next with minor modifications. Not only are the programs more portable, the programmer and statistician working on one study can understand a new study with the same structure relatively quickly compared to learning a new set of programs, macro and data structures. The productivity gain is sometimes difficult to measure but, in the long run, it will outweigh the efforts invested in standardizing.

### Conclusion

Based on the facts above, it can be concluded that CDISC implementation can cause:

- Harmonization with HL-7 in the standard for data submission, including CRT datasets, analysis datasets and programs, and metadata.
- Exchange of all clinical trial data between any two parties will be achieved by the application of the appropriate CDISC data models and standards. 2
- Generation of a document that would benefit industry and FDA by providing recommendations for the use of CDISC standards with associated processes that can promote the enhanced use of electronic source data interchange within the context of the existing regulations for regulated clinical research.

Since becoming the standard for collection, analysis and submission of clinical and preclinical trial data to the authorities in marketing application, the CDISC will become a tool to be used by the industries for their upstream processing to support Clinical Study Reports and Integrated Summaries.

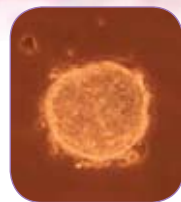
### References

1. <http://www.cdisc.org>
2. Wayne R. Kubick; <http://www.cdisc.org/pdf/CDIS-CBio0103x.pdf>
3. Wayne R. Kubick; [www.lincolntechnologies.com/Presentations/BarnettERSX.pdf](http://www.lincolntechnologies.com/Presentations/BarnettERSX.pdf)
4. Robert Weber; [www.hl7.de/veranstaltungen/kongress/xmlworkshop062002/Weber.PDF](http://www.hl7.de/veranstaltungen/kongress/xmlworkshop062002/Weber.PDF)

# MILLIPORE

## Our Line-up of Cell Culture Solutions Just Keeps Growing

Now that Upstate®, Chemicon and Linco are part of Millipore Corporation, you have one source for stem cells, sterile filtration devices, water systems and many of the other products you need for cell culture.



- The broadest selection of cell culture inserts and plates including MilliCell® single-well inserts (hanging and standing), 24-well and 96-well insert plates
- The widest selection of vacuum and pressure-driven devices for sterile filtration, including Stericup™, Steritop™, Steriflip™ and Millex™ Filter Units
- Mouse embryonic stem cells from a variety of genetic backgrounds
- The Milli-Q® Advantage Ultrapure Water System with customizable delivery options
- Media components and formulations including ESGRO® Complete Medium
- More than 10,000 purified primary and secondary antibodies

**Greater Selection. Greater Convenience.**

Visit [www.millipore.com](http://www.millipore.com) for more information on our cell culture solutions.