

Developing a CDISC Intelligence Quotient

Some CDISC members envision a time in the future when at least three of the CDISC standards will be combined into one.

**Until then, the question remains:
what does CDISC-compliant mean to you?**

KATHRYN ROY

IN AN EXCLUSIVE TO PHARMAVOICE, KATHRYN ROY, VP OF MARKETING AT PHASE FORWARD INC., DISCUSSES WHAT CDISC-COMPLIANT REALLY MEANS AND OUTLINES THE FACTORS THAT SHOULD BE CONSIDERED WHEN DEVELOPING A CDISC STRATEGY.



By breaking down the components of CDISC, people can better understand what it really means to **comply with CDISC standards** today and the implications of future CDISC implementation plans.

“People tend to refer to CDISC as though it is monolithic: ‘We are CDISC compliant,’” says Kathryn Roy, VP of marketing at Phase Forward Inc. “This is like saying, ‘I’m good at gambling’ when all someone knows is blackjack.”

There are four distinct CDISC data models — LAB, for exchanging lab data; ODM, the Operational Data Model; SDTM, the Study Data Tabulation Model; and ADaM, the Analysis Dataset Model — that ultimately will support the end-to-end data flow of clinical trials, from the sources into an operational database, through analysis to regulatory submission. The sources of data that are relevant to CDISC vary among patient records (i.e., case report form data), clinical laboratory data, data from CROs, shared data between companies, and other sources.

“By breaking down the components of CDISC, people can better understand what it really means to comply with the standards today and the implications of future CDISC implementation plans,” Ms. Roy says.

Data Coverage

“ODM is intended to help transfer CRF and audit-trail data, it also can be used to exchange data between IVRS and e-diary products and EDC solutions,” she says. “SDTM is intended to transmit patient data to reviewers at the FDA for ad hoc analysis, for example identifying unexpected trends; but it does not cover audit-trail data. ADaM is shown as containing more than the data included in CRFs and can be used to transmit derived variables and descriptions of the programmed SAS analysis.”

LAB — Exchanging Lab Data

“The objective for the LAB standard is to ease the work required to transfer data from central labs or CROs to sponsors,” Ms. Roy says. “In

What Is CDISC?

the absence of a standard, each laboratory and sponsor pair can require custom interfaces, to the extent of 18 different interactions. With a standard such as LAB, each laboratory only needs to develop one export format that can be used by all sponsors. Conversely, each sponsor only needs to support one import format for lab data. This example illustrates the value of most data exchange standards: reduce the effort to transfer data by agreeing on a common transfer format.”

The LAB standard has three file format options: delimited ASCII, SAS, and XML.

“The LAB standard does not yet facilitate partial transfers,” Ms. Roy says. “It has no facility to recognize when it is receiving the same information twice. This places the burden of oversight on the sender and recipient, rather than on the standard. One way of thinking about LAB is that it has placeholders for all common statistics reported by common lab procedures.”

ODM — Operational Data Model

ODM refers to the standard data interchange models that are being developed to support the data acquisition, interchange, and archiving of operational data. 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 applications. Unlike LAB and SDTM, ODM was XML-based from the start.

“As designed to date, ODM is a reasonable mechanism for archiving,” Ms. Roy says. “It allows the storage of clinical data, audit trails, and sig-

CDISC IS AN OPEN, MULTIDISCIPLINARY, NONPROFIT ORGANIZATION COMMITTED TO THE DEVELOPMENT OF INDUSTRY STANDARDS TO SUPPORT THE ELECTRONIC ACQUISITION, EXCHANGE, SUBMISSION, AND ARCHIVING OF CLINICAL-TRIALS DATA AND METADATA FOR MEDICAL AND BIOPHARMACEUTICAL PRODUCT DEVELOPMENT. THE MISSION OF CDISC IS TO LEAD THE DEVELOPMENT OF GLOBAL, VENDOR-NEUTRAL, PLATFORM INDEPENDENT STANDARDS TO IMPROVE DATA QUALITY AND ACCELERATE PRODUCT DEVELOPMENT.

natures. But ODM doesn't cover appearance, so retrieving an archived trial's data will not enable replication of the appearance of the original electronic or paper CRFs. Because the few trials that are reconstituted are usually at the request of a regulatory authority investigating safety issues, this should suffice.

“As sponsors start to set up their trials to conform to CDISC, they are advised to set them up to conform to ODM and SDTM at the same time,” she continues. “The two standards are complementary, and setting up trials in accordance with both standards prepares the sponsor to export ODM and SDTM easily.”

It's important to keep in mind, however, that if a company has an ODM compliant database, it does not translate automatically into SDTM export compliance. Additional manual steps will be required.

According to Ms. Roy, ODM is not specific enough to enable inter-application data transfer on its own, for example between e-diaries or IVRS and an EDC system.

“While LAB is a good standard, making it possible to write one interface that would work with multiple partners, ODM is not sufficiently specific to accomplish this goal,” she says. “It solves only part of the problem.”

SDTM — Study Data Tabulation Model

SDTM refers to the standard metadata models being developed to support the data flow from the operational database to regulatory submission. According to Ms. Roy, supporting SDTM is likely to be the highest priority for most sponsors because of interest by the FDA and other regulatory authorities. The latest revision of the SDTM standard 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.

“To understand SDTM relative to ADaM, it is useful to distinguish between raw data and derived data,” she says. “SDTM covers delivering the raw clinical data to regulatory authorities for browsing. ADaM is about delivering results from statistical analysis to the FDA.”

Ms. Roy says several sponsors participated in a pilot submission to the FDA in 2003, mostly using data previously submitted.

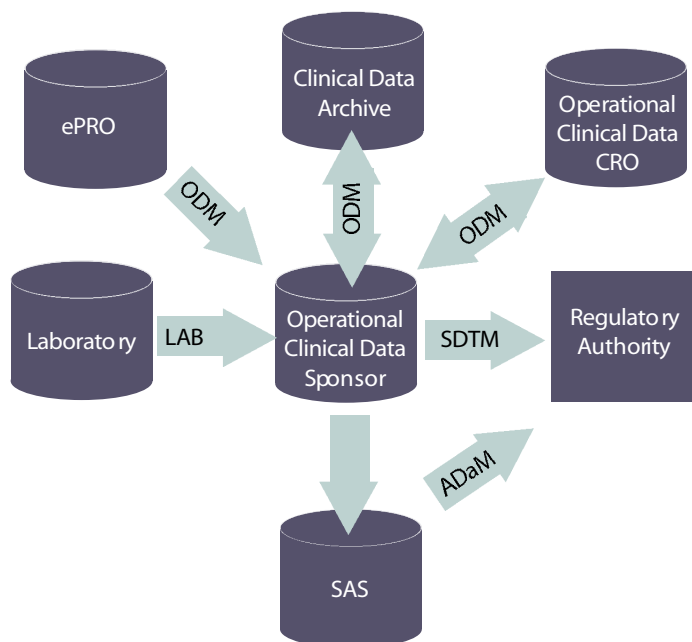
“Some sponsors extracted from OC and some from SAS,” she says.

RELATIONSHIP OF CDISC STANDARDS

Following are the four distinct CDISC standards:

LAB — for exchanging lab data
ODM — Operational Data Model

SDTM — Study Data Tabulation Model
ADaM — Analysis Dataset Model



DATA COVERAGE BY STANDARD

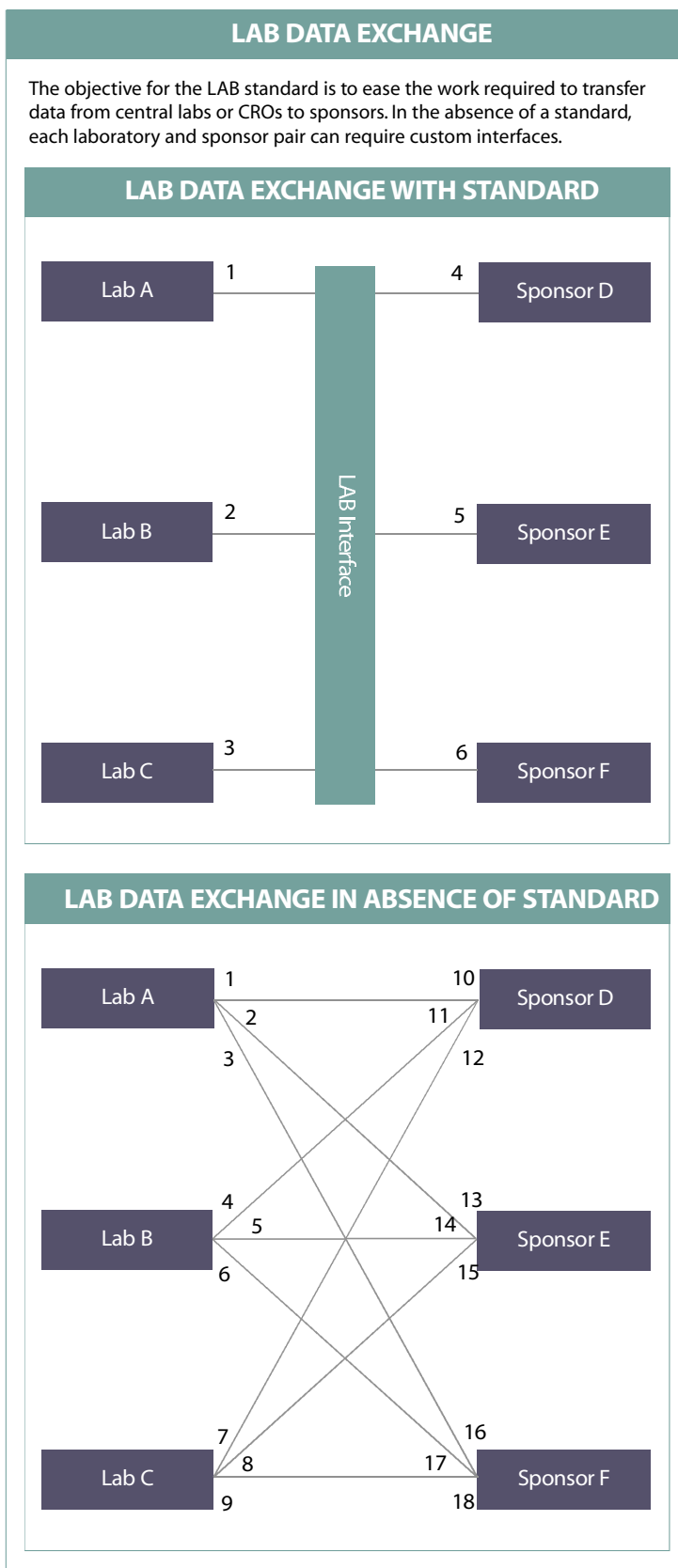
Data Type	ODM	LAB	SDTM	ADaM
CRF	■		■	■ +
Audit	■			
ePRO	■ -			
LAB		■		

Note: The table summarizes the type of data each standard is intended to help transmit.

DATA FORMAT BY STANDARD

Data Format	ODM	LAB	SDTM	ADaM
CSV/ASCII		■		
STF		■	■	■
XML	■	■		

Note: STF (SAS Transport Format) and CSV (Comma Separated Value).



“SAS can be used to prepare CDMS submissions because SAS must import all the raw data to support sophisticated analysis. But the raw data can and should be transferred directly from CDMS systems.”

Traditionally, sponsors submit summary data in several formats, such as listings and tabulations, which is a replication of a lot of data.

“The FDA, in particular, is anxious to switch to having sponsors submit the data in a normalized fashion — no repetition of data — that

SDTM 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	(currently) PDF files that include all study data collected for an individual subject, organized by time.
SUMMARY TABLES AND GRAPHS	

regulatory reviewers can apply tools to for browsing datasets, generating data tabulations and listings, and creating summaries and reports.”

Initially, Ms. Roy says SDTM, or Submission Data Standard (SDS), is being delivered via STF (SAS Transport Format); the goal is that all standards for CDISC targets are based on XML.

According to Ms. Roy, the benefits of XML are: it is machine readable (can be sent to and interpreted by computers) versus paper or fax; it is human readable, making errors easier to investigate; it is platform independent, so no special tools are required to transfer between hardware or OS platforms; it is hierarchical, so one can transfer information that normally resides in complicated tables; and it is a de facto standard that most application providers are now familiar with.

At its simplest level XML looks like this: <tag1>value</tag1>. The <tag1> refers to the name of the data item; value refers to the value of the data item; and </tag1> recognizes the end of the data item value. Therefore, <patientID>1782</patientID> translates into: the patient’s ID is 1782.

“The first priority is to enable tools to view subject profiles and later to add data listings and tabulations,” she says. “Analysis datasets are covered separately under ADaM.”

ADaM — Analysis Dataset Model

Analysis datasets, as described in SDTM previously, are typically SAS output. ADaM appears to be at the early stages of development. The objective is to not deliver analysis datasets per se to the FDA, but to deliver descriptions of the standard analysis methods used and allow the FDA to recreate the analysis based on data received via SDTM. ♦

PharmaVoice welcomes comments about this article.

E-mail us at feedback@pharmavoice.com.