

Based on research by:
Kirk Mousley, PhD and
EDC Management

Written by:
Anne Jacobson, MPH

*In the next issue of
EDC Today:*

Vendors, Products and
Suppliers: A Consumer's
Guide to the EDC
Marketplace – Part I

About EDC Management:

EDC Management was founded to assist biopharma companies plan, prepare for and implement Electronic Data Capture (EDC) strategies according to their data management goals and objectives. We do not sell or endorse any specific EDC software application or vendor.

EDC Management

P.O. Box 384
Conshohocken, PA 19428
484-530-0300 (voice)
610-567-0357 (fax)
info@edcmanagement.com
www.edcmanagement.com

Data Transfer Standards: How to Leverage Your Investment in EDC

Welcome to our eighth issue! EDC Today is an independent publication about current information and issues in Electronic Data Capture (EDC) strategies and technologies for the Biotechnology and Pharmaceutical (biopharma) industry. Each month we examine topic areas related to EDC theory, technology, practice, or implementation.

During the clinical trials process, data are gathered from a variety of sources: clinical laboratories, contract research organizations, and EDC vendors. Within the biopharma industry, data are flowing from multiple service providers to multiple sponsors, each of whom probably has its own data requirements. To streamline the flow of data traffic and create a common data “language” among sponsors and vendors, data standards are often used. In this issue, we explore the use of industry-wide data standards within clinical trials

As clinical trials become increasingly complex, sponsor companies are outsourcing part of the process of conducting a clinical trial to a contract research organization (CRO), an EDC vendor, a central clinical laboratory, or even another biopharma company. When interacting with outside service providers, sponsors must detail the type of data structures they would like to use so that transferred data will fit without complex transformation into their database.

Sponsors and service providers can choose from a daunting array of data “languages.” While data models enable coded information to be translated, there are significant costs associated with converting data so that it can be exchanged between the sponsor and the service provider. Steps involved in the data transfer process may include any or all of the following:

- Set-up, such as defining the clinical trial protocol within the different systems in terms of data structures and edit checks
- Programming or converting the data structures from one system (exporting) to another (importing)

- Actual transfer of the data between physical locations
- Verification of the transfer and rechecking of the data in the receiving system

The data transfer process represents a significant cost to the biopharma industry. In November 2000, CenterWatch estimated that the annual cost to the biopharma industry for CRO and EDC data transfers alone is approximately \$80 million; include clinical laboratory data transfers and the estimate boosts to \$156 million annually.¹ Further still, these estimates do not include the costs incurred by merged companies sharing data or the cost of preparing electronic submissions.¹

The Clinical Data Interchange Standards Consortium (CDISC) calculated even higher cost estimates in 2000. According to CDISC, approximately 3,500 trials requiring CRO and/or EDC data transfer are conducted each year.² The estimated cost for data transfer on a single trial is \$35,000. Overall, the annual industry cost of transferring CRO and/or EDC data is \$122.5 million. This figure does not include clinical laboratory data transfers, equipment costs, planning, personnel training, or error resolution.²

Contributing to the high cost of data transfer are the number of persons involved and the time they must devote to the process. Data transfers may involve a variety of sponsor and service provider personnel, including software programmers, clinical data managers and analysts, quality assurance auditors, and clinical laboratory personnel. Additional costs arise from the many activities that are necessary to facilitate data transfer:

- Clinical laboratories must support hundreds of different interfaces
- CROs must develop completely new databases to meet the unique specifications of each trial sponsor

(continued on page 2)



- EDC providers must customize the exchange format for each client
- Data managers must rely on programmers and integration steps before they can perform their functions with the database³

Data transfers can occur either once at the end of the trial or many times during the course of the trial. Though multiple consecutive transfers or transfers between vendors who are doing repeat business for a client do not cost as much as the initial transfer, they still incur additional resource requirements and time.

Clinical development programs typically last seven to ten years, during which sponsor companies must manage new technology and new systems. When technologies are upgraded or replaced, companies must address data migration and archival retention issues, which often require further data transformations. In addition, data structures may often need to be converted or transformed to conform to FDA requirements for submission of regulatory data.

The Need for Data Standards

Biopharma companies have been developing internal standards for some time, and with varying degrees of success. Company-specific standard operating procedures (SOPs) are one example. However, when companies must reach beyond their own walls and begin to communicate with other organizations, standards must be more widespread to bridge different companies and different functions.

Although implementing standard data models would not eliminate transfer costs completely, it has been estimated that using a standard transfer model – for example, as used by an EDC vendor that functions as an application service provider – could reduce costs by approximately 35%. As reported by one CRO, the repeat development of databases for a single client reduced hours for the developer by approximately 60% for the third database compared to the first.²

In addition to the obvious benefit of cost reduction, data transfer models offer a variety of benefits to sponsors, investigator sites, service providers, and regulatory reviewers.

Benefits of implementing and employing data transfer standards include:

- Streamlined conversion of operational data to submission data
- Minimized variation in structures and terminology among common safety datasets
- Elimination of the collection of unnecessary data points
- Reduction of errors due to misunderstandings or false assumptions
- Reduction of training costs and reorientation time among sites and regulatory reviewers for each new study or application
- Opportunities for the FDA to develop standard review tools leveraging standards

These benefits promise to improve the data transfer process in many ways. Sponsors can receive data in a standard format in less time and with the assurance of higher quality, benefiting from the savings realized by the streamlined processes of their vendors and partners. In addition, because submission data will appear in a more consistent and familiar format, regulatory reviewers such as physicians and statisticians can review it more quickly and easily.

Standardization also brings the potential for major improvements in data consistency and reusability. For example, rather than establishing, formulating, designing, coding, documenting, and training data management systems for each individual trial, biopharma can perform these tasks just once and use the data model in all subsequent trials. These improvements may in turn lead to further efficiencies in the performance of operational data transfers for the duration of the clinical research cycle.

By facilitating the uniform, consistent presentation of data, data standards may simplify the quality assurance process and may help ensure compliance. For the same reasons, standards can also simplify data reporting and increase the utility of the data in reporting. Further, because implementing data standards reduces the number of possible models in use, standards can significantly ease the burden of the help desk and information technology staff. Finally, organizations may find it beneficial that data collected under data standard models are archived in a “known,” uniform format.

Standards: Who Creates Them?

The development of industry-wide standards requires a concerted and multidisciplinary effort. One non-profit organization, the Clinical Data Interchange Standards Consortium (CDISC), is leading this effort. According to their mission statement, CDISC is “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.”⁴ They champion “the development of global, vendor-neutral, platform independent standards to improve data quality and accelerate product development” within the biopharma industry.⁴

CDISC details its principles on its Web site (www.CDISC.org). According to its mission statement, CDISC strives to:

- “Lead the development of standard data models that improve process efficiency while supporting the scientific nature of clinical research.
- Recognize the ultimate goal of creating regulatory submissions that allow for flexibility in scientific content and are easily interpreted, understood, and navigated by regulatory reviewers.

(continued on page 3)

- Acknowledge that the data content, structure, and quality of the standard data models are of paramount importance in this endeavor, independent of implementation strategy and platform.
- Maintain a global, multidisciplinary, cross-functional composition for CDISC and its working groups.
- Work with other professional groups to encourage maximum sharing of information and minimum duplication of efforts.
- Provide educational programs on CDISC standards, models, values, and benefits.
- Accomplish CDISC goals and mission without promoting any individual vendor or organization.”⁴

Though CDISC clearly plays a lead role in the development of standards, it is not alone in this pursuit. The International Conference for Harmonization (ICH) is also striving to develop industry-wide standards that harmonize the technical aspects of pharmaceutical development among the European Union, Japan, and the United States. The ICH is sponsored by the U.S. Food and Drug Administration, the Centers for Drug Evaluation and Research and Biologics Evaluation and Research, the Japanese Ministry of Health and Welfare, the European Federation of Pharmaceutical Industries Associations, the Japanese Pharmaceutical Manufacturers Association, and the Pharmaceutical Research and Manufacturers of America.

To avoid duplicating efforts, CDISC collaborates with other organizations such as the ICH. CDISC also collaborates with biopharma companies by communicate progress and requesting input on the development of data models and other CDISC recommendations.

Standardized Data Models

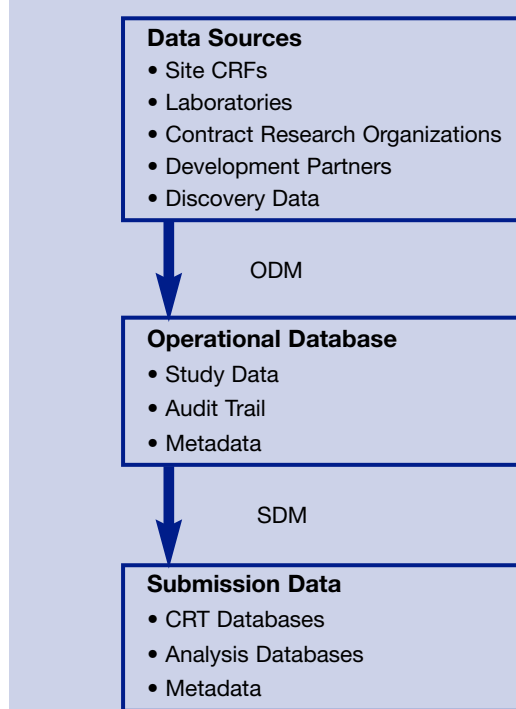
CDISC data models are designed to support the flow of data through the clinical trials process: from the data source to an operational database, through analysis, and on to regulatory submission. Data sources can include patient records, case report forms, clinical laboratory data, data from CROs, shared data by merging or partner companies, and other sources.

CDISC has developed functional models to support and facilitate several types of data exchanges within the clinical trials process. Below are common functional models:

- **Operational Data Model:** Converts data from multiple acquisition sources, including CROs and EDC tools, into an operational database
- **Laboratory Data Model:** Converts clinical laboratory data into an operational database
- **Submissions Data Model:** Converts operational data from a sponsor or operational database to a regulatory agency
- **Analysis Dataset Model:** Creates datasets to support statistical reviews of electronic submissions⁴

Table 1 depicts how two data models, the Operational Data Model (ODM) and the Submission Data Model (SDM), fit in the flow of data within a clinical trial. While the ODM facilitates data flow between source data and the operational database, the SDM translates data from the operational database into information that is primed for regulatory submission.

Table 1. CDISC Data Models and Data Flow⁴



In addition to these formal, comprehensive data models, many individual tools used in the data transfer process can be standardized, including:

- Data item (questions; attributes to standardize include, but are not limited to, data type, size, precision, and units)
- Modules (logical group of items/questions)
- Form (data entry screen of module/items)
- Derivations/calculations
- Edit Check specifications and implementations
- Clarification item wording on data clarification forms (DCFs)
- Case report form (CRF) page layouts
- DCF page layouts
- Dictionaries and codelists (e.g., MedDRA)
- Training materials
- Standard operating procedure (SOP) mandated documentation (i.e., make documentation more “fill in the blanks”)
- Computer systems (e.g., hardware, software, and operating systems) and configurations

(continued on page 4)



The degree to which individual organizations want to adopt standard data models is dependent on a range of factors, including the size of the organization and the volume of data that must be managed. For example, some organizations may employ several CDISC data models and develop internal standards for CRFs, training materials, and computer systems. Other organizations may choose a different assortment of standards depending on their specific needs.

Developing Standards

CDISC has a strict procedure for developing data standards. Though each type of data model has its own unique requirements, the CDISC process for developing standards involves some steps common to all data models. These include:

- Standard definition and team formation
- Standard development and review
- Education and support
- Standard update and maintenance
- Accreditation⁴

Anyone in the biopharma industry can identify and propose a specific standard to the CDISC board. When CDISC approves formal proposals, meaning that they meet with CDISC strategy and allocation of resources, a multidisciplinary data standard team is formed. The team of seven to twelve members develops a plan detailing the goals, deliverables, and communication strategies of the model. For example, the Submissions Data Standards team maintains the following objectives:

- “Provide regulatory submission reviewers with clear descriptions of the usage, structure, contents and attributes of all submitted datasets and variables
- Allow reviewers to replicate most analyses, tables, graphs and listings with minimal or no transformations
- Enable reviewers to easily view and subset the data used to generate any analysis, table, graph or listing without complex programming”⁴

Similarly, the Analysis Data Model team maintains these goals:

- “Build on metadata CRT models developed for safety domains, adding attributes and examples specific to statistical analyses
- Use sample statistical results as a guide for developing data set models
- Initially focus on primary and secondary efficacy variables”⁴

Draft models are first synchronized with other CDISC models in progress and then reviewed by an external review committee comprised of experts in the biopharma industry.

Testing is a critical part of the data standard model development process and is typically performed by both the development team and the external review committee. Before models reach the testing phase, they are posted in team-specific areas of the CDISC Web site. Once the team has addressed comments resulting from external tests, the model is posted on the public side of the CDISC site and is open for public comment. Once public comments have been addressed, the model is released and available for implementation by users.

Following the initial release of a data model, Version 1.0, most models undergo an annual update through a process similar to their original development. Though the public can use them freely, all models are proprietary to CDISC.

An important final step in model development is user education. The CDISC Education Committee (EDU) collaborates with both the model development teams and CDISC Public Relations/Communications task force to educate users, make sure that the models are used properly, and assist users with implementing them. EDU also provides tutorials at professional meetings for biopharma personnel who use CDISC standards. CDISC is currently developing an accreditation process for organizations and individual users who use CDISC standards.⁴

Implementing Data Standards

Because implementing data standards may require a significant investment in time and financial resources, organizations wishing to adopt data standards may meet with significant internal resistance. However, several approaches can be used to ease the adoption process and facilitate organization-wide support.

First, sponsors may attempt to introduce or renew efforts toward standardizing when implementing new systems and technologies, such as EDC. New data standards can therefore be viewed as an integral part of implementing the new system or technology.

Biopharma companies may also attempt to familiarize those who are not in daily contact with corporate finance with the importance of the financial implications of data transfer. Having employees feel that they contribute to the financial health of the organization may enable them to embrace programs designed to increase efficiency and boost financial well-being.

(continued on page 5)



Finally, companies may want to solicit input from all those who would be affected by the introduction of data standards. Listening to and addressing the concerns of employees may facilitate employee buy-in. However, organizations may want to contain the number of persons whose approval is necessary before new standards can be implemented. The greater number of persons who must approve something, the more difficult it may be to reach consensus.

Conclusion

Data transfer within the biopharma industry is a costly and time-consuming endeavor. Implementing data standards is an effective way of streamlining the process, improving efficiency, and reducing cost.

Currently, many data standards are developed by the Clinical Data Interchange Standards Consortium (CDISC), a group of industry experts whose primary goals are to improve process efficiency and facilitate regulatory submissions while supporting the scientific nature of clinical research. Through a formal development process, CDISC authors and tests data standard models with industry and public input.

Though biopharma companies wishing to implement data standards may meet some internal resistance, several steps can be taken to ease the process. Adopting data standards may help companies achieve a fundamental business goal: bringing products to market faster.

Who's behind the research?

Our lead researcher, Kirk Mousley, PhD received BS and MS degrees in Electrical Engineering from MIT and a PhD in Computer Science from Lehigh University. He has been the President of Mousley Consulting, Inc. since its founding in 1993 and has directed the company's efforts in the areas of clinical database design, data editing/cleaning, document management, and submissions.

Karl Mousley received his BS in Mechanical Engineering from Rose-Hulman Institute of Technology and a MS in Computer Science from Villanova University. He has been a senior member of the technical staff at Mousley Consulting, Inc. since 1993. Among his significant accomplishments are the investigation, evaluation, and implementation of new computer technologies for clinical data management systems and developing strategic plans for integrating these technologies into current systems. He has extensive experience preparing Standard Operating Procedures (SOPs).

References

- ¹ *Multinational Trials Bolstered by Technology. CenterWatch Newsletter. 2000;7(11).*
- ² *Kush, Rebecca. The Cost of Clinical Data Interchange in Clinical Trials. CDISC White Paper. 2001.*
- ³ *Kush, Rebecca. A Multidisciplinary Approach to Data Standards for Clinical Development -- Progress Update. Applied Clinical Trials. 2002;11(4):35-44.*
- ⁴ *Clinical Data Interchange Standards Consortium (CDISC). <http://www.cdisc.org>*

Available ***EDC In Depth*** Research Reports related to this issue:

8.1 "Beyond the Case Report Form: Standardization in EDC"

Data clarification forms, case report forms, and codelists are some of the data management instruments often standardized by sponsors and EDC vendors. In this report, we go beyond these common tools and explore additional processes that can be standardized in an effort to streamline the data management process.

8.2 "Clinical Data Interchange Standards Consortium (CDISC): A Force to Watch"

CDISC is committed to the development of standards that support the electronic acquisition, exchange, submission and archiving of clinical trials data and metadata in the biopharma industry. In this report, we detail their efforts and examine how biopharma companies might implement their recommended standards.

8.3 "Identifying and Overcoming Obstacles to Standardization"

Despite their many potential benefits, standards may be difficult for some biopharma companies to implement. In this report, we describe common obstacles to standardization and suggest techniques to overcome them.

8.4 "Implementing Standards: Finding Opportunity in Change"

Adopting any new process can be daunting, particularly when the process spans multiple divisions and groups of personnel within a single organization. In the report, we describe ways that implementing data transfer strategies can be considered opportunities to improve organization-wide processes.

See back for order information.



edc today

Issue No. 8



EDC Today and EDC In Depth

EDC Management publishes well-researched, timely information about EDC technologies and processes. Our technical bulletin, *EDC Today*, overviews a group of related in-depth research reports, *EDC In Depth*.

An annual subscription to *EDC Today* is \$175/year and can be purchased via downloadable electronic version or paper version sent via mail. Electronic versions of back issues may be purchased individually for \$25 each.

Each *EDC In Depth* research report comes with an executive summary and may be purchased individually for \$395 or as a group of related reports for \$975. Available in electronic or paper version.

To subscribe to ***EDC Today*** or purchase a specific ***EDC In Depth*** research report:

Order online at www.edcmanagement.com

Email us at info@edcmanagement.com

Call us at **1-484-530-0300**

EDC Management is a consulting and information company offering EDC consulting, including vendor selection, validation and compliance, and site assessment; Clinical and Data Management consulting; hands-on education and training; and EDC publications. Improve process today; position for tomorrow.

©2002 EDC Technologies LLC. All rights reserved. Cannot be published without prior written consent. "EDC Management" and "EDC Technologies" are service marks of EDC Technologies LLC. "EDC Today" is a trademark of EDC Technologies LLC. All other marks and trade names are the property of the respective companies.