edc management℠
education
integration
evaluation

edctoday™

A Technical Bulletin Published by EDC Management Issue No. 11

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

Written by:
Karl Mousley

In the next issue of ***EDC Today:***

Portals: How They Can Make EDC Work Better.

About EDC Management:

EDC Management is the leader in Clinical and Data Management and Electronic Data Capture (EDC) consulting services for the biopharmaceutical industry. EDC Management publishes well-researched and timely information about Electronic Data Capture technologies and processes through EDC Today™ and EDC In Depth. We do not sell or endorse any specific EDC software application or vendor. Improve process today; position for tomorrow.

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

02-07-03-4461

Technology Old and New:
Integrating Legacy Systems with EDC

Welcome to our eleventh 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.

Selecting an EDC product is a critical process for biopharma companies. Biopharmas hoping to implement EDC products have a number of considerations to make – determining how to handle an existing legacy Clinical Data Management System (CDMS) and any auxiliary systems is one of these considerations. In this issue, we present a number of integration considerations and explain why some careful thought concerning these systems is necessary when implementing an EDC product. Those sponsors with no legacy systems (e.g., new startup firms) should still find this paper relevant as it reviews functionalities that Biopharmas with years of clinical data management experience perceive as being important…yet are inexplicably missing from their CDMS.

When deciding to implement an EDC product, some thought is necessary regarding what is to be done with the old CDMS. Will it be retained and see continued use? If so, will it be "tied into the EDC system"? Will its use be ramped down over time as studies currently being managed with it wind down? What should be done with the "old data" if the legacy CDMS is decommissioned? Finally, what will be done with the auxiliary systems, if any, that were implemented to augment the CDMS?

Legacy CDMS have often been customized, either directly or by the development of supporting auxiliary systems, by sponsors and Clinical Research Organizations (CROs) to meet perceived business requirements not met by the core system and/or to provide perceived business advantages. Some of these systems may have been developed over long periods of time and some at considerable expense.

Sponsors have usually invested a substantial amount of time and effort in adapting a legacy CDMS to their business requirements (and probably a substantial amount of time and effort in adapting their business to their legacy CDMS). While legacy CDMS products were often marketed and packaged as "shrink wrapped" software, none of them ever truly were. Time spent molding the system to the Sponsor's needs include the development of many standards and procedures. Metadata describing standard questions/items, module/panels/tables, forms/pages, edit-checks, and even discrepancy correction queries needed to be developed. Instructions/training for entry system developers, data entry personnel, and data managers needed to be formulated and promulgated. Auxiliary systems providing functionalities not supported by the core legacy CDMS may have been implemented as well.

Obviously, a prudent sponsor or CRO will inventory its existing systems, metadata, and standards with an eye to reclaiming or cyber-recycling its investment in the legacy system. A smart sponsor or CRO will also use this "chance" to improve its processes by streamlining, improving, or perhaps even eliminating obsolete or inefficient processes.

(continued on page 2)

When tackling an EDC implementation project, it is important to consider the legacy CDMS and any auxiliary systems. Furthermore, it is important to assess not only how the new EDC system will integrate with the CDMS but also how it will integrate with existing auxiliary systems. Finally, after assessing the systems in place, it is important to incorporate the information just uncovered into the EDC implementation plan. This valuable information will include such things as the identification of functionalities that were not part of the CDMS (but perhaps should be part of the EDC system) as well as identify systems that may need to be re-engineered in order to continue to provide required business functionalities.

In this paper, it is assumed that any sponsor with a sizable investment in a legacy CDMS with auxiliary systems would naturally wish to retain its existing CDMS, perhaps indefinitely, or even more likely, gradually ramp down its use until many of the functions found in the auxiliary systems have been either incorporated in their new EDC product implementation, modified to work with the EDC system, or simply abandoned as no longer needed.

Degrees of EDC-CDMS-Auxiliary Integration

As shown in Figure 1, the new EDC system can fit into the data processing system architecture relative to a legacy CDMS system in one of three ways. It could be closely coupled and data entered at the investigator sites is stored directly in the CDMS database. The new EDC system might be loosely coupled with a legacy CDMS system and have its own database, requiring data transfers to occur on a (usually) pre-determined schedule. Or the new EDC system might "stand alone" and have no ties at all with a legacy CDMS system and data stored in its self-contained database would be transferred directly into a Submission database (e.g., SAS DataSets). (See Table 1)

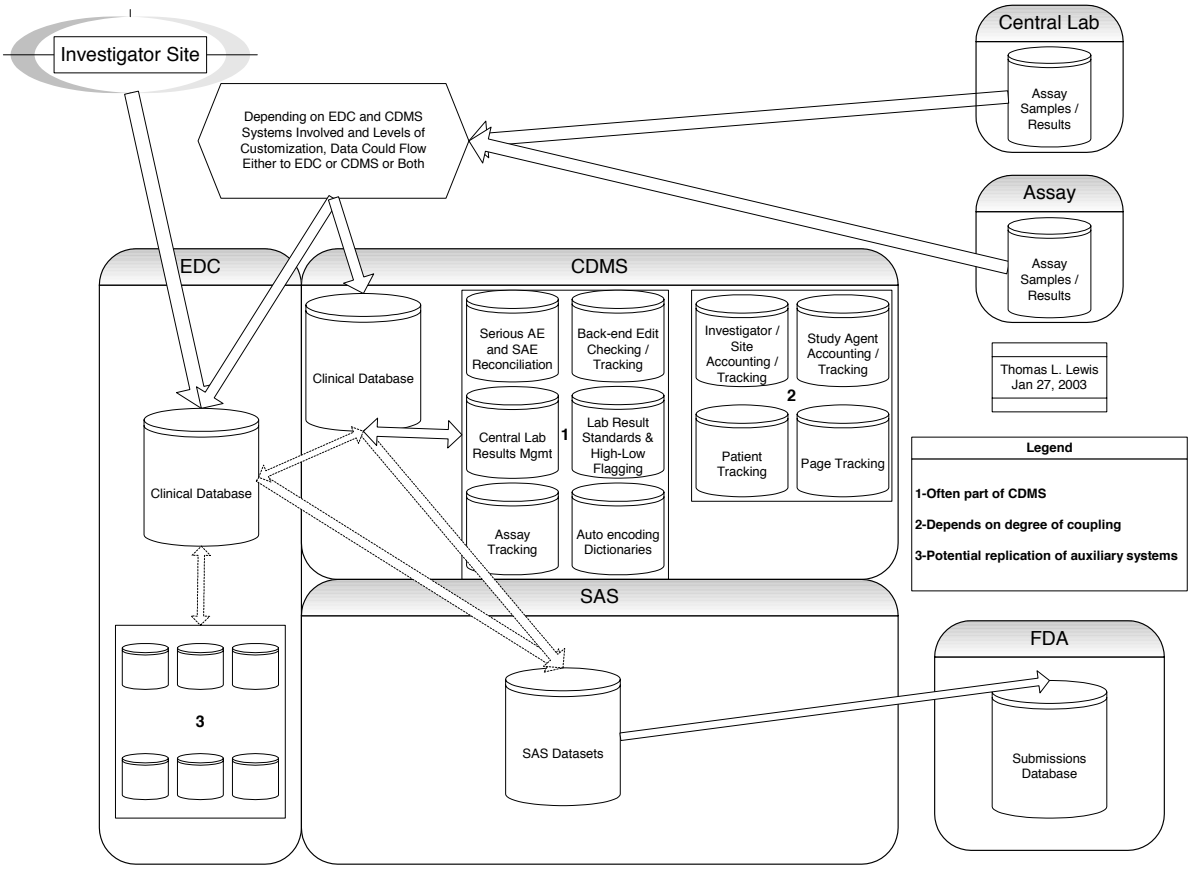

Figure 1. EDC-CDMS-Auxiliary Systems

(continued on page 3)

Table 1. EDC-CDMS Architecture

Close-coupled (EDC data stored in CDMS database)	Legacy CDMS updated to include a web-enabled component. Entails the installation of web-server hardware and other supporting services, along with the Validation / Operation Qualification of the additional hardware/software components.
Loose-coupled (Data is transferred from EDC database and uploaded into CDMS-based data repository)	CDMS-based data storage. System may allow continued use of existing data maintenance tools (e.g., re-coding of auto encoded items) and the combining of data from multiple studies for submissions.
Stand-Alone/Submission Database (SAS DataSets/Change Control)	Data transferred from EDC system into a Submissions Database (e.g., SAS DataSets) using Change Control procedures.

Each of these architectures (i.e., Close, Loose, and Stand-alone) have different trade-offs in terms of costs and benefits. (See Table 2) Many of the trade-offs center on the number of auxiliary systems and the nature of their integration with the legacy CDMS system.

In the past, a close integration between a CDMS and any auxiliary system would be seen as positive since the closer the integration, the more seamless the auxiliary system would be (e.g., when the data are shared by the two systems, duplication of entry and maintenance effort as well as synchronization issues would be minimized.) It is unfortunate that such close integration between CDMS and auxiliary system(s) may well hamper the introduction of an EDC system unless it is closely coupled to the legacy CDMS. If the auxiliary system(s) are not closely integrated to the CDMS, they may actually need less re-engineering in order to work with the new EDC system.

Table 2. Examples of Trade-offs

	Good	Bad
Close-coupled	1. Data Entry System Developer develops entry system(s) using familiar tools. 2. Auxiliary systems (including closely integrated ones) may continue to work without modification.	1. Choice of EDC products restricted to those available from CDMS vendor or from vendors that offer high quality bridge to CDMS. 2. Entails the installation of web-server and other support and services, along with the Validation / Operation Qualification of the additional hardware/software components. 3. These EDC system may not take full advantage of new design philosophy and technology. User Interface may not be optimal.
Loose-coupled	1. A wider variety of EDC products to choose from. 2. EDC system might incorporate functionalities missing from CDMS and alleviate the need for some auxiliary systems (e.g., patient tracking)	1. Necessitates the creation of new, possibly 'nearly duplicate' processes and procedures. 2. Data Entry System Developer may need to develop two entry system(s) using dissimilar tools. 3. Auxiliary systems may need extensive re-engineering.
Stand-Alone/ Submission Database	1. EDC system may be optimized. 2. Existing auxiliary systems may no longer be needed.	1. Existing auxiliary systems may no longer work and there may be a need to engineer entirely new replacement auxiliary systems that work in the EDC system environment. 2. CDMS decommissioning expenses.

(continued on page 4)

When, and if, the legacy CDMS is decommissioned, the data contained within it will most likely need to be archived. How this is accomplished will not be covered in great detail in this paper, but needless to say, the goal of data archival is to preserve both the data and a detailed description of the data in a format that will be "understandable" sometime in the indeterminate future. Traditionally this has meant the creation and storage of ASCII column-delimited files. These files would consist of not only the clinical data captured during a clinical trial, but also metadata describing the data, all associated data tables (e.g., code lists, dictionaries), and now, with the advent of 21 CRF 11 rules, electronic signature information, if it exists with the legacy data. A more current approach to archiving clinical data may be to use the Clinical Data Interchange Standards Consortium (CDISC) Operational Data Model (ODM), which would store both data and metadata in XML format.

Common Legacy CDMS Auxiliary Systems

Common auxiliary systems usually support functionalities missing from core legacy CDMS products, such as management, tracking, and non-generic reporting. (See Table 3 below for a description of many of these common legacy auxiliary systems.)

Some of the integration/functionalities listed below may be incorporated in some of the EDC products available today. Refer back to Figure 1 for a graphical representation of how these systems might tie together. In general, if the EDC, CDMS and auxiliary systems are closely coupled, significant changes in systems and operating procedures may not be necessary. If the EDC system is loosely coupled to the CDMS, there may be the need for significant amounts of auxiliary system re-engineering to support "bridging" ties or even for duplicate auxiliary system(s) working within the EDC system environment. If the EDC system stands-alone, there may be a need to engineer entirely new replacement auxiliary systems that work in the EDC system environment.

Table 3. Common Legacy CDMS Auxiliary Systems

Assay Samples/Results Tracking Management	Tracking of assay samples collected for non-central laboratory processing. An example of this sort of assay might be specialized allergy testing. This system might also be used to track when the assays were performed. The measurements (e.g., the values or results) may be entered into the EDC product; sent to the sponsor and stored within the Assay Samples/Results Tracking system, in the legacy CDMS, a stand-alone database, or in a SAS Database.
Central Laboratory Samples/Results Tracking/Management	Similar in concept to the generalized Assay Samples/Results Tracking is a system focused on Central Laboratory processing. Being more specialized, these systems may have additional functionalities built in them, for instance, management and updating of Normal Range data.
Back-end Edit/Discrepancy Checking/Tracking	Although clinical data entered via an EDC system is expected to be cleaner than similar data from a CRF/CDMS when it reaches the database, there may be edit checks that can't be (easily) written at the EDC entry form level. There most likely will continue to be a need to perform data edit-checking on data after it has been captured. The legacy CDMS may continue to provide valuable support for back-end data cleaning procedures.
Serious AE and SAE Reconciliation	Legacy CDMS were not usually able to handle the unique, expedited FDA reporting requirements of Serious Adverse Experiences (SAEs) so typically a stand-alone SAE system was implemented to perform this critical processing. Some CDMS product vendors later offered SAE support, usually in the form of a loosely integrated add-on (or stand-alone) module (e.g., PhaseForward's Clintrace™, and eRT's eSafetyNet™).

(continued on page 5)

Table 3. Common Legacy CDMS Auxiliary Systems continued...

Auto Encoding of Coded Items	(E.g., Adverse Experience (AE) verbatim descriptions, prior and concomitant medications) Since verbatim terminology is not particularly usable when analyzing clinical data, a variety of computer-based coding methods and dictionaries have been developed to encode such terminology for AE's and medications. The automation of this time-consuming manual process is a definite business advantage. A large investment may have been made in the process algorithm, implementation of the software, and/or the evolution (maturity) of the supporting dictionary(ies).
Laboratory result standardization, high-low flagging, and other semi-typical data transformations	As is the case with auto-encoding, other collected clinical data may not be the most usable form due to not being of a uniform or consistent, "cross studies", format (e.g., a specific laboratory test result not all being collected in same units due to different laboratories performing the measurement/evaluation of the sample.) Furthermore, some data may need to be derived since it was not collected directly (e.g., a laboratory measurement's High-Low Flag).
Patient Randomization	For blinded studies, patients are placed into treatment groups in random fashion. Randomization may be performed by an independent process/party. Tying randomization to the EDC and study agent tracking system may require some process re-engineering. The EDC system may provide suitable (i.e., "adequate") randomization and blinding schema(s).
Study Agent Accounting/Tracking	Study Agent is sometimes referred to as Drug. A "Study Agent Accounting/Tracking" system is one that tracks study agent inventory, study agent allocations and shipments, and/or other perceived useful study agent information. This system may be integrated with a Patient Randomization system.
Investigator/Site Accounting/Tracking	An "Investigator/Site Accounting/Tracking" system is one that Tracks investigator "contact information," site "contact information," site enrollment date, "progress in study" status, and/or other perceived useful investigator/site-related information. Many EDC systems incorporate some form of Investigator/Site Tracking. If the EDC system's functionalities suffice, re-implementing or reintegrating the legacy tracking system may not be deemed necessary.
Patient Tracking	A "Patient Tracking" system is one that tracks patient enrollment dates, "progress in study" status, CRF and data entry status, DCF status, and/or other perceived useful patient-related information. Many EDC systems incorporate some form of Patient Tracking. If the EDC system's functionalities suffice, re-implementing or reintegrating the legacy tracking system may not be deemed necessary. This system may be integrated with a Patient Randomization system.
Page Tracking	A "Page Tracking" system is one that tracks receipt of patient-related paper documentation by the sponsor. This documentation might include CRFs, DCFs, Correspondence, data entry status, CDM dispositions, and/or other perceived useful patient-related paper documentation information. Many EDC systems incorporate some form of electronic Page Tracking, however reengineering and re-implementing or reintegrating the legacy tracking system may be deemed necessary.
Study Monitoring	This usually refers to a system that integrates a collection of study tracking and, possibly, reimbursement modules.
Ad-hoc and other Reports	Oftentimes, legacy CDMS do not provide reports having the layout and/or information desired by Sponsors. If the need for any particular legacy report remains after implementing an EDC system, depending on the magnitude of changes to its data source(s) and any changes to reporting tool, the report may need significant reengineering.

Conclusion

It is important to consider the legacy CDMS and any auxiliary systems when tackling an EDC implementation project. It is important to assess not only how the new EDC system will integrate with the CDMS but also how it will integrate with existing auxiliary systems. After assessing the systems in place, incorporating any newly uncovered information into the EDC implementation plan is important. Valuable information such as the identification of functionalities that were not part of the CDMS (but perhaps should be part of the EDC system) as well as the identification of systems that may need to be reengineered in order to continue to provide required business functionalities has been presented.

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).