Electronic Data Capture—Concepts and Study Start-up

September 2008

Abstract

Electronic data capture (EDC) has emerged as a proven tool for sponsors of clinical trials. Understanding the principles of EDC is more important than ever for clinical data management (CDM) professionals. This chapter reviews the regulations and guidance that currently apply to EDC during pre-production and study start-up, and emphasizes the important role that CDM professionals have in the adoption, development, and improvement of EDC systems.

Introduction

Electronic data management for research emerged in the 1970s and has evolved into a suite of processes and tools to enhance the management, quality control, quality assurance, and archiving of clinical trial research data. In the 1990s the development of electronic data capture (EDC) tools for clinical trials research became more focused. Today, EDC is gaining in popularity, and regulatory agencies are readily accepting submissions in which validated EDC tools are used. EDC systems should be more than just a means to an end and quality EDC systems can be drivers of the entire clinical trial's information management process. Data managers provide significant value in designing processes to make the transition from paper systems to EDC systems efficient while ensuring data integrity is maintained.

The return on investment has been proven for the automation of clinical trials information management processes from data entry through the summarization and archival processes. Although remote data entry (RDE) processes emerged in the 1970s,[1] these processes languished for 20 years without significantly impacting clinical trials. By the mid-1980s, personal

computers (PCs) were introduced to clinical trials for clinical data capture, which led to a major transformation in the way clinical data was captured. Prior to that time, site professionals collected data on paper case report forms (CRFs) and sent the forms to a centralized sponsor facility where data computerization took place. This method of data capture was called "centralized" because data was entered into a computer system in a single facility by professional data entry staff. The investigators' main responsibilities were the original completion of the paper CRFs and responding to queries that arose following review of computerized data.

Having PCs at the investigator site allowed for the introduction of "decentralized" clinical data capture, which became known as remote data entry (RDE). This development began a paradigm shift in clinical trial conduct, placing the responsibility for electronic data entry on site staff. Many sponsors developed proprietary hardware and software solutions to manage RDE at investigator sites. Computerized data was routinely transferred from each investigator site to the sponsor through some type of periodic file transfer, for example, using file transfer protocol (FTP). The FTP process was usually done via phone lines and took some time to complete, depending on the volume of data to transfer.

In the late 1990s, Web-based approaches to clinical data capture were introduced in an effort to gain efficiencies that other industries had realized by moving processes to the Internet. The acronym RDE was subsequently replaced by EDC as data transfer was expedited by Internet technologies rather than FTP, resulting in more frequent and rapid data transfers.[5] The introduction of Web-based EDC led to greatly expanded use of decentralized clinical data capture.

Scope

This chapter provides information on the concepts and start-up activities related to EDC systems for companies who are considering transferring some or all processes from traditional paper data collection to EDC. It concentrates on establishing an environment conducive to incorporating EDC technologies into the clinical trial process from the viewpoint of data management. Practices, procedures, and recommendations are proposed for data managers to prepare for and start an EDC system that will properly align electronic data capture technology to support statistical and medical research needs.

Comparisons between paper data collection methods and EDC are also presented. The primary focus in this chapter is on start-up activities to support EDC for electronic CRFs (sometimes called eCRFs, although the term will not be used in this document) and data integration with non-CRF data.

Many of the tasks described in this chapter may be joint responsibilities between different groups, just as there may be many different groups involved in the implementation of various tasks. However, clinical data managers need to be conscious of whether or not these tasks have in fact been performed in a satisfactory manner.

Recommendations for proper study conduct and study closeout using EDC will be addressed in the chapters entitled "Electronic Data Capture—Study Conduct" and "Electronic Data Capture—Study Closeout." Recommendations for patient diaries and interactive voice response systems (IVRS) will be addressed in future chapters of the GCDMP.

Minimum Standards

- Ensure compliance with 21 CFR 11 and consistency with the Food and Drug Administration's (FDA) *Guidance for Industry: Computerized Systems Used in Clinical Trials*.[2, 3]

- Stated quality standards should support the utilization of automated data capture, management and archiving.

- Ensure requirements are defined for data transfers and integration with other systems.

- Software systems validation should be scheduled and completed prior to EDC study implementation.

- Ensure user acceptance testing (UAT) is completed prior to implementation and deployment to sites.

- Verify training is provided for all users of the EDC systems and that all training is documented and minimum competencies are met.

- Verify access to data is limited to authorized individuals.

- Determine roles and responsibilities in data review and query management.

- Software technical support should be provided to users and a toll free phone number should be available for the help desk.

- Ensure sites have access and control of data up to database lock.

Best Practices

- Use business process analysts (possibly external, for objectivity) to establish EDC-specific workflow processes and identify required transitions from current processes.

- Do not apply paper study processes to studies using EDC.

- Identify stakeholders in current processes, as well as additional stakeholders required for new EDC processes.

- Plan studies to avoid "last minute" system modifications that introduce errors and complexity to study-specific CRFs.

- Develop CRFs or data collection tools with teams of individuals from monitoring, data management, statistics, regulatory affairs, and medical, ensuring adequate attention to the collection of safety data.

- Ensure systems are user-friendly and flexible for data entry.

- Ensure EDC systems do not restrict answers site staff can provide in a way that introduces bias into the clinical study.

- Ensure adequate edit check procedures and query management tools are built into EDC software.

- Before the start of a study, conditions (e.g., SDV completed, all queries resolved) for locking forms and/or casebooks should be set according to a set of criteria, such as, all SDV complete, all data review complete, no outstanding queries or missing data exist.

- When coding in an EDC environment it is recommended not to display coded terms back to the site user.

● Ensure data can be traced from the time of original input through the reporting and analysis files via easily accessible audit trails.

● Ensure ease and quality of all data transfers by testing data transfers prior to deployment of EDC systems.

● Ensure your EDC system integrates as needed with other databases by testing integrations with your EDC system prior to initiating any trials using the system.

● Ensure processes are defined to integrate laboratory and other non-CRF data with data obtained from the CRF.

● Ensure all user acceptance tests are documented.

● Ensure change control procedures include complete documentation.

● Ensure all documentation for use by site staff is adequately reviewed before being provided to site staff.

● If 24 x 7 x 365 support is not available, the help desk should cover the work days/times of all regions included in the study.

● The help desk should support the minimum number of languages needed to communicate with all users and all languages, including local dialects.

● Develop and follow standard operating procedures (SOPs) for electronic data capture, data validation, and data archiving.

● Assess current SOPs for potential impact created by EDC workflow processes and update SOPs as necessary.

● Include SOP modification time in project plans for EDC implementation.

● Assume that both the new workflow and SOPs will be in transition for some period of time as the staff interact with the EDC system following any modification of SOPs.

● Identify issues that merit periodic reminders, such as user ID and password security, and schedule recurring reminders.

● Provide an instruction manual for study workflow processes.

- Verify all users have documented training prior to being granted access to the system.

- Create a training environment in which users can practice, and create training cases as examples that are pertinent to the study.

- Provide a "Train the Trainer" program for clinical research associates (CRAs), data managers or others to be able to provide training to sites.

- Provide training customized to each user's role. A study coordinator may need in-depth training of most system functions, while users with read only access may need minimal instructions.

- Document all training for trial master files as well as site files.

- Integrate metrics on process and cost/benefit into the EDC process to enable better EDC versus non-EDC comparisons and comparisons across EDC technologies.

- CRF specifications should be finalized prior to finalization of edit check specifications, although development of both should be performed concurrently.

Differences Between EDC and Paper-based Studies

Four important areas that differ between EDC and paper-based studies are the manner in which data will be collected, the timeline necessary to prepare for the study, the manner in which collected data will be verified, and disaster recovery planning.

Offline vs. Online vs. Hybrid Studies

The three primary modes of capturing data for a study are:

- Offline—the traditional paper-based method for collecting, sending, and collating or an EDC system that works without a constant Internet connection.

- Online—the EDC method, typically using networked resources to record clinical data in electronic forms, which are then stored at a central server location

- Hybrid—a combination of offline and online methods that are either a combination of paper-based systems using EDC to manage some aspect of the data-collection process, or that involve the use of both offline and online EDC methods

The mode chosen is generally dependent upon the capabilities and limitations of the sponsor and EDC software used, as well as of sites that will participate in the study. Therefore analysis and planning are essential to determine which mode should be used for a given study.

EDC solutions are inherently technical implementations that vary in their degree of complexity and level of competence required by users. The EDC process extends data collection (and in some situations, data cleaning) to the site and/or subject. It is critical to accurately assess the ability of sites to use and manage the technology on which the EDC application is based. If it is apparent that a one or more sites lack the requisite technical capabilities to use an EDC solution, the sponsor should consider a paper-based or hybrid study as an alternative.

The results of the following assessment examples, as well as any others that are pertinent to the sponsor, will guide the determination of which data collection mode is best suited to a study.

- Site readiness: including technical capability, staff training and competencies, systems infrastructure, and past EDC experience

- Edit checking complexity: the study's degree of dependency on robust edit checks and their impact on system performance

- Audit trails: the importance of capturing the *entire* audit trail electronically as stipulated in 21 CFR part 11

- Subject population: for studies that can utilize subject-oriented EDC solutions such as ePRO, an assessment of the overall subject population's ability to understand and operate the technology successfully

- Study timelines: the need for short turnaround times

- Study management strategy, for example, the level of monitoring required at each site

● An assessment of the ability of the sponsor's clinical trials management system (CTMS) to interface with an EDC solution

Study Development and Start-up Timelines

Because the study database should become active upon enrollment of the first subject, study start-up is critical for EDC studies. For those working at the sponsor facility, many start-up activities may only need to be performed when EDC is initially adopted. Many of the typical CDM start-up activities for both paper-based and EDC studies include: protocol approval, CRF design, CRF annotation, edit-check specification, user acceptance testing (UAT), and documentation preparation. The differences in CDM start-up timelines for EDC studies are based largely on the increased number of tasks that must be completed before the study may begin. In addition to typical start-up activities, several additional activities may need to be considered for EDC that could impact study development and start-up timelines, including:

● Revision of SOPs to support the EDC process (documentation preparation). For sponsors this activity may be done once, while for CROs this may occur with each sponsor with whom they contract.

● Define roles and access to data by authorized sponsor and site staff

● User account management, which may include access control logs and account management documentation

● Definition and creation of new or modified standard data collection forms

● Trial specific programming and UAT (e.g., edit checks, screen designs)

● Preparation of coding dictionaries and processes as needed

● Design, programming, and testing of reports. Establish standard reports that can be reused across compounds.

● Communicating trial status impact on timelines

● Definition and requirements testing for data transfers and integration with other systems or third party data. Utilize industry standards where possible (ODM, LAB, etc.).

- Selection of task-specific applications (e.g., a grant payment system) that may need to be integrated with the EDC system

- Site assessments for the ability to use EDC

- EDC system- and trial-specific training

- Help desk support for users

- Disaster recovery planning

With these CDM start-up activities for EDC, it is important to remember that these activities are highly cross-functional.

Source Document Verification (SDV)

The FDA has issued requirements for electronic records and signatures in 21 CFR Part 11, which provides criteria for considering electronic records as equivalent to paper records and electronic signatures as equivalent to handwritten signatures. Determining the level or amount of SDV is not within the scope of DM, however, it is important to determine if the SDV process impacts the database in any way. In principal, conducting source data verification (SDV) on electronic records is the same as for paper records. Electronic records, like paper records, must be accurate, original, legible, attributable, and contemporaneous. It is important to determine how SDV processes will function before the start of an EDC study so the database can be configured to support access, workflows and reporting requirements. Validation of computerized systems is a completely different, but very important, aspect of electronic records that must be fulfilled as well.[2] Some systems will allow different SDV strategies and some will not. This needs to be agreed upon up front in case any study specific configuration is required.

The *ICH Harmonised Tripartite Guidelines for Good Clinical Practice*, the *WHO Guidelines for Good Clinical Practice for Trials on Pharmaceutical Products*, and the *Code of Federal Regulations* require that source data verification must occur for all clinical trials in phases I–IV. An evaluation of the conformity of data presented in CRFs with source data, SDV is conducted to ensure data collected are reliable and allow reconstruction and evaluation of the study. The SDV responsibilities of the principal investigator, sub-investigator, study coordinator, monitor, quality assurance auditor, and the

clinical trial manager must be made clear at the outset of the clinical trial, and adequate training should be provided to all staff involved. So there are no misunderstandings or errors when SDV is undertaken, special emphasis should be placed on confidentiality and direct access to data. All staff involved must realize that SDV adds to the scientific and ethical integrity of a clinical trial.[1, 3] Records of what was done and found, including an evaluation of findings, must be made in the same way as for any other aspect of the trial.[4, 5]

In the SDV process, information reported by an investigator is compared with the original records to ensure that it is complete, accurate, and valid. Strictly speaking, every item of data that appears in a CRF should be documented somewhere else to allow verification, audit, and reconstruction. The main objective of SDV is to confirm that the information collected during a clinical study is complete, accurate, reliable, and verifiable so as to give confidence to the sponsor and the regulatory authorities in the data being used to support a marketing application. SDV is also required to provide confidence in any data reported, for example, in published manuscripts and at scientific conferences. Without SDV or stringently controlled electronic source data collection methods, no scientist can have confidence in the data presented and in the conclusions derived.[4, 5]

All information in original records of clinical findings and in certified copies of original records are necessary for the reconstruction and evaluation of the trial. These records may include hospital records, clinical and office charts, laboratory notes, memoranda, subjects' diaries or evaluation checklists, pharmacy dispensing records, recorded data from automated instruments, microfiches, photographic negatives, microfilm or magnetic media, X-rays, subject files, records at the laboratories and at medico-technical departments involved in the clinical trial, observations, and documentation recording activities in the clinical trial. The following data are considered key data in SDV and any gross errors in these data might be detrimental to the scientific and ethical quality of the clinical trial:

- Primary efficacy data

- Inclusion/exclusion criteria

- Medical and medication history

- Physical examination and vital signs

- Visit dates

- Adverse events

- Concomitant medication

- A record that the patient has entered a clinical study and the date of informed consent

Disaster Recovery and Business Continuity Planning

When determining a move to EDC, ensure that your facility and selected vendor has a plan in place for Disaster Recovery. Disaster Recovery Plans (DRP) are very similar between EDC and paper-based trials, but are always a key consideration. In the context of this section, a disaster is an event that significantly disrupts operations, either temporarily or permanently. This event could be due to fire, theft, or a weather-related incident that removes access to data on the servers; the sudden unavailability of key members of internal or external (e.g., vendor) staff; or the EDC vendor becomes insolvent. The goal of an organization's Disaster Recovery Plan should be to minimize the loss of operational control in the event of a disaster and to restore business activities quickly with minimal disruptions. As no single response pattern is appropriate to all organizations, a DRP should be flexible in its design.

Data management should cooperate with the information technology (IT) department to ensure that a plan is in place for all hardware and software being used to implement the EDC system. The location of all components of an EDC system should be known and documented to ensure that each of these possible points of failure are addressed by the DRP.

A tiered DRP establishes different levels of response for different levels of failure. Examples of tiers, listed by increasing levels of severity, might include the following:

- Localized failure—one system drive becomes nonfunctional

- Server failure—an entire server becomes nonfunctional

- Office building failure—all resources become unavailable at a building where business operations are conducted

- City failure—all resources become unavailable within a geographic region

Accompanying the DRP should be a Business Continuity Plan (BCP) that guides continuation of a study during the recovery of failed systems. Depending on the number of EDC vendors and contract research organizations (CROs) used by the sponsor, the BCP may be included in the EDC project's data management plan, or exist separately as a plan applicable to all EDC projects. The BCP should identify alternative processes in the event the EDC system becomes temporarily or permanently unavailable. For example, a project may revert to faxing paper CRFs and queries. The BCP should also establish the process by which sites will be informed of EDC system downtime, and the alternative method for collecting data while the system is unavailable.

EDC Deployment Considerations

Considerations for deployment of an EDC system should be taken into account at the organization level and used when researching, interviewing and assessing EDC vendors. Considerations for this step in the EDC process fall into three main categories:

- Understanding different types of software technology (pdf-based, XML-based, etc.)

- Understanding different EDC system capabilities

- Researching general information about vendors

Thin Client and Thick Client Technology Comparison

There are several issues to consider when selecting an EDC vendor and client-server application for a clinical trial. A key decision is whether the bulk of the workload will be done on the client (investigational site) computer or on the server. This decision can determine the costs of clients and servers, the robustness and security of the application as a whole, and the flexibility of the design for later modification.

A thick client (also known as a "fat client" or "rich client") is client software that performs the bulk of data processing operations and does not necessarily rely on the server. For example, a word processing program installed on a personal computer is an example of a thick client. All documents are created and stored on the PC without the need for processing by a server. For study coordinators to perform data entry, the thick-client approach requires software to be downloaded to computers at the investigational sites.

The use of the thick-client approach introduces a number of challenges. In today's hospital environment, concerns for privacy and security must be considered. Users may not have administrative rights to install software and existing firewalls may block communication with servers. A thick client may require the use of dedicated internet connections and provision of IT hardware. Each installation of the client software requires validation in accordance with 21 CFR Part 11. Thick clients can encounter versioning issues, as users have to connect to the remote server to retrieve software updates, and accurate records must be kept to ensure all users are using the most current, approved version of the client software. Also, if a user must synchronize the client with a central server to submit data, contact may be lost before synchronization is complete, resulting in inconsistencies.

However, advantages of thick clients include the following:

● Local processing: Complex logical checks and coding can be carried out immediately.

● Less burden on server: Because thick clients handle much of the application processing, a thick client does not require as high a level of server performance as a thin client. The use of a thick client also reduces server loads by being able to use the client machine for processor intensive tasks like reporting and analysis.

● Better multimedia performance: Thick clients have advantages in multimedia-rich applications, which are bandwidth intensive when delivered over a thin client.

● Flexibility: On some operating systems, software products are designed for PCs that have their own local resources. Running such software as a thin client can be difficult.

- Thick clients allow sites with low bandwidth to still remain electronic and transmit their data on-demand.

In client-server architecture, a thin client depends primarily on the server for processing activities. For example, thin clients include browser-based EDC platforms that require the user to log on with the combination of a user name and password. Information is entered and stored centrally, and no data are retained on the investigational site's PC.

There are several advantages to using a thin-client model. The study coordinator does not have to use one specific computer to access and enter data. This capability is especially helpful when staff must share a limited number of PCs. Installation of study-specific software is not required, and centrally managed updates and patches ensure all users have identical client software. Dedicated network connections are no longer considerations, allowing for much greater user flexibility.

Additional advantages of using a thin client include the following:

- Lower IT administrative costs to the project: Clients are managed almost entirely by the server, and the hardware has fewer points of failure. The local environment is highly restricted, thereby improving protection from malicious software.

- Easier to secure: The client can be designed so that application data is only displayed in the browser but never resides on the client PC in any form.

- Lower hardware costs: Thin client hardware does not contain a disk, application memory, or a powerful processor and therefore can go long periods without needing an upgrade or becoming obsolete. The total hardware requirements for a thin client system are usually much lower than for a thick client system. With thin clients, memory can be shared.

- Less network bandwidth: Since terminal servers typically reside on the same high-speed network backbone as file servers, most network traffic is confined to the server room. When a thick client is used, a large file that is accessed may be transferred in its entirety from the server to the client PC. When the same file is saved or printed, the thick client sends the entire file over the network. If sites have limited access to bandwidth, this process can be highly inefficient. When a thin client is used, only mouse movements, keystrokes, and screen updates are transmitted between the

server and end user, thereby enabling large files to be accessed with far less bandwidth.

● More efficient use of resources: A typical thick-client is designed to handle the maximum processing needs of the user. However, such a design can be inefficient when allocated processing resources are not fully used. In contrast, thin clients are typically designed to use only the amount of resources required for the user's current task.

● Simple hardware upgrade path: In the thin-client model, if the peak resource usage is above a pre-defined limit, boosting resources is a relatively simple process (e.g., adding another rack to a blade server), and existing units can remain in service alongside new units. However, in the thick-client model, resolving this issue may require an entire PC to be replaced, resulting in downtime for the user and concerns regarding the secure disposal of the old unit.

When using thick clients, the following questions must be addressed prior to implementation:

● Will the site's IT department permit external software to be installed?

● Will the site's network firewall and security systems interfere with communication between the client and server?

● Who will be responsible for maintaining software and ensuring updates are provided? Will maintenance result in any downtime for users, and if so, how will downtime be managed?

● Will a dedicated PC or internet connection be used for the study? Does the study's budget include the cost of these resources? Does the site have space for the equipment required by a thick client?

● Will there be any restrictions regarding use of Internet access, such as periods when the investigational site staff are unable to connect to the Internet due to scheduled network maintenance?

● Will technical support be provided, and if so, by whom?

When using thin clients, the following questions should be addressed prior to implementation:

- Will the site's network firewall and security systems interfere with communication between the client and server?

- Who will be responsible for maintaining software and ensuring updates are provided? Who will be responsible for maintaining records regarding new updates?

- Who will be responsible for ensuring browser compatibility? If a site does not have a compatible browser, how will this issue be addressed?

- Will a dedicated PC or Internet connection be used for the study?

- Will there be any restrictions regarding use of Internet access?

Application Service Provider (ASP) vs. Technology Transfer

The decision to use either an ASP or technology transfer model of EDC depends largely on the sponsor's long-term strategy. The determining factors are usually based on the frequency of use of the software (e.g., how many studies for which it will be used) versus the cost of purchasing and maintaining the software.

Application Service Provider

An ASP is essentially a company that offers its software for use by another company at a cost. The software itself is not purchased, only the opportunity to use that software. The vendor retains full ownership of software, and the client pays for it on a "per use" basis. When an organization uses EDC software in the ASP model, the software resides on the vendor's hardware and under the vendor's authority. It is accessed by the client through a browser or other client software provided by the vendor.

The incentive for the sponsor to adopt an EDC system that uses the ASP model is that implementing, hosting, and validating software are left to the vendor, as are the issues of upgrading, maintaining, and supporting the software. A risk-based approach should be used to determine the scope and depth of any additional sponsor software validation to be performed. The ASP pricing structure takes all of these issues into consideration, and therefore ASP pricing per study is typically higher than when using a technology or knowledge transfer system.

Some advantages to using an ASP model for an EDC system include:

● Little or no setup time is needed, and limited or no software integration is required to begin using the client software.

● Pay per use of the software or "pay as you go"

● Costs for software development and upgrades are shared by multiple clients rather than solely by one client.

● In-house experience is not required, and niche employees do not need to be hired.

● Vendor handles the challenges of system up-time, reliability, security, and scalability.

● IT costs are maintained at a predictable level, and fewer expensive or specialized IT staff are required.

● Installation of heavy infrastructure is not required.

Some disadvantages to using an ASP model for an EDC system include the following:

● Clients must usually accept the software "as is". Customizations usually do not occur unless several clients have made the same request and the vendor is willing to change the software.

● Loss of control of data: because the software is owned and maintained by the vendor, clients must ensure that proper service level agreements for system up-time and application availability are in place.

Technology Transfer

In a technology transfer scenario, software under one company's authority is moved to the environment of the sponsor or CRO. Many different levels of technology transfer are possible, ranging from transfer of just the build of a study to transferring all services from a vendor. Alternately, a sponsor or CRO may bring only certain services in-house, such as the help desk and user training. Traditionally, most companies handle building the study but not its

hosting. A sponsor determining whether to bring software in-house should consider the following questions:

- Is the sponsor ready to build studies internally?

- Is the sponsor able to provide hosting services?

- Can the sponsor provide help desk services to end users?

- How many trials using EDC are planned for this year and subsequent years?

- Does the sponsor have sufficient IT staff to provide technical assistance?

- Does the sponsor have trainers to provide necessary skills to users of the software?

- Does the sponsor have a dedicated project team to handle implementation?

- What is the overall scope/timeline for implementation of the EDC system, and can these deadlines and goals be met?

EDC System Capabilities

Adopting an EDC system offers an opportunity to implement features and functions that enhance operations. To ensure success, software must be qualified and validated, and key features thoughtfully addressed.

Software Qualification and Validation

Because CRFs are used by many end users, the functionality of each CRF should be tested and validated to ensure that data are entered, assessed, cleaned, committed to the study database, extracted and delivered in a known, regular, and repeatable fashion. Issues to consider when planning and executing the validation of an EDC system include the following:

- Whether the EDC application is "off-the-shelf" or custom-developed

- The amount of application validation that has been performed previously. A risk-based approach should be used to determine the level of validation to be performed.

- Notably, the amount of validation and qualification required for an EDC system can affect the time to start-up, the cost of start-up, site initiation and qualification, site maintenance, and software patch maintenance and deployment.

Support of Library Functionality

Library functionality is the ability to reuse forms, fields, edit checks and other functions within EDC software. The ability to reuse pieces of a study for newly developed studies will allow you to gain efficiencies in the design and build process. Choosing a system or developing a process that supports a library of CRF components will greatly enhance the speed at which you can develop studies.

Electronic Investigator Signatures

Prior to creating a study's process workflow, electronic signature capabilities of the EDC system being considered must be clearly understood. Questions to be answered may include:

- Can an investigator signature be applied at the form level, visit level, or casebook level? Is there a mechanism to easily tell which CRFs, subjects, or casebooks are awaiting signature, which have already been signed, and which have been edited since signing?

- Does the system send notification that a page is ready to be signed or has been signed? How is that notification delivered? Must the investigator be logged into the system to receive the notification?

- Can the investigator reject a request for signature and provide a comment with the rejection?

- Does the EDC study workflow require multiple signatures? If so, at what level (CRF, visit, and casebook)? Does the system have the ability to apply multiple signatures at this level?

- Does the system workflow automatically create signature notification based on specific status flags? If so, can this workflow be modified?

- Can a CRF be signed with open queries?

- Can frozen or locked forms be signed?

- Can the signature capability be turned off for a study?

- Is there an ad hoc search feature for all system users to filter by signature status flags? For example, not awaiting signature, awaiting signature, forms that have been signed, no longer signed, and so on. Can this search be performed at site level, visit level, by subject, and/or by specific form?

The ability to quickly search in this system interface may be particularly useful for CDM and CRAs when ascertaining a site's completion status at the end of a defined visit, and also at the end of the study.

Electronic CRF Archiving

The method of transferring an electronic CRF to a read-only format for sites to use in the future must be determined. Factor the process for obtaining the CRFs into the contract's study timelines and expectations. See the chapter entitled "Electronic Data Capture—Study Closeout" for more information.

Export Formats

Export formats as well as restrictions to the availability of exports should be documented. Possible formats for exported data may include the following:

- Clinical Data Interchange Standards Consortium (CDISC) Operational Data Model (ODM)

- Microsoft® Access

- SAS®

- American Standard Code for Information Interchange (ASCII)

- Character delimited files

- Oracle®

- Extensible Markup Language (XML)

- Microsoft® SQL

The timing and delivery of exports is important, therefore the process for exporting and delivery of data should be robust and flexible.

Integration

Utilizing EDC has added complexity to data integration needs. Data managers must now understand how data collected or maintained outside an EDC system will be used, who will use it, and for what purpose it will be used. Knowing the answers to these questions will determine the path integration efforts must follow. In the event data integration does not occur as expected, a clearly defined roll-back plan should be established. To ensure project goals are met, the data manager must articulate these needs to technical or IT staff in clear terms. This section discusses considerations for various types of data integration that data managers may encounter during an EDC study.

Clinical Data Management System (CDMS) Integration

Unless a fully integrated EDC or data management solution is being purchased, data managers must consider how an EDC system will integrate with new or existing data management systems. The EDC vendor may be able to help with integration through an add-on component specifically designed to meet the system needs. Some organizations should consider a custom solution that will involve technical and/or IT staff. Integration should encompass data and queries, while avoiding manual transcription of queries into the CDMS when automated edit checks occur in the EDC system.

Integrations should also consider the reporting needs for EDC data. Data from EDC, ePRO, an external vendor or other sources oftentimes must be viewed together to assess data quality. A third party reporting tool may be needed to achieve this, or your organization may need to rely on clinical programming or other support groups to merge data via SAS.

SAS® Integration

The data manager, in collaboration with other functions, should decide whether EDC data will be directly integrated into the SAS® environment, or first integrated with a back-end CDMS.

ePRO Integration

If patient reported outcomes will be collected via the Web, an e-diary device, or other data device, data managers should consider where and how data will be integrated with CRF data captured through the EDC system. Many EDC systems can import bulk data from external sources. If data collected using ePRO is of interest to the investigator, it may be worthwhile to upload ePRO data feeds into the EDC system. Integration of external data into the EDC system may also facilitate the archival and submission process by enabling all data to reside in one CRF. Consideration must be given to integrating ePRO data that has the potential to unblind the study.

CTMS Integration

Integration of the EDC system and the CTMS can be a powerful way to gain efficiency in the conduct of clinical trials. Specifically, the data manager may want to integrate user account management. If site staff information is already being captured in the CTMS, this information may be transferred to either a help desk or directly into the EDC system, thereby eliminating manual creation of EDC accounts. Additionally, integration of visit information from the EDC system to the CTMS can facilitate monitoring and tracking of patient enrollment and completed patient visits. In turn, this information can be used to trigger site payments and grants. Integration of EDC with the CTMS also creates an ideal way to consolidate metrics used to assess overall trial performance.

Paper Component Integration

If data is collected using paper CRFs, a method must exist to integrate these data with data collected using the EDC system. In most instances, data collected on paper is integrated into the back-end data management system. In some cases, it may be more appropriate to merge the data using a SAS® environment. Several EDC systems now also have the capability of integrating paper data entry into the same EDC database with EDC data.

Laboratory Data Integration

Even if central laboratories are used, it is sometimes helpful to have all or key laboratory parameters available to site staff within the EDC system. The data manager must consider this need with the clinical team. Having all data stored

in one database can facilitate more robust edit checks across other CRF data in the EDC system.

External, Non-laboratory Data Integration

If data such as an electrocardiogram will be received from external vendors other than central laboratories, data management should analyze the importance of data integration. As with ePRO integration, if sites require access to this data, the data manager should plan on uploading data into the EDC system. More information on this topic may be provided in other chapters.

Other Important Integrations

As new technological tools are developed constantly, it is important to be mindful of other systems that may need to be integrated with an EDC system. In addition to the integrations discussed above, data managers should be aware of the need to also integrate an EDC system with coding, IVRS, and reporting tools other than SAS®.In the future, electronic health records (eHR) may also become an important consideration.

International Study Considerations

EDC systems are routinely used in international studies. The role of data managers in international EDC trials is similar to the role played in paper studies. However, planning is critical if the deployment of CRFs and hardware is to be completed prior to the first site initiation. Data management must work with clinical research to understand language needs of the CRF or any components of the CRF. Issues to consider include the following:

- Ascertain whether the local language can be used in a multi-national study. Many coordinators speak more than one language. Data management may avoid unnecessary work by asking this simple question or challenging the status quo in this area.

- Plan early for CRFs that must be programmed in multiple languages. Significant lead time is required to translate CRFs and verify translations.

- If applicable, ensure hardware deployment timelines are increased. Country specific laws may delay shipments significantly.

- Establish a plan to manage time zone differences, especially in relation to time and date stamping.

- Ensure hardware and software can be used at study sites, and that sites are prepared to use the tools that will be deployed to them.

- Develop a plan to manage system upgrades, which is particularly important if the system is being used 24 hours a day.

- Consider the wording of manual queries to ensure they will be understood by speakers of other languages.

- Consider issues posed by language barriers to staff training. For example, investigator meetings could provide simultaneous translation for all languages spoken by participants, a train the trainer strategy could be employed, or training materials could be translated into the users' native languages.

System and Vendor Assessments

For most organizations, moving to EDC is a significant decision, with an effect that is not limited to only data management. When launching an EDC system assessment, stakeholders from different parts of the organization are needed to develop the requirements checklist. Included in the requirements checklist should be:

- The types of data formats to export from the system

- The formatting and process for final archived CRFs

- The required software functionality, such as types of edit checks and the process for obtaining investigator signatures

- Details of service level agreements (SLAs)

This list should include a minimum of topics and the priority of each requirement (e.g., necessary or "nice to have"). A suggested method for determining overall requirements for the software and vendor is a grid to

which perceived values can be added. Additionally, identifying the availability of alternatives for some requirements may be useful, as selected vendors and software may not meet the requirements precisely.

The suggested list of minimum topics for the vendor/system assessment grid is as follows:

● About the EDC system:

☐ User friendliness of the EDC system's interface

☐ Study start-up process, including time, expectations, and what is included

☐ Configuration limitations and the amount of customization that will be required

☐ Hardware provisions and associated costs

☐ Variable costs

☐ Upgrade options and restrictions

☐ Process for change management

☐ Capability for establishment of a standards-based library

☐ Reporting capabilities

☐ Export formats available

☐ Integration of IVRS and laboratory data (if needed)

☐ Types of edit checks possible (e.g., cross-form, cross-visit, dynamic)

☐ Handling investigator signatures

☐ Process for data archiving at the end of the trial

● About the EDC vendor:

☐ Stability of the vendor company

☐ Software and system validation and validation approaches

☐ Prospect of the vendor's continued existence for conduct of the intended study

☐ The vendor's offered help desk support, if needed

☐ Languages offered by the vendor for training of EDC users

☐ Languages offered by the vendor for support to EDC users

If a vendor is sought for the purpose of securing an engagement for more than one project or product, the grid should also include:

● The vendor's approach to addressing your product portfolio needs

● A comparison of the vendor's EDC tool suite and roadmap to the study sponsor's EDC strategy

Once developed, the grid can be evaluated to identify company requirements. The following sections will further discuss the requirements listed above.

For more information, see the GCDMP chapter entitled "Vendor Selection and Management."

Other Considerations

At the point in time that you are still investigating a move to EDC, the following topics should be discussed with your prospective vendors.

Change Control Process

These plans do not need to be fully developed prior to selection of a vendor, but you will want to know and understand that your vendor has a structured, detailed, and documented plan for change control. This includes change control for software upgrades as well as for midstudy amendments. A well developed plan should be in place for both types of change control. For a more detailed discussion of change control processes, see chapter entitled "Electronic Data Capture—Study Conduct".

Escrow Agreements

Assess if there is a need for your company to have an escrow agreement put in place with your vendor. Some companies offer source code escrow services that assure the Licensee continued availability and usefulness of the software in the event that the software vendor fails to maintain the license agreement, or becomes insolvent.

Related Services: Hosting and Help Desk

A critical factor in the success or failure of an EDC study is the technical support received by users encountering problems with the system. Technical software support is often managed through a help desk. If using an outside vendor, standards and expectations for all trials using the vendor's help desk software should be documented in the vendor's contract for this service.

As the EDC market expands, EDC vendors continue to add functionality that makes their solutions unique. However, most vendors offer the services of hosting and help desk. How services are provided and fees for services can be very different among vendors.

The basic questions of how the help desk will be handled needs to be discussed with your team at this stage of deployment. You may consider bringing EDC help desk functions in-house with your organization, or you may prefer to use an outsourced help desk. The outsourced help desk could be facilitated by the EDC vendor you choose, or you could enlist the services of a technical call center or help desk. The sponsor's decision to use an internal or external help desk is primarily determined by the amount of internal resources available to perform this function and the level of response time or needs that the study or project dictates. For example, if an EDC study is of a very simple design, of short duration, and has a small number of sites, the use of an internal help desk may be the best choice, so long as qualified staff are available to take inbound calls. Internal help desk agents will have a better understanding of the company's policies, procedures, and protocol design elements. As good service will have a direct impact on their employer and future business with that particular user or site, internal staff will also be more interested in determining the quickest and best way to resolve an issue.

A global study with numerous sites, complicated and varied forms, and new users represents an opposite example. Use of an outside help desk could

provide advantages, such as covering additional time zones and languages without taxing internal resources. Moreover, external help desk agents are typically evaluated on performance of help desk ticket resolutions, and will have a vested interest in being courteous to users. However, they will not be able to address clinical-related questions, which will need to be forwarded to internal resources.

Vendor SLAs and Performance Reports

When using an external vendor, the sponsor must emphasize the writing and managing of service level agreements for performance. A contract should be established between the sponsor and vendor providing help desk services (such as the EDC vendor or other outsourced agency). This contract and/or SLA should include, but is not limited to, the following identified functions and associated costs:

● Core languages covered

● Translation fees for additional languages

● Vendor project management fees

● Portal or other web access to see open/resolved calls and problem resolutions

● Computer telephony costs

● Determination of fees based on studies vs. by site or by call

● Documented allowable number of calls per month

● Documented allowance for number of inappropriately handled calls, which should be no higher than 4%–6% of all calls

● Study setup fees, if applicable

● Documented percentage of expected system uptime

Several reports can assist with the management of a help desk, especially an external help desk. These reports include:

● Aging reports

- Escalation reports

- Summary of activity per week

- Pareto analysis of problem areas to address (Pareto analysis is a statistical technique used for selection of a limited number of tasks that can produce a significant overall effect. It is based on the Pareto principle, which says that by doing 20% of work one can generate 80% of the advantage of doing an entire job.)

- New tickets per week report

- Ticket closure patterns

The main point to remember at this time is that technical support for end users is crucial to the success of your move to EDC. Ensure that your first level of help desk coverage is available to all users, has enough language coverage to accommodate your sites and that the hours of support are sufficient for your user community.

Detailed Help Desk Planning

Once an EDC vendor is selected, roles and responsibilities of the help desk staff and the sponsor's staff must be established. Consideration should be given to the number of help desk staff available. The number of studies conducted by the sponsor will determine the number of help desk staff required and indicate whether help desk services should be provided by an external party. If only a small number of studies require support, it may be feasible for the sponsor to provide help desk support with internal staff. The timeframe during which users should be able to contact the help desk must be considered. Typically, 24/7 coverage is not required unless the EDC system is deployed globally. When the help desk is provided by an external organization, service level agreements should be established concerning the timeframe in which each call will be answered, as well as any other metric your organization feels is important.

Software support is commonly separated into different levels or tiers based on the technical expertise needed to correct the issue. Tier 1 software support is the lowest level of support needed and includes activities such as unlocking

user accounts and resetting user passwords. Because it is the most common support required by users, tier 1 software support is vital to a study. Users that require this level of support are often unable to access the system in any fashion. Therefore, to minimize the negative impact to both users and study conduct, it is critical to provide assistance to these users as soon as possible.

For various types of anticipated user issues, clear escalation paths must be identified for second and third level support. Data managers frequently serve as the second or third level of help desk support for EDC studies. The most common issues escalated to data managers are trial-specific data entry or query resolution issues. The data manager should be prepared to discuss the problem's solution with the level one help desk agent or with the user directly. This new role may require multilingual expertise from CDM. This new role also strengthens the relationships between CDM, clinical researchers, and sites. Cooperation among all three parties may be required to solve problems related to the EDC system. To ensure that users are satisfied with the EDC system, CDM should ensure that help desk escalation procedures are followed and working correctly.

Examples of issues and their required level of support include:

- Account activation: usually requires only level one support

- Technical error messages: may require level two or level three support

- CRF design issues: level two or level three support is typically provided by CDM

- Data entry issues: sometimes may be handled by level one support, but many are escalated to CDM as a level two or level three support issue

- Supporting systems issues: usually escalated to the IT department

- Query resolution issues: usually escalated to CDM

Deployment of computer equipment and Internet connections can also be handled by your help desk. However, these services can drain resources in day-to-day operations and involve the complexities of international shipping and tracking.

In addition to the issues listed above, data managers should ensure that specific information is provided for issues concerning account management, tier one software support, and requirements for multilingual capabilities.

Tier-One Software Support

Steps must be taken to ensure that training materials for help desk staff are complete and clearly identify the correct issue-escalation procedures. Ideally, help desk staff should be trained on usage of the EDC system. For each study, data management should provide help desk staff with a document outlining study-specific areas of concern, such as common issues with data entry encountered by users. This document will enable help desk staff handle calls more efficiently and will minimize issues that are escalated to data management.

Tier 1 support may be needed whenever a user attempts to access the EDC system and encounters a problem. Therefore, the support center must be available whenever users will access the system. At a minimum, standard business hours should be represented (e.g., in the United States, 9:00 AM to 5:00 PM), but even determining what standard business hours are for specific users can pose a challenge since sites can be located in different time zones or countries. Another consideration is whether or not support will be available on weekends and holidays. While the gold standard for support availability is 24 hours per day, 7 days a week and 365 days per year (24 x 7 x 365), this ideal may provide significantly more coverage than is needed and unnecessarily increase help desk costs. As with language localization, help desk availability must to be determined prior to the start of the study.

Providing Toll-Free Support

Tier 1 software support most commonly involves individual users contacting a support center or help desk for assistance. To ensure convenient access to technical assistance, users should be provided with a toll free phone number or calling cards to contact the help desk.

Multilingual Capabilities Required

To handle calls in international studies, the help desk staff should be fluent in applicable world languages.

To be effective, the help desk managing tier 1 support must be able to successfully communicate with all system users. Clinical studies are frequently multilingual and it cannot be assumed that all users will be conversant in English. Several options are available when determining how multiple language needs will be addressed. One option is for the help desk to fully support any and all languages with on-site support staff. Another, more frequently used option, is for the help desk to support several of the more common languages with on-site staff and to make use of a translation service to provide access to translators fluent in languages less likely to be needed. If providing multilingual help desk support is not possible, CDM representatives should discuss this issue with the CRO local to the international site. In many cases, monitoring staff may be fluent in local languages and can handle certain types of support.

Gap Analysis between Existing SOPs and EDC Requirements

It is critical to determine how implementation of an EDC system will require changes in the sponsor's current set of SOPs and other controlled documentation. Identifying these gaps is an effort of technical and clinical operations that must be shared among all stakeholders.

At a minimum, requirements should be written for each new process that has the potential to impact study data. These requirements must later be tested and will form the basis of validation efforts. Functional requirements must be developed to test overall functionality of the solution, and business requirements must be developed to test how the solution meets needs of the sponsor. Examples of procedures and processes for which requirements, testing, and validation should be performed include data entry, data verification, discrepancy management, data lock, user roles, user security and privileges, data reporting, subject freezing or locking, database backup and recovery (if not covered elsewhere), financial reports, study design of data objects, edit check procedures and derivation procedures.

- In conjunction with the set of requirements you establish for your EDC solution, you can also use the metrics and performance targets you have already established. These can aid in analyzing the set of SOPs that must be modified to include new practices for EDC. The primary stakeholder should be responsible for driving each new or updated SOP, however all

impacted stakeholders should have review and approval status before the SOP is put into effect.

Requirements for updating SOPs and the study process for EDC include:

● Identifying metrics and performance targets

● Performing a gap analysis between current SOPs and requirements for EDC

Data management may also establish goals for EDC projects based on calculated return-on-investment. However, most organizations will find it necessary to modify their processes to accommodate EDC during the start-up phase. It should be expected that the start-up phase will be iterative and will be impacted by many variables, including:

● The complexity of projects implemented

● The variation between projects

● Requirements for user training

● The type of EDC system implemented

● The number of staff affected by the EDC transition

● The preparation required by each site

This transition phase of EDC initiation may be planned through careful selection of the first several projects that will implement the EDC solution. Limiting the number of projects to use EDC will enable the sponsor to transition smoothly into the solution and manage expectations of stakeholders as necessary. Using this model, the sponsor may identify discrete phases of EDC implementation and formulate each phase to increase the complexity of projects using EDC. Ideally, a review step can be established at the end of each phase to inform stakeholders of the structure and expectations of subsequent phases. This review step should include all stakeholders and should analyze how closely the phase met set targets for metrics and performance.

Staffing Evaluation and Staffing Change Plans

Before initiating an EDC trial, a sponsor should carefully compare the resources needed to manage the people, technology, and processes to the resources that are currently available. Any deficiency in resources should not be underestimated. Initiating a trial using EDC without identifying and providing necessary components may result in failure to meet study objectives in terms of time and cost. Moreover, the study team (including the sponsor, contract staff, and site staff) may be negatively affected. To ensure an EDC study is achievable in terms of cost, time, and quality of deliverables, the sponsor must commit to meeting the study's resource and training needs.

Resources available to the sponsor include staff (and their skill sets), established processes, and vendors. The staffing evaluation plan should analyze the sponsor organization for any gaps between available resources and the resources needed to conduct and manage the EDC study. During this analysis, the following issues should be considered:

- Will a vendor or the sponsor provide the EDC system?

- Who, within CDM, will serve as the liaison with project management to approve the EDC system's design and oversee its production?

- Who will be responsible for testing and validating the system: CDM, a clinical project manager, a clinical research associate, or a targeted site participant?

- Who will provide help desk support to system users? To what level? What type of questions will be answered?

- Who will be responsible for training staff and maintaining training records? How often will training be required?

After assessing the staffing needs to conduct an EDC trial, the sponsor should evaluate necessary changes, if any, that are needed within the organization. The role of CDM may change from a task manager (managing the data itself) to that of a project manager (facilitating data and communication flow between all other study participants). CRAs may take on more responsibility understanding the technical aspects of data collection, training site staff, and answering questions from site staff about the EDC system. Information technology (IT) and programming resources may also have increased

visibility to other study team members, especially during the EDC system configuration, EDC suitability at prospective sites, and/or the help desk communication to resolve user questions.

To facilitate these changes in staffing structure, the sponsor should:

● Perform and document a process flow evaluation addressing all study team members in the workflow, communication plan, information flow, and reports.

● Perform an evaluation of required staff and their skill sets as compared to required resources.

Metrics and Performance Targets

Ideally, performance targets set for EDC projects will be based on the sponsor's foundational reasons for switching to EDC. These targets should represent the first level objectives for EDC projects. The next set of objectives can be developed during rollout of the EDC solution and should include feedback from all stakeholders.

Data management must also identify any additional metrics that may not be applicable for paper-based studies but that will be needed for EDC projects. Examples of EDC metrics may include average time for discrepancy resolution by site, average number and severity of help desk calls, and percent of EDC system downtime.

For EDC and paper-based studies, sponsors need to determine which metrics reports are required and establish processes for collecting, analyzing, and reporting the metrics data. Decisions need to be made to determine reporting frequency, at which level metrics need to be reported (e.g., by phase of study, by site, by trial, by therapeutic area, by region, etc), and who is responsible for reviewing and assessing data results. All metric reports must be clearly defined, with definitions clearly understood by all individuals reviewing and making decisions about data. These reports should be validated against the sponsor's computer system policies and procedures, and should be standardized so they can be used for multiple trials.

The following is a list of minimum recommended EDC metric reports:

- Study build timeline metrics

- Number of subjects screened and/or enrolled

- Subject visit date to data entered in system for that visit (measure in number of days)

- Current status and history of SDV

- Number of queries outstanding (measure in number of days)

- Percent of data/visits clean

- Number of queries per site (manual and automated)

- Query frequency counts per data element

- Time from last patient visit (LPV) to all data entered and/or data cleaned

- LPV date to data lock date

Reports to Support Study and Process Management

Data are more readily available in studies using an EDC system. This availability is particularly advantageous as it enables the sponsor to be more active in managing data management workflow processes and timelines, as well as site progress and performance. Early data reporting capabilities enable the sponsor to perform more timely assessments and take action to drive productivity, improve site performance, and reduce overall study timelines.

Just as in paper-based trials, creating report and listing specifications should be a collaborative effort between CDM and other members of the research team, especially those responsible for clinical, statistical, and safety-related functions. Report and listing specifications should be documented in a data review or data management plan that clearly indicates who is responsible for reviewing listings and reports. This plan should also clearly indicate the frequency of review and what action may be taken (e.g., creating manual queries, contacting the site, or retraining the site).

Performance targets and goals need to be established at the organizational level and individual study level, as discussed in the previous section devoted

to metrics and performance targets. Expectations based on these targets need to be communicated to both the study team (e.g., data managers, monitors) and site staff (e.g., study coordinators, principal investigators) to drive improvements in data management processes, trial timelines, and site performance. Due to the technology differences, metrics goals and expectations for EDC studies can be more aggressive than for paper-based studies. For example, in an EDC study the overall study goals for query turnaround time and subject-visit-date to date-entered-in-system can be shorter and more aggressive. This difference is also true for data-management process metrics. For example, goals for the time from LPV to data lock can be shorter on average in an EDC study than in a paper-based study, because the data can be received, reviewed, and validated in a much more timely fashion.

To ensure efficient data cleanup activities, it is recommended that reports that aid in data cleaning be identified during the pre-production period, as well as reports necessary to gather metrics about study conduct processes. Reports which capture protocol deviations should be programmed at the beginning of the study and run frequently to monitor compliance. Tasks that should be completed include the following:

- In addition to a project plan, create a flow chart that outlines each report deliverable and the person responsible for approving that report design.

- Schedule meetings to review and obtain feedback on reports to be used during the study.

- Determine the metrics important to the study and design reports to capture these metrics.

- If the study involves a CRO, consider what reports will be required from the CRO.

- Define CRO performance requirements and design a report to track performance of and provide feedback to the CRO.

CDM Deliverables During Study Start-up

The following section concerns CDM deliverables during the start-up phase of an EDC study, including CRFs, edit checks, and coding.

CRFs

In an EDC system, electronic CRFs replace traditional paper CRFs. However, data captured through electronic instruments or computer software is not immediately considered electronic CRF data. CDISC defines an electronic CRF as a CRF in which related data items and their associated comments, notes, and signatures are linked electronically.[6]

The process for designing CRFs in an EDC system is integral to study start-up. Development of an electronic CRF is more complex than simply modeling a paper CRF in a word processor. Creating an electronic CRF entails designing a truly electronic form with user interface elements that reduce challenges posed by electronic data entry, as well as facilitating data collection to improve data quality. Interface elements such as check boxes, option buttons, and menus enable users to record data less likely to be queried, which is a goal of EDC.

Although organizations such as Clinical Data Interchange Standards Consortium (CDISC) have initiatives underway, such as CDASH, there are currently no widely adopted standards for electronic CRFs, and those standards that do exist are evolving.[6] However, as with paper-based CRFs, standardization of electronic CRFs by the sponsor can:

- Facilitate data exchange

- Remove the need for mapping during data exchange

- Enable merging of data between studies

- Allow consistent reporting across protocols and projects

- Enhance monitoring activity and investigator staff efficiency

- Provide increased efficiency in processing and analysis of clinical data

- Provide capabilities not traditionally available when using paper-based CRFs

- Promote reusability of CRFs across studies through development of CRF libraries

A well-designed EDC system should assist site staff with accurate entry of study data. An EDC system can be designed to guide site staff to appropriate forms and to correctly enter data into the fields of those forms. The design of CRFs should avoid the following shortcomings:

- Using multiple pages for an CRF which could be displayed on one page

- Requesting an excessive amount of information on one page

- Using unfamiliar jargon

- Using checkboxes that do not include every applicable choice

- Using codes that are only relevant to data processors

- Requiring overly complex edit checks

The following practices should be followed during the design of electronic CRFs:

- The protocol should determine what data should be collected on the CRF.

- All data must be collected on the CRF if specified in the protocol.

- Data that will not be analyzed should not appear on the CRF.

- Data required by regulatory agencies should be collected.

- Data questions should be clear and concise.

- Duplication of data should be avoided.

- Use of free-text responses should be minimized.

- Units should be provided to ensure comparable values.

- Instructions should be provided to reduce misinterpretations.

- For each question, choices should be provided to enable summary generation by computer.

- "None" or "Not done" should be available as answer options when applicable.

As mentioned previously, standards for electronic CRFs are still developing. Generally, an EDC system should be flexible enough to capture questions as they would appear on a paper CRF. Data management should be dedicated to identifying and maintaining standards for the design and functionality of CRFs to be used in the EDC system. Prior to a study's production release, a process should be established to ensure that development of CRFs adheres to these standards.

Multilingual CRFs

Even when an EDC system has multilingual capabilities, translation of text appearing on CRFs can be a time-consuming process. Data management should determine whether English is the best language to be used for text appearing on CRFs. Because many studies are already conducted in English, most sites do not object to the use of English for CRFs. However, data management should work with each site to evaluate its multilingual requirements. Translation should be considered for printed and electronic user manuals, as well as training materials supporting the EDC system.

Although users of the EDC system may speak English or have years of English training, they may still misunderstand text that appears on an untranslated CRF. Idioms that may cause confusion or may not translate clearly should be avoided. A medical linguist may be required to translate certain terminology. Ad hoc forms should be presented in the local language of the applicable site. To check the quality of translations, back-translation by a third party should be used.

Dynamic Forms

Dynamic forms appear only when a subject meets a certain criterion, or when a particular data point is entered. A common example of a dynamic form is a form for pregnancy information which only needs to be entered when the patient is female. However, because dynamic forms do not always appear to be available in the EDC system (e.g., the pregnancy form will not appear in the system if a patient is male), they have the potential to confuse users. For example, the sex of a patient might be entered incorrectly as female and subsequently changed to male. In this case, implementation of dynamic forms on the EDC system determines what happens to the pregnancy form and its data, which is no longer needed for the male patient.

Determining how to best implement dynamic forms depends on capabilities of the EDC system and complexity of the study. Clarity with the functionality of dynamic forms can be achieved through the following practices:

- Keep functionality of dynamic forms simple; sometimes the ability to have dynamic forms can create too many permutations, and may frustrate users.

- Ensure that the development team understands the challenges of designing and implementing dynamic forms.

- During validation and qualification of the EDC system, test the design and implementation of dynamic forms by entering "incorrect" data for a patient and subsequently changing it.

Dynamic Visit Structures

Dynamic visits are similar to the dynamic forms discussed in the previous section. However, instead of forms, visits become available based on data entered for a patient. In an oncology trial for example, when a patient meets a certain criterion he or she may move to a different treatment group. Dynamic visits enable this type of capability, and the same best practices should be followed as for the design and implementation of dynamic forms. Dynamics may impact data entry efficiency and system speed, so data managers should be aware of the possibility of overloading sites with confusing or complicated dynamic functionality.

Derived Variables

EDC systems can provide derived variables. This may be helpful for sites, as some data will not need to be entered by site data entry personnel. Some commonly used derived variables would be conversions from one measurement system to another (pounds to kilograms or inches to centimeters) as well as averages or computations from other entered fields.

It is necessary for CDM to communicate how these features work and help users understand the impact on monitoring. This holds true for dynamic forms as well as derived variables.

Edit Checks

The use of edit checks in an EDC system offers data managers a unique opportunity to resolve data issues by interacting directly with clinical site coordinators. Due to the ability to gain access to data shortly after it is entered, data managers and clinical personnel can initiate issue resolutions with site staff in a more timely manner. For example, direct contact by phone with site staff promotes an active approach to completing a CRF and resolving edit check issues. Direct contact with site staff also promotes an active approach to completing a CRF and its edit checks within a short period of time from each other. Moreover, during study development, data managers can truly collaborate with database developers when programming edit checks. The technical nuances can be explained by the database developer, and the data manager can provide necessary data management principles to ensure implementation of a functional edit check.

The approach to programming edit checks should be chosen during development of the EDC database specification, and in consultation with all stakeholders involved in data validation. To define and review edit checks prior to production release of an EDC study, data managers coordinate activities of clinical, IT, quality control, quality assurance, and other groups. This is essential to the correct functioning of edit checks in an EDC system.

Edit Checks in an EDC System versus a Back-end CDMS

The approach to programming edit checks depends on the architecture of the EDC system, which can typically be described as either of the following:

- An EDC front-end data capture system with a robust back-end CDMS. With this architecture, an analysis of whether to program the edit checks within the EDC data capture system and/or in the CDMS should be determined.

- Complete EDC data management system where all edit checks must be programmed in the EDC system

The following considerations concerning the EDC system architecture should also be made in determining the approach to programming edit checks:

- How complex is the edit check? Will performance of the EDC system be adversely affected if programmed in EDC? Suppose the data manager

needs to confirm that the last date of subject contact is the last chronological date in the database. In this case, the edit check program should pull all dates from each module in the database and compare those dates against the date of last contact. This type of edit check might access the underlying database thousands of times and noticeably degrade server response times.

● Are all the data available in the EDC system? For example, if coding of terms occurs on the back-end CDMS, edit checks requiring coded terms should be programmed on the CDMS.

● Will programming back-end edit checks require manual re-entry of data into the EDC system for query resolution? The resources needed to manage this activity should be considered.

● Is a reporting database structure better suited to handling complex edit checks? For example, if an edit check is too complex, it may be best handled through listings with manual re-entry of the query into the EDC system for resolution by the site users.

In both EDC system architectures (and for paper-based studies as well), consideration must be given to the edit check specification as a whole. However, in an EDC system some issues impact the site user rather than the sponsor. For example:

● Are all of the data elements needed to properly open a given edit check actually collected in the study? If all data are not present, a query may fire that cannot be closed without entry of a specific data point.

● When adding midstudy edit checks, consider any limitations the system might have. For example, edit checks added midstudy may only activate as a result of new or modified data. Therefore, the data manager should consider programming a listing to identify issues with existing data. In this case, sites should also be informed that they may be required to resolve issues identified in earlier visits.

● Edit checks that are more study specific (not standard across trials) may generate queries due to the timing of data entry. EDC systems using dynamic forms may cause such queries. Consider providing additional

training on this issue to sites, or program edit checks on the back-end CDMS rather than on the front-end EDC system.

Hard vs. Soft Edit Checks

In addition to planning where edit checks should be programmed in the architecture of an EDC system, consideration should also be given to potential types of edit checks and corresponding user responses. Edit checks in EDC can be classified into two broad categories, "hard" edits and "soft" edits.

Soft edit checks are usually cross-panel or cross-item rules programmed to allow data to be entered into the system, but checks for consistency upon data entry. If an inconsistent or missing data item is identified by the edit check, a visual indicator (e.g., color change, iconography) indicates that a new query exists and on-screen text prompts site staff to address the query. For soft edit checks to be programmed correctly, the data manager should clearly identify fields on each form for which data must be entered.

Hard edit checks can be classified as "browser" checks or "system" checks.

- Browser checks prevent entry of data that is inconsistent with the data item being collected. If a user attempts to enter inconsistent data, submission of the form will be prevented until the inconsistency is addressed satisfactorily.

- System checks prevent entering data that do not match form and/or item property settings. For example, when a field requires a number with 2 decimals, a value of "3" cannot be entered. Instead, "3.00" must be entered to satisfy the property requirement. A system check does not produce a query or present an error message to the user. As they can disrupt the data entry flow at the site, system checks should be used only when deemed necessary.

Coding

Decisions on how to handle coding of medications, adverse events, procedures, and other study data should be included in the specifications documents and in parallel with CRF development. The role of data managers in the coding process for an EDC study should be relatively unchanged from the coding process for a paper-based study. However, the process should be adapted to the technology used to perform coding.

The following best practices for coding on EDC systems should be followed:

- Data management should work with the pharmacovigilance and drug safety group to determine how data coding should be handled. For drug safety and clinical trials, coding should be handled centrally by data management. Alternatively, coding can be coordinated between drug safety and data management.

- During the CRF development process, all data fields to be coded should be identified.

- The capability of the EDC system to support coding should be understood. If the EDC system cannot handle coding, data management should establish a process to code study data on the back-end database.

- If the EDC system is capable of handling coding, the sponsor should decide whether the user should be able to see coded terms or only the reported verbatim terms.

- Ensure the clinical team understands who will be proposing terms for coding failures. For each study, it is recommended the data manager handles this activity.

- The coding team should review the design of electronic CRFs to ensure optimization for coding purposes. For example, CRFs frequently provide menus for the coordinator to enter terms. The medical coding team can assist with development of these menus so that available terms will code appropriately.

Site Evaluation and Qualification

The process for initiating an EDC study is not just a matter of the sponsor selecting an EDC vendor that meets certain business requirements. The sponsor must also consider the pricing model, management, deployment, and implementation of the vendor's EDC system. The sponsor is responsible for ensuring that sites are assessed and qualified to use hardware and software required by the EDC system. Site evaluation and qualification by the sponsor and EDC vendor must occur during start-up activities and must be seamless, especially when EDC is being implemented at a site for the first time.

The following sections detail criteria for evaluating and qualifying a site's readiness to implement an EDC system.

Evaluating Site Technical Capabilities

Aspects of the site's technical capabilities may include the following:

- Presence of a wired or wireless network: If trials are run at remote site locations without Internet connectivity, installation of an offline data capture application may be necessary.

- Location

- Physical space (if required for deployed hardware)

- Language(s) spoken by site staff: Software configuration may be required to accommodate non-English speakers

- Compatible software, hardware and bandwidth

- Experience of site staff with EDC software: Staff may lack experience or exhibit hostility toward EDC. Users that are technically challenged may require personalized training. For a site that is new to EDC, it may be useful to identify an internal champion who can facilitate adoption of the new system.

Evaluating Site Connectivity, CRA Connectivity at Sites

- Availability of a dial-up or high speed connection

- Proximity of the site's physical location to the EDC system's server

- Capability of the site to access and synchronize with the EDC system on a scheduled basis

Site Provisioning, if Necessary

- Equipment and hardware (e.g., laptops, PCs, phones)

- Training

- Training manuals

● Access IDs (e.g., user account set up using secure access IDs)

Site Hardware and Broadband Provisioning if Necessary

Good clinical practices advise that site assessments be completed, including current trends and practices. These assessments are performed to ensure that, prior to study initiation, sites selected for a study are completely prepared to enroll patients. Although site assessments are not required, not assessing a site could be very costly—a site may have qualified patients and learn that they do not have the equipment and/or knowledge to enter patients correctly within an EDC system. Learning this too late would be detrimental to the site, study, and sponsor. Hardware and broadband provisioning can be undertaken by the sponsor, CRO, or can be outsourced to a 3rd party provider that specializes in this area.

End User Preparation

This section concerns activities that should be conducted before the study begins to ensure staff are prepared to use the EDC system.

Setting System Rights Determined by Roles and Privacy

Internet-based access to an EDC system presents additional challenges that must be clearly documented to ensure security and confidentiality of study data. The browser must ensure that all data connections cannot be breached or corrupted by an unauthorized user or external software. All audit trails detailing user access must remain unmodified and intact. User management begins with the sponsor's evaluation of the roles and responsibilities for each task within the system, based on criteria outlined in staff evaluation plans. Because the EDC system is used by different site staff and sponsor team members, access needs to be considered for all. Where input or review of data is required within the system, user roles and responsibilities should be defined and documented to identify specific access privileges or rights. Factors to be considered when defining these user roles include the following:

● Data entry rights by both site staff and sponsor team members. For example, dictionary coding requires that sponsor staff be able to enter or modify certain fields on a form. To ensure that the integrity and reliability of data is maintained, sponsors should carefully consider which fields will

be modifiable by the sponsor team. If sponsors will have such access, clear process documentation and a robust audit trail are also critical.

● Investigator signature rights

● Query generation—for example, in-house review by CDM versus reviews performed at the site by monitors

● Query resolution—for example, sites may only be able to resolve specific types of queries, while CDM can close queries after reviewing site responses

● SDV rights

● Read-only access—for example, blinding and patient privacy regulations may require user access to be limited to only certain CRFs

● Report creation, generation, or view-only access at both the site and by the sponsor should be considered. Some possible scenarios include limiting access so that each site can only generate reports for their subjects, limiting report generation across countries or regions, or limiting report creation to CDM staff who have received more advanced training.

User IDs and Passwords

Conventions for user IDs and passwords need to be determined up front in study planning. In addition, processes for dissemination of IDs and passwords to users must be established. These processes should include tracking that users have been properly trained prior to receiving access to the system. The system should force users to change their password at first log-in.

Training or system documentation should educate users as to the rules and regulations on keeping user ID and password information confidential, as well as requirements for changing their passwords. Lastly, the training materials should instruct users on what to do should they lose or forget their ID and/or password.

Account Management

Data management should participate in designing the account management process so they can train clinical staff on how they and their site coordinators

will obtain access to the EDC system. The process should minimize the number of manual steps that are included. Consideration should be given to linking the CTMS to the account creation and activation system, thereby eliminating the need to transfer user information between systems.

The typical account activation process is as follows:

1. A user is trained and authorized to be granted access to the system.

2. The user calls the help desk to request activation of his or her account.

3. The help desk confirms that EDC training has been completed by the user.

4. The help desk creates the account and assigns a temporary password to it.

5. The help desk guides the user through the process of logging on to the EDC system and selecting a new password.

6. The user confirms access to the EDC system.

Training Prior to System Access

If study team members and site staff are not fully trained to use the EDC system, they are unlikely to use it properly. Therefore, users must complete required training before being provided access to the EDC system. If possible, a certification exam can be included at the end of training to certify competence. Certification forms should be given to trainees as appropriate.

User training on both the system and study setup within the system is important. There are various views on the extent to which these two components should be included in the training plan. At a minimum, each user with the ability to modify study data should have documented training on basic system functionality, such as logging on, opening a CRF, entering data, and responding to a query. User training can be provided through methods such as the following:

● Self-study: reading materials, e-learning materials, using sample forms in a training environment

- Training environments that provide training exercises with examples that are generic or customized to the study-specific workflow

- Web-based instruction or demonstration

- Face-to-face training: Conduct training for users in a central training facility, such as investigators' meetings or other centralized training meetings

Recommended Standard Operating Procedures

- EDC Design Specifications

- System Setup, Installation and Support

- EDC Training

- Medical Coding

- Data Collection and Handling

- Data Backup, Recovery, and Contingency Plans

- Data Review and Validation

- Prequalification Requirements including 21 CFR Compliance

- User Access Creation, Modification and Revocation

- Systems and Hardware Security

- Guidelines for Outsourcing with Vendors/Vendor Management

- Handling External Data

- Coding Medical and Clinical Terms

References

1. International Conference on Harmonisation. *ICH Expert Working Group: Guideline for Good Clinical Practice.* Geneva, Switzerland: Author; 1996.

2. Code of Federal Regulations, Title 21, Part 11, Electronic Records; Electronic Signatures. Washington, DC: US Government Printing Office; 1997.

3. US Food and Drug Administration. *Guidance for Industry: Computerized Systems Used in Clinical Trials.* Washington, DC: US Department of Health and Human Services; 2007.

4. Khosla R, Verma DD, Kapur A, Khosla S. Efficient source data verification. *Indian Journal of Pharmacology.* 2000;32:180–186.

5. Schuyl, Malcolm L, Engel, Thim. A review of the source document verification process in clinical trials. *Drug Information Journal.* 1999;33:789–797.

6. CDISC (Clinical Data Interchange Standards Committee) web site. Available at: http://www.cdisc.org/. Accessed on March 31, 2008.

Further Reading

Code of Federal Regulation, Title 21, Volume 1, Part 11: Electronic Records, Electronic Signatures. Washington, DC: Government Printing Office; 1998.

Data Basics. 2001:Summer;8–9.

Drug Information Association. *Workshop on "Electronic Data Capture: Technologies, Implications and Future Requirements", November 1998.* Baltimore, MD: Author; 1998.

eClinical Forum PhRMA EDC/eSource Taskforce (Formerly the Electronic Data Management Forum (EDM Forum). *The Future Vision of Electronic Health Records as eSource for Clinical Research, Version 1.0, September 14, 2006*

Electronic records; electronic signatures. *Federal Register.* 1997:62;13429–13466.

Hopp DI. Three topics integral to the use of the Internet for clinical trials: Connectivity, communication, & security. *Drug Information Journal.* 1998:32; 933–939.

Hyde AW. The changing face of electronic data capture: From remote data entry to direct data capture. *Drug Information Journal*. 1998:32;1089–1092.

Hyland ME, Kenyon CAP, Allen R, et al. Diary keeping in asthma; Comparison of electronic and written methods. *British Medical Journal*. 1993:306;487–489.

IIR Symposia on "Automated Data Capture", February 1998, Orlando, FL and September 1998, Washington, DC.

Kubick WR. The elegant machine: applying technology to optimize clinical trials. *Drug Information Journal*. 1998:32;861–869.

Kush RD. Electronic data capture: A survey. *The Monitor*. 1999:Fall;37–40.

Lampe AJ, Weiler JM. Data capture from investigators' & sponsors' perspectives: Balancing quality, speed, & cost. *Drug Information Journal*. 1998:32;871–886.

Latham D. EDC: Establishing standards for electronic data capture. Interview of K. Vaillant. *Pharmaceutical Visions.* 1998:Autumn.

Ma, JM. *A Modeling Approach to System Evaluation in Research Data Management*. Unpublished doctoral dissertation. Chapel Hill, NC: University of North Carolina; (1986).

Marks, RG. Validating electronic source data in clinical trials. *Controlled Clinical Trials.* 2004;25:437–446.

Palm U. Controlling the increased complexity of electronic clinical data collection and management technology by data modeling and system engineering. *Drug Information Journal*. 2002:36;683–692.

Regalado A. Re-engineering drug development, II: Clinical data collection and management. *Start-up*. 1998:27.

Stokes T. Computer systems validation (6-part series). *Applied Clinical Trials*. September 1996, January 1997, February 1997, April 1997, June 1997, August 1997.

US Food and Drug Administration. *Guidance for Industry: Computerized Systems Used in Clinical Trials.* Washington, DC: US Department of Health and Human Services; 2007.

US Food and Drug Administration. *Guidance for Industry: Electronic Submissions of Case Report Forms (CRFs), Case Report Tabulations (CRTs) and Data to the Center for Biologics Evaluation and Research.* Washington, DC: US Department of Health and Human Services; 1998.

Chapter Revision History

Publication Date	Comments
September 2003	Initial publication.
May 2007	Revised for style, grammar, and clarity. Substance of chapter content unchanged.
September 2008	Revised to reflect the orientation of chapter towards the concept and start-up phase of EDC. Content updated and organization of material revised. Study conduct and study closeout content moved to separate chapters.

This page is intentionally blank.

.