GCDMP

REVIEW ARTICLE

Electronic Data Capture—Selecting an EDC System

Maxine Pestronk*, Derek Johnson†, Muthamma Muthanna‡, Olivia Montano§, Denise Redkar-Brown‖, Ralph Russo¶, Shweta Kerkar** and David Eade††

Web-based electronic data capture (EDC) has become the preferred method for capture of key-entered data in clinical studies. This chapter reviews the considerations for selecting an EDC system including evaluation of systems and vendors, user requirements, and process change, as well as initial implementation of systems within organizations. The goal of system selection is to assure that organizational needs are identified and documented and ultimately that that the desired functionality is available and appropriately supports clinical studies conducted by the organization. Multiple roles on study teams use the EDC system and should be involved in software selection and initial implementation.

1) Learning Objectives

After reading this chapter, the reader should understand

- The regulatory basis for practices in selecting an EDC system
- Common requirements and functionality domains of EDC systems
- Key domains and criteria for pre-selection evaluation of EDC systems
- Process impact and redesign considerations at evaluation and selection time
- Initial system implementation within an organization

2) Introduction

Historically, data for clinical trials have been manually abstracted from medical records, electronically extracted from medical records, or recorded directly on Case Report Forms.[1] For multicenter clinical studies a variety of approaches have been reported and have evolved toward decentralized entry of data at clinical sites.[2] Advantages of relocating data entry closer to the data source with data checks that flag discrepant data on the user interface during entry have been well articulated as have barriers and challenges.[2] The latter have evolved over time and with technology advances and adoption have mostly been

* P4CDM, US

† Inflamax Research, LTD, CA

‡ Accenture Service Pvt Ltd, IN

§ Pathways for Rare and Orphan Studies (PROS) Foundation, US

‖ Independent

¶ Pfizer, US

** IQVIA, US

†† UBC, CA

Corresponding author: Maxine Pestronk, CCDM
(mpestronk@verizon.net)

reduced to operational challenges overcome in most research contexts.

Reports of EDC predecessor systems and process, called Remote Data Entry (RDE) started appearing in the 1970s with the then increasing accessibility of computers.[3,4,5,6,7,8,9] Widespread access to the internet at the beginning of this century offered opportunity to centralize technology management while maintaining the ability for clinical sites to enter and respond to queries about data. Several informative historical accounts of the development of web-based EDC as a technology are available.[2,10,11] Most EDC systems today are fully web-based and accessed through web browsers.

Choosing an EDC system can and should be a significant decision for an organization. EDC system selection decisions are complex. Such decisions involve choices about processes for collecting and managing data that involve the entire trial team. Adopting new information technology offers new opportunity for process redesign such as workflow automation and other decision support. It has been argued that the real gains from EDC come with use of the technology to re-engineer processes.[10,12–19] Organizations must decide to what extent a new system needs to support existing processes for data collection, management, and monitoring versus offer new ways of working. EDC system selection and implementation usually involves many stakeholders including, but not limited to, project management, data management, clinical management, biostatistics, and information technology.

While sometimes EDC systems are chosen for an individual study, more frequently the chosen system will be used for multiple studies conducted by an organization and will impact operating procedures at clinical investigational sites and throughout clinical study operations. As the primary system for data collection and management in

clinical studies today, EDC systems can be a cornerstone of, integral component of, and leaping-off point toward greater safety, quality, and efficiency in clinical studies.

3) Scope

This first of three chapters on web-based Electronic Data Capture (hereafter EDC) covers considerations in and criteria and processes for selection of software for web-based EDC in clinical studies. Topics covered include common EDC system functionality, evaluation of candidate systems and vendors, and consideration of process impact and potential for process redesign at the time of system selection. The primary focus in this chapter is the identification of requirements for the EDC platform itself. These include functionality important to human subject protection, data integrity, study conduct – including needs of system users, and Title 21 Part 11 compliance. Other important aspects of choosing an EDC system such as software delivery, acquisition models, vendor stability, and cost estimation are addressed as are initial software implementation considerations.

Recommendations for building a study within an EDC system, testing a built study, study start-up, and provisions for change control for an EDC study are covered in the second EDC Chapter titled "Electronic Data Capture – Study Implementation and Start-up." Aspects of study conduct and study closeout are addressed in the third and final EDC Chapter entitled "Electronic Data Capture – Study Conduct, Maintenance, and Closeout". The EDC chapters contain applications of good clinical data management practices *specific* to data collection and management using web-based EDC. General good clinical data management practices are not re-articulated here.

4) Minimum Standards

The **International Council for Harmonisation (ICH) E6 (R2)** contains several passages particularly relevant to EDC software selection and initial implementation.[20]

Section 2.8 states, "Each individual involved in conducting a trial should be qualified by education, training, and experience to perform his or her respective tasks."[20] Echoing similar statements elsewhere in ICH Good Clinical Practice (GCP) and in Title 21 CFR Part 11, this requirement applies to EDC software selection in that it applies to individuals involved in EDC system selection, installation, testing, use, and maintenance whether they are performed in-house or elsewhere. Where tasks are performed by other organizations, this requirement is met through vendor qualification assessments, usually part of software selection decision making. Functionality to record and track over time system privileges assigned to users, i.e., tasks that users are allowed to perform in the system, becomes criteria used in software evaluation and selection.

Section 2.10 states, "All clinical trial information should be recorded, handled, and stored in a way that allows its accurate reporting, interpretation, and verification."[20] Functionality to meet this requirement becomes criteria used in software evaluation and selection.

Section 2.11 states, "The confidentiality of records that could identify subjects should be protected."[20]

Functionality to meet this requirement becomes criteria used in software evaluation and selection.

Section 4.9.0 states, "The investigator/institution should maintain adequate and accurate source documents and trial records that include all pertinent observations on each of the site's trial subjects. Source data should be attributable, legible, contemporaneous, original, accurate, and complete. Changes to source data should be traceable, should not obscure the original entry, and should be explained if necessary (e.g., via an audit trail)."[20] This investigator site requirement applies to EDC systems because the EDC system can serve as the original capture of information, in which case the EDC system is maintaining source data. Where the EDC system is intended to be used in this way, functionality to meet this requirement becomes criteria used in software evaluation and selection.

Section 4.9.2 states, "Data reported on the CRF, that are derived from source documents, should be consistent with the source documents or the discrepancies should be explained."[20] Functionality in EDC systems for triggering data to be source document-verified and tracking such verification supports this GCP requirement and becomes criteria used in software evaluation and selection.

Section 5.0 in the following passage recommends use of quality management systems and advocates risk management.

> "The sponsor should implement a system to manage quality throughout all stages of the trial process.
>
> Sponsors should focus on trial activities essential to ensuring human subject protection and the reliability of trial results. Quality management includes the design of efficient clinical trial protocols, tools, and procedures for data collection and processing, as well as the collection of information that is essential to decision making.
>
> The methods used to assure and control the quality of the trial should be proportionate to the risks inherent in the trial and the importance of the information collected. The sponsor should ensure that all aspects of the trial are operationally feasible and should avoid unnecessary complexity, procedures, and data collection. Protocols, case report forms (CRFs), and other operational documents should be clear, concise, and consistent.
>
> The quality management system should use a risk-based approach."[20]

A Quality Management System necessitates that executive leadership articulate and support a quality policy that documents leadership intent with respect to quality management. Because methods used to collect and manage data impact quality, executive leadership support for EDC selection and use is imperative. Further, leadership should assure that the quality management system extends throughout the organization and to vendors, suppliers, and sub-contractors where appropriate through a vendor qualification and management program.

Section 5.0.1 further advocates a process-oriented quality management system approach stating that "During protocol development the Sponsor should identify processes and data that are critical to ensure human subject protection and the reliability of trial results."[20] Processes reliant upon EDC software may meet this requirement. EDC functionality to indicate, process, and report on data deemed critical is supportive of meeting this requirement and becomes criteria used in software evaluation and selection.

Section 5.1.1 further states, "The sponsor is responsible for implementing and maintaining quality assurance and quality control systems with written SOPs to ensure that trials are conducted and data are generated, documented (recorded), and reported in compliance with the protocol, GCP, and the applicable regulatory requirement(s)."[20] Title 21 CFR Part 11 also requires SOPs for data collection, entry, and changes;[21] in the case of EDC, these apply directly to clinical investigational sites as well as data sponsors. Functionality supporting auditable processes becomes criteria used in software evaluation and selection.

Section 5.1.2 protects access to source data and documents stating, "The sponsor is responsible for securing agreement from all involved parties to ensure direct access (see section 1.21) to all trial-related sites, source data/documents, and reports for the purpose of monitoring and auditing by the sponsor, and inspection by domestic and foreign regulatory authorities."[20] Where the EDC system is used to collect and maintain source data, this criterion applies. EDC software functionality to support controlled and direct access to source data and documents supports this requirement and becomes criteria used in software evaluation and selection.

Section 5.1.3 states, "Quality control should be applied to each stage of data handling to ensure that all data are reliable and have been processed correctly."[20] EDC system functionality to support automated detection, alerting, and tracking resolution of discrepant data directly supports this requirement as does functionality to support source document verification and reconciliation of data captured through EDC with externally collected or managed data. As such, this functionality becomes criteria used in software evaluation and selection.

Section 5.5.1 refers to qualifications of study personnel and states, "The sponsor should utilize appropriately qualified individuals to supervise the overall conduct of the trial, to handle the data, to verify the data, to conduct the statistical analyses, and to prepare the trial reports."[20] This general requirement applicable to all data requires that personnel qualifications, including site users of EDC systems, with respect to the EDC software and its use be documented. The role of personnel qualifications in software selection decisions is described above in ICH E6R2 section 2.8.

Section 5.5.3 concerns validation of computerized systems and states, "When using electronic trial data handling and/or remote electronic trial data systems, the sponsor should, a) Ensure and document that the electronic data processing system(s) conforms to the sponsor's established requirements for completeness, accuracy, reliability, and consistent intended performance (i.e., validation)."[20] This requirement echoes Title 21 CRF Part 11 and requires that the installation of the EDC system used for a study be validated.

Section 5.5.3's first addendum states that validation of computer systems should be risk-based. "The sponsor should base their approach to validation of such systems on a risk assessment that takes into consideration the intended use of the system and the potential of the system to affect human subject protection and reliability of trial results."[20] This general GCP requirement promotes right-sizing the type and extent of validation of system functionality to the assessed risk associated with functionality. In EDC systems, building a study within validated software has significantly less risk than developing new software. Open source software has different risks than commercial software or in-house custom-developed software. These risk differences are considerations in EDC software selection and initial implementation including system validation.

Section 5.5.3 addendum b states that an organization "Maintains SOPs for using these systems."[20] The 5.5.3 addendum c-h introductory statement enumerates topics that should be covered in SOPs. "The SOPs should cover system setup, installation, and use. The SOPs should describe system validation and functionality testing, data collection and handling, system maintenance, system security measures, change control, data backup, recovery, contingency planning, and decommissioning."[20] These requirements apply to system selection and initial implementation in that the processes covered by the requirement can be significantly impacted by the functionality available in an EDC system being used by a sponsor. Further, individual requirements in the section such as 5.5.3 addendum (e) "Maintain a list of the individuals who are authorized to make data changes", (g) "Safeguard the blinding, if any", and (h) "Ensure the integrity of the data, including any data that describe the context, content, and structure"[20] enumerate functionality that become evaluation criteria to the extent EDC system support of these requirements is required by the organization.

Section 5.5.4 concerns traceability and states, "If data are transformed during processing, it should always be possible to compare the original data and observations with the processed data."[20] This requirement directly states functionality needed in EDC systems for GCP compliance.

Section 8.0 states that documents that "individually and collectively permit evaluation of the conduct of a trial and the quality of the data produced"[20] are considered essential documents and shall be maintained as controlled documents, i.e., "should provide for document identification, version history, search, and retrieval". This requirement impacts EDC software selection and implementation in that EDC functionality to manage and control data specifications, including definition and specifications for programmatic operations performed on data, eases the external document control burden and

includes the records maintained via this functionality as essential documents.

Title 21 CFR Part 11 also identifies regulatory requirements for traceability, training and qualification of personnel, and validation of computer systems used in clinical trials.[21] Requirements in 21 CFR Part 11 Subpart B are stated as controls for closed systems (21 CFR Part 11 Sec. 11.10), controls for open systems (21 CFR Part 11 Sec. 11.30), signature manifestations (21 CFR Part 11 Sec. 11.50), and signature/record linking (21 CFR Part 11 Sec. 11.70). Requirements for electronic signatures are stated in in 21 CFR Part 11 Subpart C. The requirements in Title 21 CFR Part 11 directly impact EDC software selection and initial implementation. Where Part 11 compliance is required, the technical controls stipulated become software evaluation and selection criteria. Where the EDC system needs to be Title 21 CFR Part 11 compliant, which in the United States includes all studies submitted to the FDA for regulatory review as well as many studies funded by the National Institutes of Health (NIH), the Part 11 technical controls become software evaluation and selection criteria. In addition, a common interpretation of these regulations is that important functionality be documented on a traceability matrix that explicitly documents the functions and ways in which they were tested.[21]

The March 2018 **FDA Study Data Technical Conformance Guide Technical Specifications** Document is incorporated by reference into the *Guidance for Industry Providing Regulatory Submissions in Electronic Format – Standardized Study Data.* The appendix of the Study Data Technical Conformance Guide states, "In addition to standardizing the data and metadata, it is important to capture and represent relationships (also called associations) between data elements in a standard way".[22] As such, documenting associations between data elements becomes an EDC software selection criterion.

The **Medicines and Healthcare products Regulatory Agency (MHRA) 'GXP' Data Integrity Guidance and Definitions** provides considerations and regulatory interpretation of requirements for data integrity, such as:

Section 5.1 "Systems and processes should be designed in a way that facilitates compliance with the principles of data integrity."[23] In E6(R2) the FDA defines data integrity as, "completeness, consistency, and accuracy of data" and goes on to state that "Complete, consistent, and accurate data should be attributable, legible, contemporaneously recorded, original or a true copy, and accurate (ALCOA)."[20] Section 6.4 of the MHRA guidance similarly defines data integrity as, "the degree to which data are complete, consistent, accurate, trustworthy, reliable and that these characteristics of the data are maintained throughout the data life cycle. The data should be collected and maintained in a secure manner, so that they are attributable, legible, contemporaneously recorded, original (or a true copy) and accurate".[23] Functionality to assure data integrity becomes EDC software evaluation criteria.

Section 6.9 of the MHRA guidance states, "There should be adequate traceability of any user-defined parameters used within data processing activities to the raw data, including attribution to who performed the activity."[23]

Functionality to assure data traceability becomes EDC software evaluation criteria.

The **General Principles of Software Validation; Final Guidance for Industry and FDA Staff** points out relevant guidelines regarding documentation expected of software utilized in a clinical trial. This guidance discusses software validation processes rather than functionality; these are not enumerated here. For commercial systems, this criterion can be met through a Quality Management System assessment during vendor assessment. For in-house developed EDC software, this criterion is met through an internal Quality Management System that covers the software development lifecycle.[24]

Good Manufacturing Practice Medicinal Products for Human and Veterinary Use (Volume 4, Annex 11): Computerised Systems provides the following guidelines when using computerized systems in clinical trials:

Section 1.0 echoes ICH E6(R2) stating, "Risk management should be applied throughout the lifecycle of the computerised system taking into account patient safety, data integrity and product quality. As part of a risk management system, decisions on the extent of validation and data integrity controls should be based on a justified and documented risk assessment of the computerised system." This impacts EDC system implementation in the level of testing and controls applied and, like ICH E6(R2), implies that a documented risk assessment should exist for the system.[20,25]

Section 4.2 states, "Validation documentation should include change control records (if applicable) and reports on any deviations observed during the validation process."[25] Functionality to track changes to the system set-up for a study becomes EDC software evaluation criteria.

Section 4.5 states, "The regulated user should take all reasonable steps, to ensure that the system has been developed in accordance with an appropriate quality management system."[25] For commercial systems, this criterion can be met through a Quality Management System assessment during vendor assessment. For in-house developed EDC software, this criterion is met through an internal Quality Management System that covers the software development lifecycle.

Section 7.1 states, "Data should be secured by both physical and electronic means against damage. Stored data should be checked for accessibility, readability and accuracy. Access to data should be ensured throughout the retention period."[25] The technical controls for data access become software evaluation and selection criteria. Readability, accuracy, and other data quality dimensions may be impacted by incorrect system operation. Thus, functionality enabling data accessibility, readability, and accuracy assessment during a study becomes EDC software evaluation criteria.

Section 7.2 states, "Regular back-ups of all relevant data should be done. Integrity and accuracy of backup data and the ability to restore the data should be checked during validation and monitored periodically."[25] These functions become software evaluation and selection criteria.

Section 10.0 states, "Any changes to a computerised system including system configurations should only be

made in a controlled manner in accordance with a defined procedure."[25] For commercial systems, this criterion can be met through a Quality Management System assessment during vendor assessment. For in-house developed EDC software, this criterion is met through an internal Quality Management System that covers the software development lifecycle.

GAMP 5: A Risk-based Approach to Compliant GxP Computerized Systems pre-dated ICH E6(R2) in suggesting scaling activities related to computerized systems with a focus on patient safety, product quality, and data integrity. GAMP® 5 provides guidance for maintaining compliant computerized systems fit for intended use. While GAMP® 5 does not articulate additional functional requirements that impact EDC software selection, the evaluation of vendors, open source products, and development and evaluation of in-house software can all be informed by the approaches in GAMP® 5. GAMP® 5 provides the following guidelines relevant to systems used to collect and process clinical trial data:

Section 2.1.1 echoes the risk-based approach articulated in ICH E6(R2) and states, "Efforts to ensure fitness for intended use should focus on those aspects that are critical to patient safety, product quality, and data integrity. These critical aspects should be identified, specified, and verified."[26] Like ICH E6(R2) this promotes right-sizing the type and extent of validation of system functionality to the assessed risk associated with functionality.

Section 4.2 states, "The rigor of traceability activities and the extent of documentation should be based on risk, complexity, and novelty; for example, a non-configured product may require traceability only between requirements and testing."[26] Like ICH E6(R2) this promotes right-sizing the type and extent of traceability-related system functionality to the assessed risk associated with functionality.

Section 4.2 states, "The documentation or process used to achieve traceability should be documented and approved during the planning stage, and should be an integrated part of the complete life cycle."[26] As such, these functions become software evaluation and selection criteria.

Section 4.3.4.1 "Change management is a critical activity that is fundamental to maintaining the compliant status of systems and processes. All changes that are proposed during the operational phase of a computerized system, whether related to software (including middleware), hardware, infrastructure, or use of the system, should be subject to a formal change control process (see Appendix 07 for guidance on replacements). This process should ensure that proposed changes are appropriately reviewed to assess impact and risk of implementing the change. The process should ensure that changes are suitably evaluated, authorized, documented, tested, and approved before implementation, and subsequently closed."[26] Functionality to track changes to the system set-up for a study becomes EDC software evaluation criteria.

Section 4.3.6.1 states, "Processes and procedures should be established to ensure that backup copies of software, records, and data are made, maintained, and retained for a defined period within safe and secure areas."[26] These functions become software evaluation and selection criteria.

Section 4.3.6.2 states, "Critical business processes and systems supporting these processes should be identified and the risks to each assessed. Plans should be established and exercised to ensure the timely and effective resumption of these critical business processes and systems."[26] Business resumption becomes an evaluation and selection criterion for EDC vendors and an implementation consideration for internal and vended solutions.

Sections 5.3.1.1 and 5.3.1.2 inform organizational risk assessment procedures and respectively state, "The initial risk assessment should include a decision on whether the system is GxP regulated (i.e., a GxP assessment). If so, the specific regulations should be listed, and to which parts of the system they are applicable. For similar systems, and to avoid unnecessary work, it may be appropriate to base the GxP assessment on the results of a previous assessment, provided the regulated company has an appropriate established procedure." and that "The initial risk assessment should determine the overall impact that the computerized system may have on patient safety, product quality, and data integrity due to its role within the business processes. This should take into account both the complexity of the process and the complexity, novelty, and use of the system."[26]

The FDA guidance, **Use of Electronic Health Record (EHR) Data in Clinical Investigations**, emphasizes that data sources should be documented and that source data and documents be retained in compliance with 21 CFR 312.62(c) and 812.140(d). It further states, "FDA's acceptance of data from clinical investigations for decision-making purposes depends on FDA's ability to verify the quality and integrity of the data during FDA inspections."[27]

Section IV defines interoperability as, "the ability of two or more products, technologies, or systems to exchange information and to use the information that has been exchanged without special effort on the part of the user"[27] and recognizes that "EHR and EDC systems may be non-interoperable, interoperable, or fully integrated, depending on supportive technologies and standards."[27] Where integration or interoperability is desired by the sponsor, these become EDC software selection requirements. Such requirements may include system support for data interchange according to specific standards such as the Health Level Seven (HL7) Fast Healthcare Interoperability Resource (FHIR®) standards or the Clinical Data Interchange Standards Consortium (CDISC) Operational Data Model (ODM) standard.

Section IV.C states, "FDA encourages sponsors to periodically check a subset of the extracted [from EHRs] data for accuracy, consistency, and completeness with the EHR source data and make appropriate changes to the interoperable system when problems with the automated data transfer are identified."[27] Functionality in EDC systems accepting electronic EHR data to enter re-abstracted data and itemize and track resolution of discrepancies would support this guidance recommendation. Where desired,

such functionality becomes an EDC software selection criterion.

Section V.C.1 echoes the eSource guidance and states that the EDC system should have the ability to identify the EHR as "the data originator for EHR data elements gathered during the course of a clinical investigation"[27] Where EHR interoperability is desired, this functionality becomes an EDC software selection criterion.

Section V.C.2 echoes the eSource guidance and states, "After data are transmitted to the eCRF, the clinical investigator or delegated study personnel should be the only individuals authorized to make modifications or corrections to the data."[27] The section echoes Title 21 CFR Part 11 and further states, "Modified and corrected data elements should have data element identifiers that reflect the date, time, data originator, and the reason for the change" and that "Modified and corrected data should not obscure previous entries".[27] The same section further states, "Clinical investigators should review and electronically sign the completed eCRF for each study participant before data are archived or submitted to FDA", that "If modifications are made to the eCRF after the clinical investigator has already signed the eCRF, the changes should be reviewed and approved by the clinical investigator", and that use of electronic signatures for records subject to Title 21 CFR Part 11 must comply with that regulation.[27] Where EHR interoperability is desired, this functionality becomes an EDC software selection criterion.

Further echoing the eSource guidance Section V.C.2 states, "If a potential for unblinding is identified, sponsors should determine whether the use of interoperable systems is appropriate or whether other appropriate controls should be in place to prevent unblinding."[27] Where EHR interoperability is desired, this functionality becomes an EDC software evaluation and selection criterion.

Similarly, the FDA's **Guidance on Electronic Source Data Used in Clinical Investigations** provides guidance on the capture, review, and retention of electronic source data in FDA-regulated clinical investigations. The guidance "promotes capturing source data in electronic form, and it is intended to assist in ensuring the reliability, quality, integrity, and traceability of data from electronic source to electronic regulatory submission".[28]

In the background section, the guidance states, "Source data should be attributable, legible, contemporaneous, original, and accurate (ALCOA) and must meet the regulatory requirements for recordkeeping."[28] Record keeping requirements for clinical investigators and sponsors are detailed in Title 21 CFR 312.50, 312.58, 312.62, and 312.68 for drugs and biologics and Title 21 CFR 812.140 and 812.145 for medical devices.

Section III.A.1 states that all data sources (called originators in the guidance) at each site should be identified. "A list of all authorized data originators (i.e., persons, systems, devices, and instruments) should be developed and maintained by the sponsor and made available at each clinical site."[28] The guidance states that each data element is associated with a data originator. It goes on to state, "When a system, device, or instrument automatically populates a data element field in the eCRF, a data element identifier should be created that automatically identifies the particular system, device, or instrument (e.g., name and type) as the originator of the data element."[28] For a Title 21 CFR Part 11 audit trail, the association with the data originator and data changer is required at the data value level. Where the EDC system is the source, the EDC system should be on this list. In addition, EDC system functionality, where desired, for maintaining the data source list for each site, becomes a requirement in software evaluation and selection decisions.

Section III.A.1 of the guidance requires controls for system access and states, "When identification of data originators relies on identification (log-on) codes and unique passwords, controls must be employed to ensure the security and integrity of the authorized user names and passwords. When electronic thumbprints or other biometric identifiers are used in place of an electronic log-on/password, controls should be designed to ensure that they cannot be used by anyone other than their original owner."[28] This functionality becomes an EDC software selection criterion.

Section III.A.2.d states that when data from an EHR are transmitted directly into the eCRF, i.e., electronically, the EHR is considered the source.[28] The stated rationale is that algorithms are often needed to select the intended data value and that this processing step necessitates verification. As such and where EHR-to-eCRF eSource data will be used, the ability to designate the EHR as the source for data values originating from the EHR becomes a requirement in EDC software evaluation and selection.

Section III.A.3 states, "Data element identifiers should be attached to each data element as it is entered or transmitted by the originator into the eCRF"[28] and that data element identifiers should contain (1) Originators of the data element, (2) Date and time the data element was entered into the eCRF (this data receipt milestone is also time point at which the EDC system audit trail begins), and (3) association with the subject to which the data belongs.[28] The guidance further states that the EDC system "should include a functionality that enables FDA to reveal or access the identifiers related to each data element".[28] As such and where eSource data will be used, the ability to maintain these associations becomes a requirement in EDC software evaluation and selection.

Section III.A.4 states, "Only a clinical investigator(s) or delegated clinical study staff should perform modifications or corrections to eCRF data".[28] Echoing Title 21 CRF Part 11, the section goes on to state, "Modified and/or corrected data elements must have data element identifiers that reflect the date, time, originator and reason for the change, and must not obscure previous entries", that "A field should be provided allowing originators to describe the reason for the change", and that "Automatic transmissions should have traceability and controls via the audit trail to reflect the reason for the change."[28] Where the EDC system is used to capture source data, these items become EDC software evaluation and selection requirements.

Section III.A.5 states that the FDA encourages "use of electronic prompts, flags, and data quality checks in the eCRF to minimize errors and omissions during data entry".[28] The rationale is that for eSource data, without an independent recording of the observation, the opportunity to make corrections to the source is gone after the time of the original observation or measurement has passed. For this reason and where the EDC system is used as the original capture of source data, this becomes a requirement for EDC software evaluation and selection.

The same section states, "clinical investigator(s) should have the ability to enter comments about issues associated with the data".[28] Where the EDC system is used as the original capture of source data, this becomes a requirement for EDC software evaluation and selection.

Section III.B.1.a states that to comply with the requirement in 21 CFR 312.62(b) for drugs and biologics and 812.140(a)(3) for devices to maintain accurate case histories, "clinical investigator(s) should review and electronically sign the completed eCRF for each subject before the data are archived or submitted to FDA" and that such electronic signatures must comply with Title 21 CFR Part 11.[28] This requirement applies more broadly than just where the EDC system is used as the original capture of source data and as such is routinely a requirement in EDC software evaluation and selection.

Section III.B.1.b goes on to state that in the case where clinical investigators need to be blinded to certain data, the data are exempt from the aforementioned investigator review requirement.[28] Where the EDC system is used to capture eSource and blinding of investigators to data is intended, the functionality to do so becomes a requirement in EDC software evaluation and selection.

Section III.B.2 anticipates the eventuality that data changes may be needed after a clinical investigator's review and signature. The guidance states that in this case, "the changes should be reviewed and electronically signed by the clinical investigator(s)".[28] This requirement applies more broadly than just where the EDC system is used as the original capture of source data and as such is routinely a requirement in EDC software evaluation and selection.

Section III.C states, "clinical investigator(s) should retain control of the records (i.e., completed and signed eCRF or certified copy of the eCRF)", that "clinical investigator(s) should provide FDA inspectors with access to the records that serve as the electronic source data", and that data and documents to corroborate source data captured in the EDC system may be requested during an inspection.[28] As such, where the EDC system is to be used as the original capture of source data, provision of a certified copy of all such data and relevant context such as the audit trail, data element identifiers, and data originators to the clinical investigator becomes a requirement in EDC software evaluation and selection.

Section III.D emphasizes that the FDA encourages viewing of the data early and by sponsors, CROs, data safety monitoring boards, and other authorized personnel to prompt detection of study-related problems. While prompt data review is not itself a regulatory requirement, adequate clinical trial monitoring is; i.e., "… ensuring proper monitoring of the investigation(s), ensuring that the investigation(s) is conducted in accordance with the general investigational plan and protocols contained in the IND…".[29] As such, functionality to support, facilitate, or automate timely review of data become criteria in EDC software evaluation and selection.

Section III.D also suggests aspects of access control: (1) a list of the individuals with authorized access to the eCRF should be maintained, (2) only those individuals who have documented training and authorization should have access to the eCRF data, (3) Individuals with authorized access should be assigned their own identification (log-on) codes and passwords, and (4) log-on access should be disabled if the individual discontinues involvement during the study.[28] These access control requirements echo Title 21 CFR Part 11 and apply more broadly than just where the EDC system is used as the original capture of source data. As such these are common requirements in EDC software evaluation and selection.

With these requirements in mind, in **Table 1** we state the following minimum standards for the selection of EDC systems.

Table 1: Minimum Standards.

---

1. Secure upper management support for the selection of an EDC system and the functionality to be sought in the selection.

2. Identify system requirements to be used as software evaluation and selection criteria including but not limited to functionality needed for human subject protection, data integrity, study conduct, and regulatory compliance.

3. The identified functionality should be comprehensive for the organization and serve as the starting point of the traceability matrix used in software validation.

4. If a commercial product is to be used, identify vendor selection criterion; these include functionality, business, Quality Management System, or other characteristics to be assessed as part of vendor qualification.

5. Perform and document a risk assessment with respect to the EDC system as intended to be used.

6. Undertake and complete software validation activities and other controls commensurate with the risk.

7. Assess and complete changes to Standard Operating Procedures (SOPs) and other governance policy and procedures necessitated by EDC-enabled processes.

8. Create and maintain SOPs covering the software evaluation, selection, and implementation processes.

---

5) Best Practices

Best practices, as stated in **Table 2**, were identified by both the review and the writing group. Best practices do not have a strong requirement based in regulation or recommended approach based in guidance, but do have supporting evidence either from the literature or consensus of the writing group. As such best practices, like all assertions in GCDMP chapters, have a literature citation where available and are always tagged with a roman numeral indicating the strength of evidence supporting the recommendation. Levels of Evidence are outlined in **Table 3**.

6) Stakeholders in the EDC System Selection Process

Choosing an EDC system involves multiple considerations including strategy, goals, business relationships, and capability of the selecting organization and their desired provisioning models, software functionality, system performance, and available funds. Thus, the list of stakeholders in EDC selection and implementation can be extensive. Operationally, data collection and management decisions such as whether EDC should be used and, if so, for which data and through what types of processes can impact most operational groups involved in clinical trials. Impacted operational groups include, but are not limited to, project management, data management, clinical operations, research pharmacy, biostatistics, information technology, and contracting. Individuals from all of these functional areas are possible stakeholders in the EDC selection and implementation process. [VI] The likelihood that all stakeholder needs will be reflected in the selection of an EDC system increases when those stakeholders and their responsibilities are identified early in the decision making process. [VI]

7) EDC System Selection Considerations

There are over 60 EDC system vendors in the EDC market.[32] Companies enter and leave the marketplace often. Though functionality is the most important category of selection criteria with respect to data quality, other considerations such as regulatory compliance, cost, vendor stability, business model compatibility, system flexibility, implementation timing, global experience, and availability of support may narrow the number of appropriate vendors. Kush et al. categorize EDC system selection criteria as aspects of vendor background or product specifications. Product specifications are further categorized as regulatory, standards, usability, operational, and business.[10] A sample Request for Proposals (RFP) including (1) vendor, software, and service characteristics; (2) business relationship information including the licensing model, service, maintenance, and other agreement terms; (3) pricing details; and (4) functional requirements. The Appendix illustrates the information one company found useful in EDC software and vendor selection and the functional requirements necessary for the conduct of their studies. Most organizations use a combination of functional and other criteria. Functionality can be a key differentiating factor among different EDC systems. However other factors such as vendor experience, past performance, financial stability, and pricing can significantly sway if not determine selection decisions. In this case, other business requirements may become a differentiator. Perhaps a lower priced system will be

Table 2: Best Practices.

1. Identify appropriate personnel participation in the evaluation and selection of EDC systems. [VI]

2. Ensure that internal or external skills are a match for the chosen EDC software acquisition model and are present to support the system and studies for which the system is used. [VI]

3. Separate infrastructure development from the set-up and conduct of individual studies.[13] [VII]

4. Identify and leverage EDC functionality to improve data collection and management; e.g., decision support, process automation, and interoperability with other information systems.[10,30] [III]

5. Where software is acquired as a service or with services, prepare a thorough Scope of Work (SOW) to document and ensure complete understanding of expectations and deliverables.[31] [III]

Table 3: Grading Criteria.

Evidence Level	Evidence Grading Criteria
I	Large controlled experiments, meta, or pooled analysis of controlled experiments, regulation or regulatory guidance
II	Small controlled experiments with unclear results
III	Reviews or synthesis of the empirical literature
IV	Observational studies with a comparison group
V	Observational studies including demonstration projects and case studies with no control
VI	Consensus of the writing group including GCDMP Executive Committee and public comment process
VII	Opinion papers

adequate for the simple, short-term studies, whereas a higher end system may be best for a complicated long-term study.

a) Business and Financial Factors

The cost of licensing an EDC system ranges from free to hundreds of thousands or millions of dollars. Pricing models vary among vendor organizations and historically have included charging by data value, data element, or CRF screen while others charge by study or even implementation. Other factors in pricing may include number of simultaneous or named user licenses and the services provided by the vendor. In addition, consideration should be given to service level agreements for availability of platform, system response time, and accessibility of support. The purchasing organization may prefer making an upfront one-time investment versus paying by study or incurring monthly fees. Organizational vendor contracting processes may not allow for the type of service contract offered by the vendor. Organizational processes may only authorize paying invoices for services that have been provided already, but a vendor may require a quarter payment upfront. For some organizations budget constraints may be a large factor in a decision.

b) Timeline Factors

The implementation timelines will also drive the selection process. The timing for the installation and validation of the EDC system must be considered to ensure that the system is ready to start a study build when needed and in full compliance with applicable regulations. Additional timing considerations include training, length of time to build a study specific application, and the extent of changes to organizational processes and SOPs required.

c) Vendor Background and Stability Factors

If the vendor is a public company, their financial and historical performance can be obtained freely. Knowing about the vendor is essential; information such as the following may provide insights:

- experience with the type or types of studies of interest
- experience in one or more therapeutic areas of interest
- experience in one or more regions of the world
- number of current/past customers
- software development experience
- aspects of the vendors software development quality management system
- number of employees
- ratio of development personnel to entire staff
- length of time in business
- financial stability
- past performance on similar volume and scope of functionality or services

Obtaining recent customer references is also strongly advised.

d) Process Compatibility Factors

Current SOPs may dictate a specific process that the system may or may not support. For example, organizational SOPs may allow a data manager to review and close queries, regardless of how they were generated, whereas a system may only allow the "monitor" role to close queries generated during the source document verification process. Such a difference would require a work-around or change to organizational process. There are many opportunities for role-based functionality like that employed in most EDC systems to conflict with current organizational practices, roles, and workflow. These conflicts necessitate changes in organizational practices, generation of work-arounds, or changes in software functionality prior to implementation. These organizational and business considerations may become selection criteria.

e) Software Functionality-based Selection Criteria

There is a core set of common EDC functions such as entering data and identifying data discrepancies covered by most EDC systems. See early functionality lists by McFadden et al., Kush et al., El Emam et al., and Franklin et al.[10,33,34,35] [VI] However, vendors differ in how basic EDC functionality is offered as well as continue to use current and planned functionality to differentiate their system from other marketed systems. An initial step in selecting an EDC system is determining organizational functionality requirements. Some vendors attempt to cover a vast array of requirements through system configurability such as multiple options for handling workflow and data flow for diverse data streams. High levels of configurability may increase the complexity of setting up, maintaining, and migrating studies as well as software cost. The effort necessary to achieve a function within a system may be just as important as whether a system can accomplish it. Thus, in comparing overall cost, configuration effort (i.e., study set-up effort) should be included as well as the cost for achieving needed features that are not covered by a system. In many cases a role-based workflow analysis, i.e., who does what and when, and data flow analysis will be helpful to determine configuration and implementation complexity and cost.

Common EDC product features are described but are not limited to those listed below.

Support for a hybrid data entry model: Some study scenarios include collection of data on paper. Several EDC systems have the capability of utilizing paper data entry into the same EDC database as eCRF data. These are called hybrid systems and allow for "Paper CRF" data entry (i.e., single or double data entry) by sponsor or designated personnel into the EDC system. Many EDC systems have the ability to set up two different types of form entry within the same build, e.g., double data entry for some users and single entry by others where entry rights are set by user permissions. Historically, the ability to enter data while not connected to the internet was offered with some systems. While not

a large concern today in urban areas, offline capability may be important in remote regions of the world.

Identification and resolution of data discrepancies: Most EDC systems have the functionality to specify and execute missing and range checks during entry. Systems vary in the support for tracking discrepancies and their resolution, management of more complex checks, and management of checks against imported data. The timing of when the checks are executed, i.e., in real-time as data are entered, after forms are submitted, nightly automatic execution, or manually initiated batch execution differs between vendors. The timing may also vary within an EDC system dependent on the type of check, e.g., univariate checks, multivariate checks, cross-form checks, and checks run against imported or interfaced data. Likewise, how discrepancies are tracked may vary by timing of execution or type of check.

Medical Coding: Medical coding features may include the ability to encode data during entry, facilitate encoding by a medical coder after entry, or facilitate exporting data for external coding. Some EDC systems include facilities for dictionary management and versioning.

Safety Data Management: Some organizations use separate safety data management systems and others manage the entire process within the EDC system. Where an organization uses a separate system for Serious Adverse Event management, the ability of the EDC system to send information about AEs meeting the criterion of "serious" may offer efficiency and data quality gains. Where the EDC system is used to manage follow-up and reporting of SAEs, rule-based detection, notification, and workflow management for AEs and potential SAEs is often desired as is functionality to export populated MEDWATCH or CIOMS SAE-forms to Sponsor systems or for external reporting.

Principle Investigator (PI) Signature functionality: Some systems have methods to allow the PI to sign forms, visits, or casebooks in a controlled fashion. For example if data is changed after the PI has signed, the signature will be revoked.

Importation of external data: Study data often come from external organizations and devices such as laboratory data, Patient Reported Outcome (PRO/ePRO) data, and data collected from external devices. These data may need to be imported and at minimum associated with the correct study subject and time point. EDC systems vary in the effort required to import data and in the functionality to manage imported data.

Integration with other systems: Real-time exchange and use of data is required for some studies. Some EDC systems include configurable application interfaces for the real-time exchange of data. Systems with which EDC systems interface may include clinical trial management systems (CTMS), randomization systems, systems at central labs, PRO/ePRO, and other systems.

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 clinical 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. Site status in the CTMS system, such as a site having clearance to enroll subjects, may be used as input into an algorithm to initiate access in the EDC system. 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.

Randomization: IVRS/IWRS (IRT System) Integrations: Randomization features may include randomization algorithms within the EDC system that automatically assign the treatment group, importing randomization lists, or interoperability with an external system used for randomization. Randomization functionality may support simple random sampling but may not support more complicated sampling strategies that include extensive blocking or balancing.

Interactive Voice Response Systems (IVRS), Interactive Web Response Systems (IWRS), or Interactive Response Technology (IRT) systems may be used for randomization in trials. Some EDC systems have built-in randomization functionality that is fully integrated as part of the baseline configuration Other systems do not have this feature or do not support the type of randomization chosen for a study necessitating use of an external randomization system. In either case, the combination of this functionality integrated with the EDC provides a powerful solution that reduces data entry for site staff and ensures no transcription errors in research subject identifiers.

Integration of this data from external vendors involves building a secured pathway via a secure File Transfer Protocol (sFTP), Web Services (SOAP/REST), or other secure mechanism that will transfer the files generated from the vendor system and place these files in the EDC compatible host. This process is sometimes referred to as developing an IVR/IWRS program.

ePRO/eCOA Integration: If patient reported outcomes or site-scored clinical outcome assessments will be collected via the Web, an e-diary device, or other data device, clinical data managers should consider where and how this data will be integrated with eCRF

data captured through the EDC system. If data collected using ePRO/eCOA is needed for decision-making by the site, it may be necessary to upload all or some, e.g., scores for scored assessments or ePRO data into the EDC system. More ePRO and eCOA systems are now integrating data directly with EDC systems, as the efficiencies gained in managing use have helped reduce overall study costs.

Laboratory Data Integration: It is sometimes helpful to have all or key laboratory parameters available to site staff within the EDC system even if central laboratories are used. The clinical data manager should consider this need with the clinical team. Integrated laboratory data stored in the EDC system can facilitate more timely and robust edit checks across other eCRFs.

Other External Data Integration: If electrocardiogram, medical device, or other data are collected from external vendors, the clinical data manager should evaluate whether data integration or data acquisition status is appropriate. Again, integrations into an EDC system should only be performed if the data has direct impact on subject management.

Clinical Data Management System (CDMS) Integration: At some point in a study, data are integrated into one database. In some organizations this is done during a study to facilitate data cleaning and a Clinical Data Management System may be used. Unless a fully integrated EDC or clinical data management solution is used, clinical data managers should consider how an EDC system will integrate with new or existing clinical 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. Integration should encompass data and queries, while avoiding manual transcription of queries into the CDMS when automated edit checks occur in the EDC system. Understanding general referential integrity and form design differences will assist the clinical data manager with this decision.

The point at which data are integrated should also be informed by reporting needs. Data from EDC, ePRO, vendors, or other sources often need to be viewed together to assess data status, completeness, or payment milestones. Where such data integration is not required for study operations, data from disparate sources may be merged prior to analysis.

Electronic Health Records: As the world is moving towards digitalization, data collection is evolving toward electronic data sources such as devices or EHRs. While building an EDC application it is important that consideration is given to the current and planned capability of the EDC system to receive data using healthcare-based exchange standards such as the Health Level Seven (HL7) Fast Healthcare Interoperability Resources (FHIR®) standards. The system may be used purely for transfer of data from an Investigator controlled source to EDC. It is important that the EDC system be able to capture the metadata from the source to maintain traceability.[27] [I]

Other Important Integrations: As new technological tools are developed, 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, clinical data managers should be aware of the need to also integrate an EDC system with reporting tools other than access to data by SAS® software or R for analysis.

Different types of data: Some data such as images or local labs bring additional requirements into consideration. For example local lab ranges are lab-specific, may change over the course of a trial, and are collected from each site. Images may require special processing and functionality to display and annotate. Data from some devices are acquired over time and the sampling rate needed for the study should be considered; data acquired at high sampling rates may be better visualized as waveforms. Special provisions may be needed to affix the identifier for the research subject or participant to the data.

Services related to the data: Some vendors or CROs offer data management services related to the study build, data management, and review of the study. Accessibility to data by and from these should be considered along with data integration decisions.

Accessibility of data: Data within the EDC system need to be accessible for analysis. EDC systems have facilities for export or accessibility of data. Vendors providing hosting or data management services charge for each data transfer. Other systems offer self-service exports or direct access, e.g., read only views of the data. In addition, functionality is required either from the EDC system or from a warehouse for authorized users to query the data directly. Situations often arise during a study that require ad hoc querying of the data to identify, confirm, or troubleshoot system or process problems. This functionality should always be available so that situations can be handled quickly when they arise.

Reports: EDC systems often offer reports or report development and delivery functionality. Some systems offer standard operational reports, e.g., outstanding forms or queries that cannot be customized for a specific study while other systems may offer varying degrees of report customization. Some systems offer facilities for visualization or ad hoc reporting of the clinical data. A third party reporting tool or custom programming may be needed where system reporting functionality does not support organizational needs.

Tools for study building and writing data validation checks: Some EDC systems include a graphical user interface for setting up eCRF screens and data vali-

dation checks while others import information used in set-up. Still others require programming in various languages and to varying extents. Some EDC systems use a combination of these. The workflow for study set-up can indicate the level of effort required. The latter is a common requirement in EDC software evaluation and selection.

Tools for documenting and managing user accounts: Some systems track user training and have technical controls to enforce completion of training prior to system access while others do not. Other EDC systems additionally offer facilities for granting and revoking privileges and tracking these changes over time.

Tools for managing risk based monitoring or partial Source Data Validation (SDV): Some systems allow for and have features to help manage SDV. Common features include having the system apply random sampling to identify a study-specific percentage of pages to be verified and having the system indicate variables that selected for SDV. Some systems support workflow for SDV as well as tracking SDV progress and results.

Tools for multi-lingual forms: While data in many studies are collected in English across the globe, that is not always the case. Organizations may have a need for rendering study data collection forms in different languages and for maintaining synchronicity between form and rule versions across forms rendered in multiple languages. Where needed, these become a requirement in EDC software evaluation and selection.

Data export functionality: Getting data out of EDC systems is as important as putting it in. Functionality is usually needed to provide the site a pdf or other human readable enduring copy of the data entered from their site at the close of a study. EDC system usually have functionality to provide customized and on-demand data exports in varying formats including CSV, SAS data sets, and xml including CDISC ODM. Many EDC system also support CDISC certified imports and exports of both data and metadata in bulk or "snapshot" or in transactional manner.

When considering features and functionality requirements, data managers should be aware that most vendors are in the continual process of improving/updating their system functionality. Therefore it may be important to understand features and their development timelines, such as whether the feature is a standard feature of the product in current release; will be a standard feature of the product in future release within the next 12 months; the feature needs a specific modification to the software; or the feature will not supported within the next 12 months. These are included in the sample Request for Proposals in the Appendix.

A good starting point for feature understanding is to think about needed system features in the context of a specific or typical project. The high-level business needs and desired system features form the requirements used in EDC system selection process. Although much older than the list above, McFadden et al. provide a pragmatic, basic list of needed functionality.[33]

Other specialized needs may include functions such as the following:

· Document management such as providing a webpage for making current versions of study documents available to sites
· Automatic reporting, sending out populated forms (e.g., CIOMS/SAE) to external Pharmacovigilance (PV) systems
· Collection and management of reference ranges by site for local labs
· Collecting data for or triggering milestone base for Site Payments.

El Emam et al. offers a hierarchical list of EDC functionality where systems with more advanced functionality could be assumed to have more basic functionality.[34] Such an approach suggests that a maturity model to describe EDC systems might be possible or even useful. Separating advanced embedded functionality, such as real-time interfaces with other separate operational systems (e.g. Clinical Trial Management Systems), allows for configurable workflow automation and decision support, which may be beneficial for evaluating systems.

8) EDC Methods of Delivery

There are multiple business models for delivery of the EDC software or a study application. These include technical transfer, software as a service, and software as a service with services. The most used models today are software as a service and software as a service with services.

a) Sponsor/CRO purchases, installs, and maintains the software

This is when an organization purchases a license to use software, acquires the software, and installs the software in an environment of their choosing. Requirements include having or having access to the appropriate hardware and network connectivity. Where Title 21 Part 11 compliance is required, the new installation has to be qualified, validated, and maintained under change control. There should also be procedures for the use and support of the system as well as training for system support staff. In this model, all study builds and upkeep are done in-house.

b) Software-as-a Service (SaaS) option

The SaaS option is sometimes referred to as hosting, or more recently as cloud computing and was formerly known as an Application Service Provider (ASP) model. Software as a service, e.g., provided via the internet, involves providing the hosted application as a service. The EDC vendor has all the software in their environment

and it is accessed via the internet. A SaaS model may provide early benefits when starting out utilizing an EDC system. The EDC vendor can offer support and expertise for early trials based on previous experience. EDC trials may be initiated faster since they are not dependent on a technical implementation. The pricing models differ between vendors; however, typically there will be the licensing fee to use the software, per study fees based on the time of trials, and per service costs. Study builds can be done by the sponsor or designee, e.g., a CRO or the EDC vendor.

c) SaaS with Services
Software as a service with services includes the vendor additionally providing services such as building the study application within the EDC system, training, or data management. Newer EDC applications that utilize cloud computing may offer "per-study" models. In this scenario the sponsor or designee pays for each study to be hosted. There will often be a one-time fee for building the study, the EDC vendor will build the study, there will often be a monthly subscription charge to maintain and support the study, and there may or may not be a per user fee or per site fee. In this model, the agreement with the vendor may be limited to one study or may cover multiple studies.

d) Transition from one model to another
With new technology, organizations tend to progress from fully outsourcing through various stages of bringing services and sometimes the technology itself in-house.[36] [VI] Where there is sufficient volume of studies, there is a natural progression for a SaaS with Services provider option to a SaaS model. After experience running studies with the system, an organization may gain capability and be comfortable performing study builds in-house rather than having the vendor provide that service. Similarly, an organization may gain experience and capability to maintain the software installation. These transitions offer potential cost reduction, increased convenience and control of eCRF build activities, increased access to integrated data, and opportunity for broader process optimization.

Software and support services delivery models play a large part in setting roles and responsibilities between a vendor and customer. Determining and clearly articulating the roles and relationship(s) among the EDC vendor, sponsor, and CRO is a fundamental step in selecting an EDC vendor. This is particularly important in situations, where a sponsor uses multiple EDC systems with multiple models across multiple projects.

e) Organizational Iterative Evaluation and Assessment of Current Course
New technology is being delivered to the marketplace at a fast pace. Vendors are expanding existing functionality and developing partnerships for augmented offerings. New technology to handle new sources of data are also emerging. Organizations need a forum for ongoing dialog about the current strategy, environmental scanning for new opportunities, assessment of need, value offered, and risk, and decision-making regarding piloting or integrating a new technology into the organizational tool box.[36] [VII] The fast pace of new technology delivery forces organizations into continual scanning, assessment, and decision-making.

At some point, decision-makers want to know and weigh the additional value added by a new technology. Waife points out that doing so requires organizational awareness and discipline to accurately calculate the cost of the current process as a baseline against which possible advances can be compared.[36] Such comparisons must take into account total cost within the sponsor and externally at sites and study partners.[36] [VII] Total cost also includes cost of software development of acquisition, vendor oversight and management costs, vendor charges, as well as the cost of work by all engaged in the process. Calculating cost and return on investment is further complicated by the difficulty of assigning cost to things made possible by new technology that cannot be accomplished today. Sometimes the advantages of new ways of working are not evident. This process of environmental scanning and assessment is facilitated by having resources in clinical development, informatics, or biometrics who undertake these activities,[36] undertake the communication necessary to maintain organization awareness, and have ready access to organizational decision-making processes. Such a resource could also work between operations and finance to map processes and maintain an understanding of the current costs and pain points and to develop solid measures with which to evaluate new technology.

9) Selecting and Contracting
There are multiple activities associated with selecting and contracting an EDC vendor.

a) Selecting an EDC System
Specific project needs should be identified and a Request for Information (RFI) and/or Request for Proposal (RFP) should be sent to specific EDC vendors for consideration. An RFI is a document that an organization sends to prospective vendors to ask for specific information or clarification on services or products. Often the information obtained will be used to shorten the list of vendors or contractors from which proposals will be requested. A request for proposal, or RFP, is the document that an organization sends to prospective vendors to formally request a proposal and associated cost.

Both documents are normally used in the early stages of vendor selection, with the RFI typically sent earlier than the RFP. The proposals received in response to an RFP are often used as the basis for selecting the vendor outright or for creating a short-list of vendors from whom presentations will be requested. Using a structured RFP such as that in the Appendix helps to collect the desired information from vendors and in comparing them along factors and criteria important to the organization.

Often the vendors will have the opportunity to demonstrate the EDC functionality as well as the vendor

capabilities. Key vendor personnel will be present. It is important to include key stakeholders in these presentations and to request demonstration of functionality required for your organization. Often organizations will rate proposals and functionality against a grid such as the example provided in the Appendix. Once all of the presentations have concluded, additional selection criteria may include past, ongoing, or future projects within the program, vendor performance history, vendor experience with previous industry studies, and references from current/former customers. See the Vendor Selection and Management Chapter of the GCDMP for more information on vendor selection processes.[31]

b) Contracting an EDC vendor
The Statement of Work or Scope of Work (SOW) is often an attachment to the contract with a vendor. The SOW describes in detail the work to be performed, for example, implementation support or study set-up and execution, often with estimated timelines and milestones. The SOW is usually negotiated and the negotiations on the SOW or contract language may go for several rounds before the final version is signed by all parties. Upon finalization of the SOW, the sponsor project team may choose to kick-off the project or relationship with a meeting between the EDC vendor and sponsor or CRO teams to introduce team members, review SOW highlights, discuss vendor processes and templates, identify anticipated complexities and risks, and agree on the implementation timelines and a meeting schedule. See the Vendor Selection and Management Chapter of the GCDMP for more information on contracting and management processes.[31]

c) Software Validation
In EDC studies the data collection method is much more than a piece of paper – it is an entire software application that includes workflow automation, decision support, and other study conduct aids designed to work as data are generated. Therefore the entire scope of validation activities should be completed up-front before a study starts. One of the biggest benefits of utilizing EDC is that the data are cleaned at entry and are available for review in real time. If the software is not fully validated or the study set-up is not complete when the first patient is enrolled, these advantages are lost. Refer to the GCDMP's Chapter, "Electronic Data Capture – Study Implementation and Start-up" for discussion and evidence supporting this best practice.

When using an EDC application there are important considerations that the sponsor needs to consider regarding the validation of the chosen system. The EDC system itself needs to be validated[21] [I] and each application or study that is built using this system should be tested [VI]. The US Food and Drug Administration Guidance for Industry: Computerized Systems Used in Clinical Investigations provides recommendations for the validation of the program(s) used.[37] For more information on validation refer to the aforementioned FDA regulation and guidance as well as GAMP5.

Whether developed by the sponsor or not, it is the responsibility of the sponsor to ensure that the system has been developed and is maintained per an acceptable Software Development Life Cycle (SDLC) process, including validation. Where a hosted vended system is used, to ensure the vendor's system meets the requirements, an in-person or virtual audit of the vendor's Quality Management System, including SDLC processes and documentation, is usually performed. Where software is installed locally, a vendor audit is usually conducted and is followed by installation and installation qualification and validation in the local environment. The validation process usually follows contract execution.

d) Infrastructure Implementation
After a decision has been made to implement a new EDC system, organizational infrastructure for appropriate use of the system must be developed. A key recommendation is to separate infrastructure development from the set-up and conduct of a study.[13] [VII] This section covers infrastructure development.

EDC and most new technology offers an advance of some sort. The value added from new technology including EDC is realized when the new technology enables things that were not possible or were difficult prior to the technological advance. Mechanisms through which such benefits are realized include automation, connectivity, decision support, and generating knowledge from data, such as mining data to identify new trends or undetected problems.[30] Leveraging technology to do new things or to do things in new ways changes, i.e., gaining benefit from the technology requires re-engineering organizational processes.[10,12–19] [III] As such, organizational processes should be expected to change as the organization implements new technology. Organizational process changes when implementing EDC can be significant because EDC touches organizational processes involving the collection, processing, and use of data during a study. Even where an organization is changing from one EDC system to another, different functionality and differences in how the new system supports workflow can also require significant organizational process adjustment.

Multiple people have articulated frameworks for thinking through organizational changes needed to support EDC implementation. Richardson categorizes the key implementation and support functions where change is to be expected with EDC adoption into: capable processes, human resources, qualified staff, and, implicit in his model, the technology itself.[13] He emphasizes the best practice of completing implementation prior to production use because products from these implementation and support functions are needed to build a study database and execute a study using new technology.

Capable processes are required for study build and conduct tasks. Processes documentation describes the sequence of process steps, when they are performed, and the roles that perform them and supports EDC or any new technology to be used as intended. Processes interacting with EDC impact data collection and processing, and, as stated explicitly now in ICH E6(R2), are required to be documented. Process documentation, job aids, and training facilitates consistency across sites and staff. In addition, documented processes evidence that data collection and handling procedures were established a

priori and not decided in an ad hoc manner that invites bias. Capable processes are those that have demonstrated desired and consistent behavior and quality. Process capability is established and maintained by measuring and monitoring key aspects of the process. This documented evidence demonstrates that all study functions leveraging or interacting with the software are operating as intended. Richardson refers to processes as validated processes.[13] The adjective capable is used here instead to distinguish establishing new processes or process changes from formal process validation expected in GMP contexts.

Human resources includes understanding what types of professionals and support staff will be needed to run an EDC study and gaining the commitment that they will be dedicated for the required percentage of time and duration.[13] This includes sites capable of using the technology and site staff with experience or training in EDC studies.[13] To fulfill the "human resources are ready" criteria, all individuals should have training or experience relevant to their role in running an EDC trial. All individuals should understand organizational processes for EDC studies and be prepared to apply applicable knowledge, skills, and abilities to the conduct of an EDC trial.

As described by Richardson, qualified staff means that staff are trained in the study procedures for the particular study and are prepared to execute that study.[13] This includes training sites in the conduct of the study using the EDC system, e.g., how the data collection forms operate, how discrepancies are identified and resolved, how data not entered by the site appear (if this is the case) in the EDC system and their responsibilities regarding that data, any dynamic behavior of the system such as alerts or automated workflow, and training in system security.

Spink offers a similar framework that includes technology, logistics, process, and organization.[38] In this framework, technology includes hardware, network capacity, security protocols, and integration of the EDC system with other operational infrastructure systems such as safety and randomization.[38] Logistics include the support for initial deployment of EDC to the sites and helpdesk operations. Global reach and having a Service Level Agreement (SLA) in place were noted as important to logistical support.[38] Process in the framework is concerned with scaling-up the use of EDC across an organization through process changes supporting the EDC software and ensuring that processes needed for EDC are worked into existing organizational processes. Organization includes the impact of EDC on the existing organizational roles and structure. For example, more processes, and thus roles, work in parallel and study set-up must be completed prior to the first enrollment.[38]

Though not explicitly stated in the frameworks, would-be technology adopters in organizations should assure that the technology adoption and associate process re-engineering is aligned with organizational goals and has the support of top leadership. Leadership should be briefed on the cost, benefit, and risks with new technology adoption so that decisions are well-informed by the risks and potential advantage with initiating new technology adoption and that the timing for organizational adoption is at a point in the technology-adoption curve that leadership supports. Informed decisions help assure enduring support through the challenges and set-backs inherent in new technology adoption.

Once the EDC system has been validated in the local environment or a vendor's validation has been accepted for a hosted platform, after organizational processes around study set-up and conduct using the system have been documented and tested, after professionals and support staff have been allocated and are "EDC ready", and after the study and site personnel have been trained in use of the EDC system for the study, then the study can proceed. At this point, the new technology is considered implemented. These aspects apply to set-up of a study within an EDC system, i.e., set-up can proceed after validation acceptance, process finalization, and allocation and training of professionals in their set-up tasks.

Even though pilots indicated benefit, information systems need to be evaluated in production to know if the anticipated gains have been achieved, to identify and assess any unintended consequences, and to identify opportunities for improvement.[18,39] Further, the world changes around information systems and performance or overall benefit can diminish over time. Processes and staff change. Implemented systems are stretched to support new uses. For these reasons, EDC performance should be monitored over time to drive continued improvement and to detect diminishing performance.

This will occur prior to sites being trained on the study application. The second EDC chapter, "Electronic Data Capture – Study Implementation and Start-up," picks up at this point – after EDC software is implemented with the appropriate surrounding infrastructure and ready for use on a clinical study.

10) Recommended Standard Operating Procedures

In section 5.0.1, ICH E6 states, "During protocol development the Sponsor should identify processes and data that are critical to ensure human subject protection and the reliability of trial results."[20] This implies that organizations should map out the processes involved in study design, start-up, conduct, and closeout and make explicit decisions about which are considered to impact human subject protection and the reliability of study results. Organizational processes may be partitioned differently leading to different scope and titles for SOPs. We provide the following as a list of processes commonly considered to impact human subject protection and the reliability of study results. Organizations may differ as to how these processes are covered in SOPs.

- Vendor or EDC system selection[20,21,40] (Refer to Title 21 CFR 312.52, ICH E6 R2 5.0, and Title 21 CFR Part 11.)
- EDC System Setup, Installation and Support[20,21] (Refer to Title 21 CFR Part 11 and ICH E6 R2 5.5.3 b.)
- Software Development Lifecycle including maintenance and change control[20] (Refer to ICH E6 R2 5.5.3 b.)
- Vendor auditing and oversight[20,40] (Refer to Title 21 CFR 312.52, ICH E6 R2 5.0.)

11) Literature Review Details

This revision is based on a systematic review of the peer-reviewed literature indexed for retrieval. The goals of this literature review were to (1) identify published research results and reports of EDC methods and evaluation and (2) identify, evaluate, and summarize evidence capable of informing the practice of selecting and implementing a web-based EDC system.

The following PubMed query was used:

("electronic data capture" OR "EDC" OR (internet AND "data collection")) AND ("clinical trial" OR "clinical trials" OR "clinical study" OR registry OR registries OR "observational study" OR "interventional study" OR "phase 1" OR "phase 2" OR "phase 3" OR "phase 4" OR "phase I" OR "phase II" OR "phase III" OR "phase IV" OR "first in man" OR "clinical research" OR "device study" OR "interventional trial" OR "phase 1" OR

"phase 2" OR "phase 3" OR "phase 4" OR "phase I" OR "phase II" OR "phase III" OR "phase IV" OR RCT OR "randomized clinical trial" OR "non-interventional" OR "post-marketing authorization" OR "post authorization" OR "adaptive trials" OR "feasibility study" OR "phase 2/3" OR "phase II/III" OR "phase 2a" OR "phase 2b" OR "phase IIa" OR "phase IIb" OR "phase IIb/IIIa" OR "phase 2b/3a")

The search query was customized for and executed on the following databases: PubMed (777 results); CINAHL (230 results); EMBASE (257 results); Science Citation Index/Web of Science (393 results); Association for Computing Machinery (ACM) Guide to the Computing Literature (115 results). A total of 1772 works were identified through the searches. The latest search was conducted on February 8, 2017. Search results were consolidated to obtain a list of 1368 distinct articles. Because this was the first review
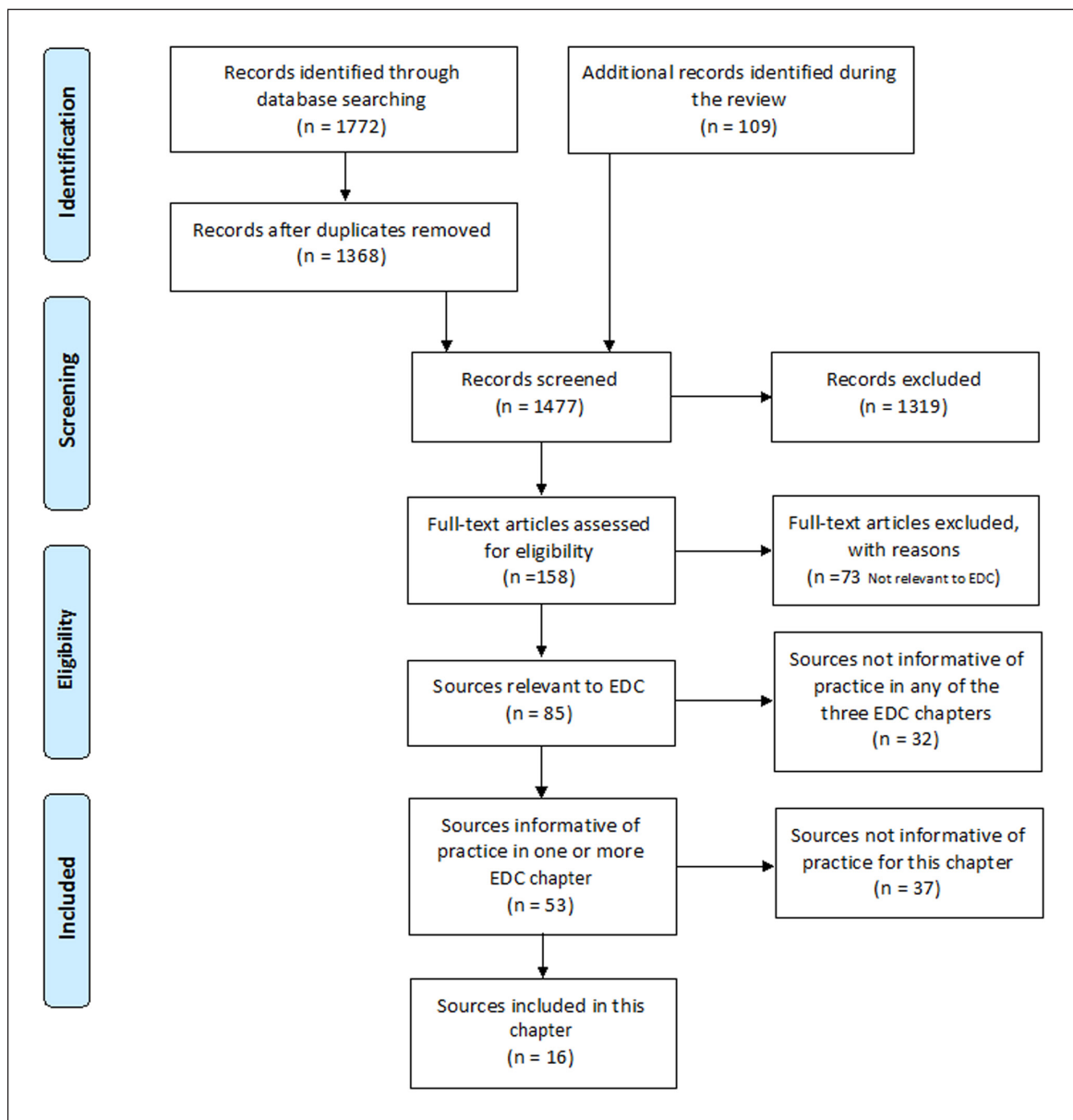

Figure 1: Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) Diagram for Web-based EDC – Selecting an EDC System.

for this chapter, the searches were not restricted to any time range.

Two reviewers used inclusion criteria to screen all abstracts. Disagreements were adjudicated by the writing group. Forty-nine sources (mostly articles) meeting inclusion criteria were selected for review. The selected sources were read by the writing group and 109 additional sources identified through the review. Each of these 158 (49 + 109) sources was read for mention of explicit practice recommendations or research results informing practice. A total of 85 sources were deemed relevant to EDC and 73 were excluded by the full text review as not relevant to EDC. Of the 85 relevant sources, 53 were identified as informative for practice in one or more of the EDC GCDMP chapters and 32 were relevant but not informative of practice in any of the three EDC chapters. Sixteen articles provided evidence for this EDC chapter (**Figure 1**). Relevant findings from these sixteen articles have been included in the chapter and graded according to the GCDMP evidence grading criteria as described in **Table 3**. This synthesis of the literature relevant to web-based EDC was performed to support the transition of the EDC chapters to an evidence-based guideline.

12) Version 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.
January 2021	Content updated and organization of material revised to reflect orientation of the material to system selection. Study start-up material moved to separate chapter.

Additional File
The additional file for this article can be found as follows:

Competing Interests
The authors have no competing interests to declare.

Acknowledgements

References

1. **Collen MF.** Clinical research databases – a historical review. *Journal of Medical Systems*. 1990; 14(6): 323–344. DOI: https://doi.org/10.1007/BF00996713

2. **Zozus M.** EDC past present and future: An introduction to the revised GCDMP chapters. *Data Basics*. 2020; 26(3).

3. **Kronmal RA, Davis K, Fisher LD, Jones RA, Gillespie MJ.** Data management for a large collaborative clinical trial (CASS: Coronary Artery Surgery Study). *Comput Biomed Res*. 1978; 11(6): 553–66. DOI: https://doi.org/10.1016/0010-4809(78)90034-4

4. **Jennett RJ, Gall D, Waterkotte GW, Warford HS.** A computerized perinatal data system for a region. *Trans Pac Coast Obstet Gynecol Soc*. 1978; 45: 51–7.

5. **Fendt K, Helms RW, Morrissey LJ.** Some Examples of Implementing Longitudinal Data Management. In *SAS American Statistical Association 1979 Proceedings of the Statistical Computing Section*.

6. **Helms RW.** A distributed flat file strategy for managing research data. In: *ACM '80 Proceedings of the ACM 1980 annual conference*. 1980: 279–285. New York: Association for Computing Machinery. DOI: https://doi.org/10.1145/800176.809982

7. **Neaton JD, Duchene AG, Svendsen KH, Wentworth D.** An examination of the efficiency of some quality assurance methods commonly employed in clinical trials. *Stat Med*. 1990; 9(1–2): 115–23. DOI: https://doi.org/10.1002/sim.4780090118

8. **Christiansen DH, Hosking JD, Dannenberg AL, Williams OD.** Computer-assisted data collection in multicenter epidemiologic research: The atherosclerosis risk in communities study. *Control Clin Trials*. 1990; 11(2): 101–15. DOI: https://doi.org/10.1016/0197-2456(90)90004-L

9. **Reynolds-Haertle RA, McBride R.** Single vs. double data entry in CAST. *Control Clin Trials*. 1992; 13(6): 487–94. DOI: https://doi.org/10.1016/0197-2456(92)90205-E

10. **Kush RD, Bleicher P, Kubick W,** et al. *eClinical trials planning and implementation*. Boston: Thompson Centerwatch Inc.; 2003.

11. **Helms RW.** Data quality issues in electronic data capture. *Drug Inf J*. 2001; 35: 827–837. DOI: https://doi.org/10.1177/009286150103500320

12. **Dimenas E, Johansson D, Palmblad M.** Wrangstadh M. Clinical operations online (COOL)-a world wide web-based approach to running clinical trials: Results from two international multicenter gastrointestinal trials. *Drug Inf J*. 2001; 35: 745–753. DOI: https://doi.org/10.1177/009286150103500313

13. **Richardson A.** Planing and running the eClinical trial. *Appl Clin Trials*. 2003; January.

14. **Green JA.** Realizing the value proposition of EDC. *Innovations in Clinical Trials*; 2003.

15. **Lu Z.** Electronic data-capturing technology for clinical trials, experience with a global postmarketing study. *IEEE Eng Med Biol Mag*. 2010; 29(2): 95–102. DOI: https://doi.org/10.1109/MEMB.2009.935726

16. **Mitchel JT, Kim YJ, Choi J,** et al. The impact of electronic data capture on clinical data management perspectives from the present into the future. *MONITOR.* 2008; August.

17. **Mitchel JT, Dean Gittleman D, Park G,** et al. The impact on clinical research sites when direct data entry occurs at the time of the office visit: A tale of 6 studies. *InSite.* 2014; Second Quarter.

18. **Kush R.** Electronic Data Capture-Pros and Cons. *BioExecutive International.* 2006; 2(6): S48–52.

19. **Laky D.** The evolution of EDC into eClinical. *epc magazine.* 2007; 11(2). Available from: http://www.samedanltd.com/magazine/11/issue/2. Accessed September 29, 2020.

20. **Food and Drug Administration, US Department of Health and Human Services.** ICH E6(R2) Good Clinical Practice: Integrated Addendum to ICH E6(R1), March 2018. Available from https://www.fda.gov/regulatory-information/search-fda-guidance-documents/e6r2-good-clinical-practice-integrated-addendum-ich-e6r1.

21. **Food and Drug Administration, US Department of Health and Human Services.** Electronic Records; Electronic Signatures, 21 CFR §11 (1997). Available from https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm.

22. **Food and Drug Administration, US Department of Health and Human Services.** Guidance for industry: Providing Regulatory Submissions in Electronic Format – Standardized Study Data. December 2014. Available from https://www.fda.gov/regulatory-information/search-fda-guidance-documents/providing-regulatory-submissions-electronic-format-standardized-study-data.

23. Medicines and Healthcare products Regulatory Agency (MHRA) 'GXP' Data Integrity Guidance and Definitions, Revision 1. March 2018. Available from https://www.gov.uk/government/publications/guidance-on-gxp-data-integrity.

24. **Food and Drug Administration, US Department of Health and Human Services.** General Principles of Software Validation; Final Guidance for Industry and FDA Staff, January 2002. Available at https://www.fda.gov/regulatory-information/search-fda-guidance-documents/general-principles-software-validation.

25. European Commission Health and Consumers Directorate-general, EudraLex; The Rules Governing Medicinal Products in the European Union Volume 4: Annex 11: Computerised Systems (2011). Available from https://ec.europa.eu/health/documents/eudralex/vol-4_ga.

26. *GAMP® 5 A Risk-based Approach to Compliant GxP Computerized Systems.* North Bethesda, MD: International Society for Pharmaceutical Engineering (ISPE). 2008.

27. **Food and Drug Administration, US Department of Health and Human Services.** Guidance for industry: Use of Electronic Health Record Data in Clinical Investigations. July 2018. Accessed September 17, 2020. Available from https://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-electronic-health-record-data-clinical-investigations-guidance-industry.

28. **Food and Drug Administration, US Department of Health and Human Services.** Guidance for Industry: Electronic Source Data in Clinical Investigations. September 2013. Accessed August 8, 2018. Available from https://www.fda.gov/regulatory-information/search-fda-guidance-documents/electronic-source-data-clinical-investigations.

29. **Food and Drug Administration, US Department of Health and Human Services.** Investigational New Drug Application, General responsibilities of sponsors, 21 CFR §312.50 (2019). Available from https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?fr=312.50.

30. **Stead WW, Lin HS.** National Research Council. *Computational Technology for Effective Health Care: Immediate Steps and Strategic Directions.* Washington, DC: The National Academies Press. 2009. DOI: https://doi.org/10.17226/12572

31. **Amatya S, Edgerton D.** Vendor selection and management. *Data Basics.* 2020; 26(1).

32. **Capterra Inc.** Clinical Trial Management Software search website. Available from: https://www.capterra.com/clinical-trial-management-software/?utf8=%E2%9C%93&users=&feature%5B2%5D=31834&commit=Filter+Results&sort_options=. Accessed January 7, 2019.

33. **McFadden ET, LoPresti F, Bailey LR, Clarke E, Wilkins PC.** Approaches to data management. *Control Clin Trials.* 1995; 16: 3OS–65s. DOI: https://doi.org/10.1016/0197-2456(94)00093-I

34. **El Emam K, Jonker E, Sampson M, KrležaJeric K, Neisa A.** The use of electronic data capture tools in clinical trials: web survey of 259 Canadian trials. *J Med Internet Res.* 2009; 11(1): e8. DOI: https://doi.org/10.2196/jmir.1120

35. **Franklin JD, Guidry A, Brinkley JF.** A partnership approach for electronic data capture in small-scale clinical trials. *J Biomed Inform.* 2011; 44: S103–S108. DOI: https://doi.org/10.1016/j.jbi.2011.05.008

36. **Waife RS.** Transitioning clinical data management from the 1980s to the 2010s: Strategies for corporate decision making. *Drug Inf J.* 2001; 35: 713–719. DOI: https://doi.org/10.1177/009286150103500309

37. **Food and Drug Administration, US Department of Health and Human Services.** Guidance for Industry: Computerized Systems Used in Clinical Investigations. May, 2007. Available from https://www.fda.gov/regulatory-information/search-fda-guidance-documents/computerized-systems-used-clinical-investigations.

38. **Spink C.** Electronic Data Capture (EDC) as a means for e-clinical trial success, IBM Global Services

(report BIWUK00002-2-0202-AP). *Pharmaceutical Clinical Development.* 2002; March.

39. **Friedman CP, Wyatt JC.** Evaluation Methods in Biomedical Informatics, Second Ed. New York: Springer Science Business Media, Inc.; 2006. DOI: https://doi.org/10.1007/0-387-30677-3

40. **Food and Drug Administration, US Department of Health and Human Services.** Investigational New Drug Application, Transfer of obligations to a contract research organization, 21 CFR §312.52 (1997). Available from https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm.