Clinical Data Interchange Standards Consortium

Electronic Source Data Interchange (eSDI) Group

Leveraging the CDISC Standards to Facilitate the use of Electronic Source Data within Clinical Trials

Version 1.0

Date 20th November 2006

Table of Contents

Executive Summary

The application of information technology has been shown to improve data quality and patient safety, particularly in the healthcare environment and with electronic source data (eSource) collection instruments. The use of eSource can improve the assessment of patient compliance in trials using electronic Diaries. To augment the benefits of IT, the use of standards can facilitate data interchange among various parties using disparate systems and data sources/databases and enable a better information link between research and healthcare. Therefore, the FDA would like to encourage the use of such technology, standards and processes for clinical trials.

Unfortunately, existing regulations and guidelines (which were developed in the world of paper), are not entirely clear on the processes and accountabilities, when new technology is introduced; they do not specifically address many of the issues involving eSource or how investigator and sponsor responsibilities should be fulfilled when electronic data capture is used in clinical trials. Furthermore, the roles that data standards can play to synergistically improve the clinical trial process and to meet regulatory requirements when new technologies are implemented need to be articulated; they are not included in current regulations and guidance.

The Clinical Data Interchange Standards Consortium (CDISC) is an open, multidisciplinary group that has led the development of global, vendor-neutral, platform-independent standards, to improve data quality and accelerate product development in our industry. The CDISC submission standard has been acknowledged by FDA for submitting clinical trial data to FDA. The current mission of CDISC, ***to develop and support global, platform-independent data standards that enable information system interoperability to improve medical research and related areas of healthcare***, speaks of the desire of CDISC to facilitate the use of eSource, particularly in the context of electronic health records and patient reported outcomes, for clinical research as well as healthcare.

Because the FDA is interested in leveraging standards throughout the clinical trial process, as evidenced through the FDA Critical Path Initiative and Opportunities List, CDISC and the FDA are exploring upstream uses of the CDISC standards and the value beyond just regulatory data submission. Such benefits include efficient, economical storage and archive of electronic data (along with audit trail, administrative information and edit checks) and enabling standard means of audit/review of this information. With the encouragement of the FDA, CDISC therefore initiated the eSource Data Interchange Group to discuss current issues related to eSource data in clinical trials and to make recommendations for the use of standards and processes to encourage eSDI within the context of existing regulations. ***The specific objective of the eSDI Group was to produce a document that aligns multiple factors in the current regulatory environment, to encourage the use of eSource collection and industry data standards to facilitate clinical research for investigators, sponsors and other stakeholders***.

The eSDI group has done an assessment of the existing regulations in the context of eSource data, identified issues that may inhibit adoption, explored the value and benefits of implementing standards for data acquisition, exchange and archive of eSource. Along with reviews and input from external reviewers, the eSDI group has developed

1. User requirements that can be used as a checklist to ensure that regulations are being addressed with these solutions included as recommendations from the eSDI group;

2. Five scenarios for processes, including benefits from standards, to address key areas of the eSource data interchange;
3. Outlined recommendations for updating the existing regulatory framework;
4. Provided a checklist for investigators creating awareness of their responsibilities when dealing with eSource; and
5. Provided a template for sponsor companies to document their compliance with the source data regulations.

The scenarios developed by the group center around: a) the storage of eSource at the investigative site; b) use of an eSource system provider (contracted supplier); c) the Single Source Concept (leveraging standards to enter eSource data simultaneously into an electronic health record system or system at a site and a clinical study systems, EDC or database); d) eSource extraction and investigator verification (using electronic health records; and e) direct extraction of clinical trial data from electronic health records (EHR), as an alternative to acknowledge the ultimate vision for research-healthcare data flow. For each scenario, the benefits and value of standards are also included.

A key goal was to provide these recommendations as a benefit to the industry and the FDA to encourage the adoption of data interchange standards – in particular the ODM – and suitable associated processes to facilitate and encourage participation in electronic clinical research, including bridging the gap between medical research and healthcare. Desired outcomes would be to facilitate investigator participation in clinical research and ultimately to improve patient care.

Purpose

This document is intended to align multiple factors in the current global regulatory environment to encourage the use of electronic source data (eSource) collection and industry data standards to facilitate clinical and biomedical research for investigators, sponsors and other stakeholders.

It is recognized that the existing regulations are largely based upon paper-based processes and that changes could clarify and/or streamline electronic data collection-based processes; however, these changes will take time while new technologies are available today. This document is focused on today's environment and what is feasible with respect to electronic source (eSource) trials.

Objectives

Specific objectives for the electronic Source Data Interchange (eSDI) project, in the context of the above-stated purpose, are to:

a) provide benefit to the industry and regulatory authorities by leveraging the clinical research expertise in the eSDI Group and CDISC to clarify the value of data interchange standards and appropriate processes to streamline trials employing eSource data collection;

b) provide a set of base regulatory requirements to assist those conducting trials using eSource data collection in their planning and execution of such trials in today's regulatory environment;

c) provide potential scenarios that exemplify the use of CDISC data standards and appropriate processes for eSource data collection and interchange;

This will help pave the way towards a vision of 'research at the point of care and care at the point of research' and ultimately to enable information system interoperability to improve medical research and related areas of healthcare; the core of the CDISC mission

Scope

With respect to scope, the eSDI Initiative covers the eSource data interchange processes and standards as they relate to data collection/acquisition, interchange and archive of eSource for global regulated clinical and biomedical research.

Document References

[1] Code of Federal Regulations, Title 21 CFR, Part 11: Electronic Records; Electronic Signatures; Final Rule, Federal Register, March 20, 1997.

[2] Code of Federal Regulations, Title 21 CFR, Part 312: Investigational New Drug Application, Federal Register, April 1, 2002.

[3] International Conference on Harmonization, Good Clinical Practice: Consolidated Guideline, Federal Register Vol 62, No. 90, 25711, May 9, 1997.

[4] Food and Drug Administration. Guidance for Industry: Computerized Systems Used in Clinical Trials. FDA April 1999.

[5] Food and Drug Administration. Guidance for Industry: Part 11, Electronic Records; Electronic Signatures – Scope and Application. FDA August 2003.

[6] Paul Bleicher, "eSource Redux", *Applied Clinical Trials,* August 2002, 30-31.

[7] Teri Stokes and Jean Paty, "Electronic Diaries, Part 1. What is a Subject Diary, and How Do Regulations Apply" *Applied Clinical Trials*, September 2002, 38-43.

[8] Stephen A Raymond and Gerald F. Meyer, "Interpretation of Regulatory Requirements by Technology providers. The Case for Electronic Source Data" *Applied Clinical Trails*, June 2002, 50-58

[9] Dave Iberson-Hurst, "Electronic Diaries: Source Data Out in the Open" *Applied Clinical Trials,* EDC Supplement, February 2004, 16-21

[10] Research Project Results: 2002 (CDISC-CenterWatch), 2003 (CDISC-CenterWatch) and 2004 (CDISC).

[11] FDA Critical Path Initiative. FDA website. http://www.fda.gov/oc/initiatives/criticalpath/ Site accessed on 25 May 2005.

[12] The Economist, April 30th 2005 "The no-computer virus" Special report :IT in the health-care industry, pages 65-67.

[13] Stone et al. Patient non-compliance with paper diaries. BMJ 2002;324;1193-1194

[14] Dave Iberson-Hurst, "The CDISC Operational Data Model: Ready to Roll?" *Applied Clinical Trials,* EDC July 2004, 48-53

[15] Diane Carr, Queens Health Network, Presentation for CDISC-DIA eClinical Conference, October 2004.

[16] Joanne L. Rhoads, M.D., MPH. Director, Division of Scientific Investigations, CDER. e-PRO Source Documentation: FDA Regulatory Concerns. Presentation at the DIA Workshop, April 5, 2005.

[17] Rebecca Kush and David Hardison, "How Necessary are Data Standards?", Scrip Magazine, May 2004 (http://www.cdisc.org/pdf/ScripMay04.pdf)

[18] CDISC LAB Standard (http://www.cdisc.org/models/lab/v1.0.1/index.html)

[19] HL7 ECG Waveform Standards and HL7 V3 Message for Periodic Reporting of Clinical Laboratory Data (www.hl7.org)

[20] "FDA Announces Standard Format That Drug Sponsors Can Use to Submit Human Drug Clinical Trial Data", FDA News, Department of Health and Human Services Announcement, 21 July 2004.

[21] Case Report Tabulation Study Data Tabulation Model (CRTDDS) (http://www.cdisc.org/models/def/v1.0/index.html)

[22] Study Data Specifications for the eCTD for submissions using the SDTM available at http://www.fda.gov/cder/regulatory/ersr/ectd.htm.

[23] CDISC Operational Data Model (http://www.cdisc.org/models/odm/v1.2.1/index.html)

[24] FDA Final Guidance "Providing Regulatory Submissions in Electronic Format—Human Pharmaceutical Product Applications and related Submissions Using the eCTD Specifciations" (http://www.fda.gov/cder/guidance/7087rev.pdf)

[25] Critical Path Opportunities List (http://www.fda.gov/oc/initiatives/criticalpath/reports/opp_list.pdf)

[26] Clinical Data Interchange Standards Consortium (www.cdisc.org)

[27] CDISC Glossary, Abbreviation and Acronyms (http://www.cdisc.org/glossary/index.html)

Acronyms and Abbreviations

CDISC	Clinical Data Interchange Standards Consortium
CDER	Center for Drug Evaluation and Research
CDMS	Clinical Data Management System
CFR	Code of Federal Regulations
CRF	Case Report Forms
CRO	Clinical Research Organization
CRT DDS	Case Report Tabulations – Data Definitions Specification
CSUCT	Computerized Systems Used In Clinical Trial
ECG	Electrocardiogram
eCRF	electronic Case Report Form
eCTD	electronic Common Technical Document
ePRO	Electronic Patient Reported Outcomes
EDC	Electronic Data Capture
EHR	Electronic Health Record
EMR	Electronic Medical Record
eSDI	electronic Source Data Interchange
FDA	Food and Drug Administration
GCP	Good Clinical Practice
HIPAA	Health Insurance Portability And Accountability Act
HL7	Health Level Seven
ICH	International Conference on Harmonisation
IT	Information Technology
ODM	Operational Data Model
PDA	Personal Digital Assistant
PDF	Portable Document Format
PhRMA	Pharmaceutical Research and Manufacturers of America
PRO	Patient Reported Outcome
SDTM	Study Data Tabulation Model
XML	Extensible Markup Language

Glossary

For a glossary of terms and additional abbreviations, please see reference [27]

Rationale and Introduction

Information technology demonstrated improved data quality and patient safety in electronic medical documentation [12, 15]. The use of electronic diaries vs. paper diaries improved the accuracy of assessing subject compliance in clinical trials [13]. EDC can reduce incoming errors by approximately two-thirds, if edit checks are implemented at the point of data collection at the site, since the errors are addressed at the point of data collection rather than later in the process. With electronic data capture, legibility becomes far less of an issue and more rapid feedback on database requirements reduces the burden of query resolution on sites and monitors. Ready access to the data facilitates project management, and electronic data reduces capacity issues with archive. The FDA recognizes these and additional benefits of information technology and does not want to inhibit the biopharmaceutical industry from also benefiting. In fact, streamlining clinical trials and leveraging standards are at the core of the FDA Critical Path Initiative [11].

There are additional benefits of electronic data capture that can be gained when Clinical Data Interchange Standards Consortium (CDISC) standards are leveraged with the technology and appropriate processes, particularly when exchanging data among various organizations (sites, sponsors, vendors, regulatory authorities) and when using different technologies.[17] Sponsors and site personnel have indicated their desire to encourage the use of standards, not only for reporting and submission but also at the sites in the data collection processes and in the use of new technologies. .[10] Out of 300 sponsors and 192 contract research organizations globally, in a research project conducted by CDISC and CenterWatch, over 90% agree that *"Standards should be extended to facilitate data collection at investigative sites."* Ninety-four percent of site representatives, who responded to the CDISC research project surveys in 2004, agreed with the statement "*Sponsors should collaborate in the standardization of practices and data collection systems for investigative sites."*

The value of eSource data collection and interchange extends beyond assessing patient compliance and improving data quality. Subjects might enter their own information electronically, thus, 'opting into' trials voluntarily and streamlining data collection. There is the potential for a significant reduction in time for monitoring in terms of source data verification. Integrating clinical research capabilities into electronic health record systems also increases the potential to obtain more safety surveillance information and reduces redundant data collection and transcription for investigators and site personnel. If investigators were using electronic health record systems to facilitate clinical research in addition to patient care, this could also decrease the number of systems in their office space. Sixty-eight percent (68%) of the sites that responded to the 2004 research survey have more than one system/application operating concurrently for collecting data into electronic case report forms for clinical trials; 17% sites have five or more. [10]

The vast majority of sites and sponsors agree that eSource is the future and that it is time to pave the way for these opportunities and benefits. Eighty-three percent of sponsors answered 'Yes' to the statement "*Would you advocate the use of eSource now or in the future, i.e. the entry of data (excluding Laboratory and electronic Patient Reported Outcomes data) electronically without first capturing the data on paper?"* Also in these research projects, conducted by CDISC and

CDISC-CenterWatch, sites ranked *"Electronic source documentation online"* first in a list of ways to better leverage technology.

Clinical laboratory data and ECG waveform data have been collected as eSource data for years, and there are now interchange standards available (CDISC and HL7) to support the transfer of these data in standard formats among stakeholders.[18,19] These are readily accepted by regulatory authorities. Hence, we must now ask what issues are different between the collection of laboratory and ECG data and the collection and interchange of other research data, i.e. what needs to be implemented to further progress eSource data collection and interchange and interoperability among research and healthcare systems.

The work of the eSource Data Interchange Group began with a desire to determine how to leverage standards to facilitate eSource data acquisition, exchange and archive in the context of today's regulatory environment. To be able to leverage the standards appropriately, it is essential to understand the relevant regulations and requirements. This document, therefore, provides a set of user requirements that have been generated through extensive analyses of existing regulations. (Appendices are available with these analyses and tables of how the user requirements map to different technologies.) The document then provides scenarios that can meet the user requirements with today's technologies applied to eSource trials within the context of existing regulations and where the standards can be leveraged to facilitate the relevant processes. Considerations for paving the way to the future, when clinical research and healthcare systems are interoperable and the standards harmonized, have been included in the discussions towards these scenarios.

Background Relevant to the eSDI Initiative

CDISC is an open, global, non-profit organization that has now established industry standards to support the electronic acquisition, exchange, submission and archiving of clinical and non-clinical study data and metadata for medical and biopharmaceutical product development. [26] CDISC was initiated in late 1997 and its accomplishments can be attributed to countless volunteers from multidisciplinary functions. CDISC standards are vendor-neutral and platform-independent. The mission of CDISC was expanded in scope in 2004 to develop and support *global, platform-independent data standards that enable information system interoperability to improve medical research and related areas of healthcare.* This vision was considered in embarking on the eSDI project.

Those involved in regulatory submissions to the U.S. Food and Drug Administration (FDA) have become aware of the CDISC Study Data Tabulation Model (SDTM) to standardize the format in which electronic data can be provided to facilitate regulatory reviews. [20] The SDTM metadata can also be submitted using the CDISC XML transport standard (the Operational Data Model). For this purpose, the CDISC Case Report Tabulation Data Definition Specification (define.xml) is deployed. [21] Both SDTM and CRTDDS (define.xml) are now listed as specifications in FDA Final Guidance, "Providing Regulatory Submissions in Electronic Format—Human Pharmaceutical Product Applications and related Submissions Using the eCTD Specifciations" [22, 24].

The CDISC Operational Data Model was designed to support the acquisition, exchange and archive of electronic data in a standard XML-based format.[14, 23] A core requirement of the standard was that it be able to support all existing regulations applicable to these clinical trial processes. The ODM provides an effective means to archive electronic data, at an investigative site or a sponsor setting, without requiring that the system be 'mothballed' in order to retrieve and review the data at a later point in time. The ODM includes the ability to capture audit trail information in a standard format, and it supports e-signatures and other requirements of the regulation 21 CFR 11. In addition, ODM provides a means to use the standard audit trail to facilitate data review (by sponsors or regulators), with indicators of data integrity (number of times data fields have been changed) and a means to automatically generate CRFs, in which the data fields can be based upon the SDTM metadata (to facilitate data collection and/or monitoring or data review).

The FDA Critical Path Initiative and Opportunities List [11, 25] specifically reference the need for standards to streamline clinical trials. It was specifically attractive to FDA representatives involved in eSource and ePRO initiatives to obtain more information and input from sectors of the industry that had been relatively quiet with respect to the concerns of electronic data capture, in particular eSource. They were especially interested in hearing from site representatives and sponsors. The suggestion was made for CDISC to convene a multidisciplinary group with representatives from these and other sectors. Another important consideration was to ensure that the discussions remain neutral with respect to any particular solution or application. Hence, the eSource Data Interchange Group was formed for the specific purpose of convening a group to generate recommendations on processes and standards that could facilitate the use of eSource Data Interchange for regulated clinical trials in the context of the existing regulations. The process for generation of this document is detailed in Appendix 6.

Source Documents and Data

Introduction

The increasing use of computers within clinical trials, driven by the desire to speed drug development times and reduce costs, has resulted in an increasing interchange of electronic data. A significant part of that data falls within the scope of the Food and Drug Administration's (FDA) predicate rules[1] and, due to its electronic nature, the scope of 21 CFR Part 11. One class of data in particular, source data, has caused particular concern within the industry. Whereas in the traditional world, data were recorded on the tangible and comfort-giving piece of paper, now the same data are stored electronically, allowing the information to be quickly copied, transferred, changed or deleted. Therefore, the industry must consider how the potential benefits of electronic source can be realized, while minimizing the risks and impact on current practices and personnel that it brings.

When undergoing periods of change, and with the exciting lure that new technology can bring to clinical trials, it becomes easy to focus solely on the advantages of the change, at the expense of ignoring potential exposures that may arise with the use of the new technology. It is also possible that one of two scenarios may occur within the regulatory environment: a) as the technology claims tend to focus more on the speed of process, and the data availability advantages are touted, regulations may be ignored; or b) exaggerated attention is given to the exact wording of the regulations, and implementation activities grind to a halt with spiraling analysis of the software features and debate over regulatory interpretation. Either scenario can result in the delay of a timely, appropriate and compliant introduction of new technology.

When considering the issues arising with the use of computer systems in clinical trials, and especially eSource and documents, it becomes important to heed the regulatory expectations. However, since many of the clinical trial regulations were developed prior to the extensive use of computer systems, it may not be as easy to understand how to apply the terminology in the regulations to the "e" environment. It becomes important to fully understand the intent of the FDA regulations and to look at the objectives rather than the precise detail.

In conducting the trial at a particular trial site, the clinical investigator generates, collects and records data in support of the trial. This source data may be in a variety of records, including medical records, patient charts, laboratory test results, case report forms (if original entries are recorded on those), ECGs, x-rays, digital photographs, and patient diaries. The source data serves as the basis for subsequent decisions and analysis, both by the clinical investigator for the subject's medical care and by the sponsor to reach conclusions on a drug's safety and efficacy. Sponsors use copies of the data recorded at the clinical site, summarize it, derive new variables,

[1] **Predicate rule:**

This term refers to underlying requirements set forth in the Federal Food, Drug, and Cosmetic Act, the Public Health Service Act, and FDA regulations (other than 21 CFR part 11). Regulations governing good clinical practice and human subject protection relevant to eSource can be found at 21 CFR parts 50, 56, 312, 511, and 812.

Source: FDA

and perform various analyses to reach their conclusions. Regulators need to reconstruct the trial by comparing the data submitted to the agency by the sponsor with the source data prepared and maintained at the investigational site. Significant data movement and manipulation can occur between systems (clinical data management database, analysis database, derived datasets) and between business partners such as sponsors, CROs, clinical labs, and image reading services. Therefore, it becomes critical to ensure that regulators can always return to the original data and follow the trail to the ultimate conclusions drawn. It is for these reasons that the regulatory agencies place such significance on the trustworthiness of the data collected during a trial. There are likely to be multiple ways in which the regulations can be met, both with technical as well as procedural controls. It is important that we maintain an understanding of the regulatory objectives and look to the spirit of the regulations, as we investigate and evaluate changes in process discussed within this document[2].

Electronic source data (eSource) is source data that are captured in an electronic form[3] rather than on paper. Source data[4] are all data held in original records, or certified copies thereof, necessary for the evaluation of the trial. Such data are held in source documents[5].

---

[2] The fundamental question is "why do we have source data?" One of the general principles from ICH GCP should be noted:

All clinical trial information should be recorded, handled, and stored in a way that allows its accurate reporting, interpretation and verification.

Source: ICH GCP, section 2.10

[3] **eSource:**

Source Data captured initially into a permanent electronic record.

Source: CDISC

[4] **Source Data:**

All information in original records and certified copies of original records of clinical findings, observations, or other activities in a clinical trial necessary for the reconstruction and evaluation of the trial. Source data are contained in source documents (original records or certified copies).

Source: ICH GCP, Section 1.51

[5] **Source Documents:**

Original documents, data, and records (e.g., hospital records, clinical and office charts, laboratory notes, memoranda, subjects' diaries or evaluation checklists, pharmacy dispensing records, recorded data from automated instruments, copies or transcriptions certified after verification as being accurate copies, microfiches, photographic negatives, microfilm or magnetic media, x-rays, subject files, and records kept at the pharmacy, at the laboratories and at medico-technical departments involved in the clinical trial).

Source: ICH GCP, Section 1.52

The scientific integrity of the trial and the credibility of the data from the trial depend substantially on the trial design. A description of the trial design, should include:

…

The identification of any data to be recorded directly on the CRFs (i.e. no prior written or electronic record of data), and to be considered to be source data.

Source: ICH GCP, Section 6.4. and 6.4.9

Source data and the associated source documents are the foundation of clinical research. The research undertaken is driven by the clinical protocol and the design documented therein. This design will define which data are to be collected for each subject enrolled within the trial. That data are collected and carefully recorded to ensure that the data are of sufficient quality such that the regulatory authorities and the public can trust the conclusions based on the data[6]. It is this original (the first initial) recording that constitutes source data[7]. It should be noted that, irrespective of the technology used to capture the data, be it paper or electronic means, the important criteria is that the data can be trusted.

In the paper world, source data captured by an investigator is collected within one of three types of source documents: a) the subject's own medical record; b) directly onto a Case Report Form[8] (CRF); or c) onto some other piece of paper that is neither part of the medical record nor a CRF[9]. The data will typically be copied to a CRF (if not directly entered on a CRF) and submitted to the sponsor, while the investigator will retain a copy of the CRF that will be incorporated into the subject's case history[10].

In the case of a paper diary[11], the subject will enter data into the diary and return the completed diary to the investigator. The investigator will then forward the diary to the sponsor with a copy being kept as part of the subject's case history.

With the advent of electronic diaries, it can be seen that the paper source document disappears; the paper diary is no longer present, being replaced by a cell phone, a Personal Digital Assistant (PDA) or some other electronic data capture technology. The question arises as to where do the

---

[6] During the work three criteria where developed that outlined the key concerns of the agency: a) How do we ensure that the data submitted are the data captured? b) How do we ensure the data captured is accurate?; and c) How do we ensure the subject's safety?

[7] It should be noted that this original data can be copied and that new copy considered the source if the copying operation is verified as being accurate.

[8] **Case Report Form:**

A printed, optical, or electronic document designed to record all of the protocol required information to be reported to the sponsor on each trial subject

Source: ICH GCP, section 1.11

[9] This third case is less common but has been included so as to cover all eventualities. One comment on the document noted that "This is actually not so rare. CIs often use worksheets, shadow charts".

[10] **Case History:**

A Case History contains all observations and other data pertinent to the investigation on each individual administered the investigational drug or employed as a control in the investigation. Case histories include the case report forms and supporting data including, for example, signed and dated consent forms and medical records including, for example, progress notes of the physician, the individual`s hospital chart(s), and the nurses notes. The case history for each individual shall document that informed consent was obtained prior to participation in the study.

Source: 21 CFR part 312, Section 62(b)

[11] Diary and electronic Diary (eDiary). Terms used to refer to the mechanisms used to collect Patient Reported Outcome (PRO) data. encompassing diaries, diagnostic instruments, therapeutic area specific severity measures, quality of life assessments and pharmaeconmonics or work productivity assessments

source data reside? Within the electronic scenario, there are data being collected and stored, perhaps for very short periods of time, on electronic devices, that are eventually transmitted[12] to servers located at the vendor's or sponsor's premises. This could result in multiple copies of the source data, leaving one to question which one is the source? In addition, the investigator may not have had visibility of the subject-reported data, raising concerns about the ability to provide effective patient care.

Concerns have also been expressed regarding electronic CRF systems where data are entered directly into an electronic system without ever being captured on paper. As with electronic diary systems, questions arise over the location of the source data and the responsibilities of an investigator to maintain suitable case histories.

[12] It is noted that some technologies, such as Interactive Voice Response Systems (IVRS), transmit the data directly to servers.

Analysis of Existing Practice and Regulations

In order to answer the questions raised by eSource and present options to industry, an examination of the process and the associated regulations has been undertaken as a means of deriving the fundamental objectives for source data and source documents. By deriving requirements that are technology independent – and it should be remembered that paper can be considered a technology – the requirements can be taken forward and used to assess whether a given implementation and/or technology will meet the needs of the FDA, sponsors and investigators alike. Three analyses were undertaken: a) examining the paper process; b) examining the electronic regulations; and c) looking at existing industry practices, to derive a total of 12 requirements; nine from the first, one from the second and two from the third analysis.

Analysis of Paper Source Documents and Process

The first analysis, presented in Appendix 1, is based around the use of paper source documents. It is well understood that the paper process can be imperfect, see reference [13]. However, the premise for undertaking such an analysis is that a well-structured process, based on the use of paper source documents, can meet the agency's current predicate rules. Therefore such a process exhibits the key principles that the agency requires in the collection of clinical trial data and, by detailing these principles, they can then be taken forward into the electronic world.

The analysis presented in Appendix 1 results in a set of user requirements for source data held within source documents, irrespective of the media or technology used to hold the data:

Requirement 1: An instrument used to capture source data shall ensure that the data are captured as specified within the protocol.

Requirement 2: Source data shall be Accurate, Legible, Contemporaneous, Original, Attributable, Complete and Consistent.

Requirement 3: An audit trail shall be maintained as part of the source documents for the original creation and subsequent modification of all source data.

Requirement 4: The storage of source documents shall provide for their ready retrieval.

Requirement 5: The investigator shall maintain the original source document or a certified copy.

Requirement 6: Source data shall only be modified with the knowledge or approval of the investigator.

Requirement 7: Source documents and data shall be protected from destruction.

Requirement 8: The source document shall allow for accurate copies to be made.

Requirement 9: Source documents shall be protected against unauthorized access.

The requirements are mapped to the regulations in Appendix 5.

Existing Electronic Record Regulations

The above analysis has examined source documents and source data from the perspective of a paper world. However, it is also necessary to examine the impact on the user requirements when electronic records are considered.

21 CFR Part 11 details requirements for records identified within the predicate rules and held in an electronic form. As such, eSource falls under the requirements of the regulation. 21 CFR Part 11 can, for the purpose of the discussion within this paper, be split into two parts: a) the controls for electronic records (in Open or Closed Systems); and b) the requirements for electronic signatures.

Electronic signatures, while important, do not impact the underlying predicate regulations for the storage of source document and data. The regulations for source documents and data are technology independent. If records were stored using a paper-based system then 21 CFR Part 11 would not apply. If those same records were stored electronically then 21 CFR Part 11 would apply and there would be a potential need for electronic signatures. We can therefore think of the requirements driven by 21 CFR Part 11 as being layered on top of the source data regulations, they are in addition to the predicate rules. Given that the analysis is considering the predicate regulations irrespective of the form in which they are stored, electronic signatures are not considered further. However, should source documents and data be stored electronically; the demands of the regulation will need to be met.

Appendix 2 details the analysis undertaken with respect to the 21 CFR Part 11 regulation. The analysis results in the addition of a single new core requirement.

Requirement 10: The sponsor shall not have exclusive control of a source document.

A full explanation of the regulatory basis for this requirement is contained within Appendix 2[13].

[13] See also reference [16]: Joanne L. Rhoads, M.D., MPH. Director, Division of Scientific Investigations, CDER. "e-PRO Source Documentation: FDA Regulatory Concerns". Presentation at the DIA Workshop, April 5, 2005. It should be noted (as indicated in the next section on User requirements and Definitions) that this requirement does not preclude such circumstances as phase 1 units operating within sponsor organizations. These trials still require individuals operating in the roles of sponsor and investigator and their responsibilities are as per other trials.

Existing Practice with Electronic Technologies

A number of technologies are already deployed as part of clinical trials. These include systems used to capture subject data, such as eCRF, diagnostic data as well as electronic diary systems. Given the development of the user requirements, it seems logical to assess existing electronic systems and practices against one another. Appendix 3 contains such an analysis.

Three issues emerge from the analysis of the key requirements against typical technology architectures, namely that: a) a single instance of source data located at the sponsor does not meet the key requirements; b) the mechanism used in copying source data is important; and c) with electronic systems there is a need to designate the location of the source data.

Single Copy

A single copy of the source data located at the sponsor organization has been shown not to meet the regulatory requirements as they are phrased today simply because, from an investigator's perspective, the requirement to ensure that source data are accurate cannot be met. This is because such source data can be modified without the investigator's approval and thus would be inaccurate in the eyes of the investigator.

Copying Source Data

Within the analysis above, the issue of copying source data arose. Two issues are raised when source data are copied. How do we ensure that the copied entity is an accurate copy of the original, and can the copy take on the role of source data?

The FDA's guidance document "Computerized Systems used in Clinical Trials" defines a certified copy as *"a copy of original information that has been verified, as indicated by dated signature, as an exact copy having all of the same attributes and information as the original."*[14]

With an automated electronic copy there is no ability to apply an individual's signature to the copy of the data, as there is no individual initiating the operation. Bar the signature requirement, an electronic system can meet the requirements specified in that it can make an exact copy with all of the same attributes and information as the original. It must be ensured that the process is reliable and accurate. In theory, one could manually review the copies (the same as with photocopies of paper) and then e-sign them. However, this would be problematic if not impossible to achieve given the number of copy operations that take place and their location (e.g. an electronic diary located with a subject). Subsequent software reading the copies would need to authenticate the signature, again, problematic. A validated copy process should be able to be relied on to prove that copies are accurate and complete. However, this means that the copy process should be pre-verified to operate correctly under the variety of conditions that may be encountered. The copy must retain all the components of the original, including any associated metadata and any changes made to it.

The copy operation raises the issue as to which is the source data. The obvious answer is the original, but, given that the two items are the same, it may be desirable to consider the copy as

[14] Computerized Systems used in Clinical Trials, April 1999

the source. An example may be when source data are copied from an eDiary device to a machine located at an investigator site. In this circumstance, it may be desirable to consider the new copy on the investigator machine to be the source data that is to be maintained. If this was accepted as a method, then such ideas would need to be documented and the transition of location of the source data would need to be visible to all concerned.

This idea of moving source data is connected with the idea of the Transitory Data Collector that has been proposed by some within the industry. This concept proposes that data collected on a device, but destined for a central server, not be considered source data since the period of time that the data are on the device is finite[15]. The source data would be that stored on the central server. However, the data, while on the device, is the only copy of the source data and should be considered as such. There is little difference between the concepts of designating the source data or considering it transitory. What is different is the emphasis placed on the data when it is considered source data and the controls that need to be in place while it is. Assurance is needed in both cases that the source data are copied or transmitted accurately without error. And, if the source data can be changed or deleted during the brief time it is stored in the original collection mechanism, then an audit trail should be in place, and audit trail entries copied to the server with the data.

Requirement 11: The location of source documents and the associated source data shall be clearly identified at all points within the capture process.

Requirement 12: When source data are copied, the process used shall ensure that the copy is an exact copy preserving all of the data and metadata of the original.

Considerations for Statistical Analysis and eSource

As with any data being collected and reviewed during the conduct of a clinical trial, data captured by electronic means, including ePRO, must only be viewed within the protocol defined requirements for blinding of the study data. The rapid collection and availability of data collected using electronic means does NOT give license to review unblinded trial data (i.e. with any knowledge of treatment groups). The blinding of study treatment codes is an essential part of the scientific integrity of many clinical trials since it serves to greatly reduce or eliminate bias in the evaluation of treatment groups. Please refer to FDA Guidance for Clinical Trial Sponsors (Draft Guidance - 2001; Section 4.2) and ICH E9 - Statistical Principles in Clinical Trials (Section 4.5).

[15] The data are held for the period of time while a connection is made from the device to the central server, the data are transferred and reception of the data confirmed by the server. This period may be very short down to seconds or minutes but could also be lengthy, hours or days.

User Requirements and Definitions

The requirements developed within this document place certain interpretations on key words and phrases and require some shifts in thinking. This summary provides the associated explanatory notes and definitions that accompany the user requirements. The notes are written in a technology-independent manner so as to be applicable to paper or electronic systems and processes.

Note: A mapping between the user requirements and the regulatory text appears in Appendix 5.

1. **An instrument used to capture source data shall ensure that the data are captured as specified within the protocol.**

   Any instrument used, be it a paper form or an electronic method needs to be verified against the requirements (the clinical protocol) to ensure the correct data are being captured and that the investigator or subject is not being influenced or biased when they respond.[16]

2. **Source data shall be Accurate, Legible, Contemporaneous, Original, Attributable, Complete and Consistent.**

   The process and tools must include features and controls to enable the collection of data that meets the necessary levels of data quality and integrity. In particular:

   a) Accurate: The data captured shall be accurate and the reporting of such data should be accurate.

   b) Legible: Data must be held such that, when retrieved, it can be read and understood. This includes not only storing the data such that it can be retrieved, but also storing any metadata such that the meaning of the data is clear.

   c) Contemporaneous: Data are recorded as soon as possible after the event to which it refers.

   d) Original: The data should be the original data and not falsified.

   e) Attributable: Data should be attributable to the individual, both to the subject being reported on, and those who have modified that data.

   f) Complete: The data must be whole, an entire set.

   g) Consistent: The data must be self-consistent and free from self-contradiction.

[16] Reference is used only in relation to subjects being influenced. However it was pointed out that investigators can be influenced in similar ways. Since draft 0.5 of this document was released the FDA have released a draft guidance on Patient Reported Outcomes.

3. An audit trail shall be maintained as part of the source documents for the original creation and subsequent modification of all source data.

The maintenance of an audit trail is important to ensure that the quality of the data is maintained and that changes to the source data are approved and traceable.

The term "source document" is a term very much related to the world of paper and the term is deeply embodied within the ICH GCP guidance document. Within the electronic world, the term should encompass a logical collection of source data.

The audit trail would incorporate the date and time of the change, the identity of the individual making changes, the action being undertaken, the old and new data values and the reason for change.

DEFINITION	
Source Document	The mechanism used to bind together a logical collection of source data items.

4. The storage of source documents shall provide for their ready retrieval.

Source documents should always be available to authorized individuals to meet their regulatory obligations. Ease of use, be it of a manual processes or an electronic system, is a key factor in allowing access to source data. Processes or systems that are difficult to use may make it difficult to locate the desired data.

5. The investigator shall maintain the original source document or a certified copy.

The principle behind this requirement is that the investigator controls the source document or a certified copy, thus ensuring protection against unauthorized changes to the data once it has been passed to another party.

DEFINITION	
Maintain.	The action of capturing, recording, amending and storing source documents

10. The sponsor shall not have exclusive control of a source document.

This requirement is associated with requirement number 5 and is important in ensuring that, at no point in time, is there only a single copy of the data that is only under the control

of a sponsor. This protects against source data being modified in circumstances where it should not.

It should be noted that this does not preclude such circumstances as phase 1 units operating within sponsor organizations. These trials still require individuals operating in the roles of sponsor and investigator and their responsibilities are as per other trials, i.e. the investigator must still retain control over the source data.

DEFINITION	
Control	The ability to decide when source data are created, amended, viewed or copied.

6. Source data shall only be modified with the knowledge or approval of the investigator.

The investigator is responsible for the source data held within source documents. The data should only be modified with the investigator's approval.

7. Source documents and data shall be protected from destruction.

Source documents must never be destroyed during the period within which they must be retained under the regulations. However, a copy can be (see below), and this is an important concept in that, in an electronic world, there may be a case for copying a record, designating the new copy as the source and removing the original[17]. In this circumstance, extreme care should be taken to ensure the new copy is available prior to the deletion of the original.

8. The source document shall allow for accurate copies to be made.

The need for accuracy when copying source data cannot be over emphasized. Once an error has been introduced it will propagate down the chain. Copies need to be made for examination by authorized parties, for example, regulatory authorities but also when source data are to be migrated, see 11 below.

9. Source documents shall be protected against unauthorized access.

Source documents need to be protected so as to maintain subject confidentiality and to prevent unauthorized persons modifying the data.

11. The location of source documents and the associated source data shall be clearly identified at all points within the capture process.

[17] For example, an eDiary where the data may be copied from the device to a PC located at the site or to a central server.

It should be a requirement that if data are to be copied, then the locations of source data and the points when such data are copied are well documented and understood. There is only ever one source, therefore when source data are copied, it must be well understood as to which is considered the source. There should only be a single source such that it is clear what is under the control of the investigator. It is this single copy that is used in ensuring that the data submitted to a regulatory agency is the data collected by an investigator or derived from such.

12. When source data are copied, the process used shall ensure that the copy is an exact copy preserving all of the data and metadata of the original.

When source data are captured on paper it is possible to copy the paper documents, verify that no information has been lost and consider the copy the source document.[18] Within the electronic world we wish to have the same concept but two issues place barriers in our way: a) the intangible nature of the data in that it resides unseen within a computer and b) the copy process.[19]

Therefore it is recommended that the definition for Certified Copy be revised. The following is being reviewed by the CDISC Glossary Group and will be modified in accordance with their recommendations.

DEFINITION	
Certified Copy	A copy of original information that has been verified as having the same metadata and data as the original. The copy may be verified by dated signature or by a validated electronic process.[20]

[18] See ICH GCP 1.51 and 1.52 and the definition of certified copy within the FDA Guidance Document Computerized Systems used in Clinical Trials

[19] It was noted during review that not "all" the properties of source data captured on paper are typically copied onto a paper copy. Some information is added (different ink, new paper, contrast, size, format, etc.)

[20] The FDA, in the withdrawn Guidance for Industry 21 CFR Part 11; Electronic Records; Electronic Signatures Maintenance of Electronic Records, introduced the concept of 'accurate' and 'complete' copies. Although this guidance was withdrawn, there is some useful information in this guidance.

Scenarios

The analysis presented within this document examines the current regulatory framework, develops a set of user requirements and assesses those requirements against current industry practice. This assessment of current practices has pinpointed certain areas where, if a very strict interpretation of the regulations is taken, it could be argued that some solutions may not meet all of the current regulatory requirements. However, in a time of transition, there is a need to reflect upon the spirit of the regulations (and to keep in mind that some of these regulations were created for paper based documentation only) rather than using a literal interpretation. This view is necessary to adapt to the current environment and thus gain the benefit of new technology, while maintaining the necessary measures to ensure that clinical trial data continues to be of the highest quality and integrity.

Based on the analysis presented, it is recommended that those implementing processes or systems used to capture source data as part of a clinical trial assess the processes and systems against the key requirements identified within this paper. The user requirements have been designed to be independent of a given technology and reflect the needs of the regulatory requirements found in both ICH GCP and the FDA predicate rules[21].

Before describing specific scenarios in detail, it should be recognized that, as technology advances at differing rates across various domains (patient diary/eDiary, EDC, EHR, psychometric tools, internet based methods, personal health records, regional health records, phone and cell-phone based data collection, central labs, sponsor based labs, new approaches to instrumentation, etc.), any **given** trial could end up utilizing **multiple** techniques (sometimes called mixed mode) in an orchestrated way to elaborate the overall amalgamated data environment and/or to leverage the best technology for different portions of the process for a given clinical trial. In that light, it has been suggested that the eSDI group take a decidedly more **permissive**, yet **managed** approach to allow for a menu of options for the integration of new technologies in clinical trials. Such an approach provides for creativity and the use of state-of-the art technologies along with more traditional or known technologies. Presumably, clinical trials employing multiple technologies leverage data interchange standards to facilitate data flow from source to reporting.

It is highly recommended that sponsors clearly document[22] the process they are following for data flow, retention, access and archive to clearly delineate how their specific process adheres to the 12 requirements in this document, including authority and all appropriate regulations. This should be completed for each approach used and be made available to regulators at appropriate times. A standard form would be useful for documenting the process and adherence to the user requirements.

The eSDI working group has developed five scenarios that the group believes will permit all stakeholders to deploy new technology for the capture of eSource data within the spirit of the

[21] It should be remembered that ICH GCP is a guidance document within the US regulatory framework. However, within the European Union, ICH GCP is referenced from the EU Clinical Trials Directive 2001/20/EC and from the new GCP Directive 2005/28/EC.

[22] This could take the form of process map, data flow diagram, system and process diagram etc.

existing regulations while ensuring the necessary level of control and ensuring data quality and integrity, thus providing the public at large confidence in the drug development process. In addition, it is felt that these scenarios are forward-thinking and can help pave the way for utilization of electronic health records for clinical research in the future, to facilitate interoperability between clinical research and healthcare systems and information sharing between these two patient-focused arenas. These are certainly not the only possible scenarios for implementing eSource trials; there are alternate combinations and additional scenarios that will meet the user requirements. For each trial conducted, the scenario used and the processes put in place should be reviewed for adherence to the 12 requirements in this document and applicable regulations and predicate rules.

Keeping the aforementioned general considerations in mind, five potential solutions for employing eSource data technologies within the context of the existing regulations are described in more detail in the rest of this section. These are:

1. Source at Site
2. eSource System Provider (Contracted Supplier)
3. Single Source Concept
4. Extraction and Investigator Verification (Electronic Health Record Data)
5. Direct Extraction from Electronic Health Records

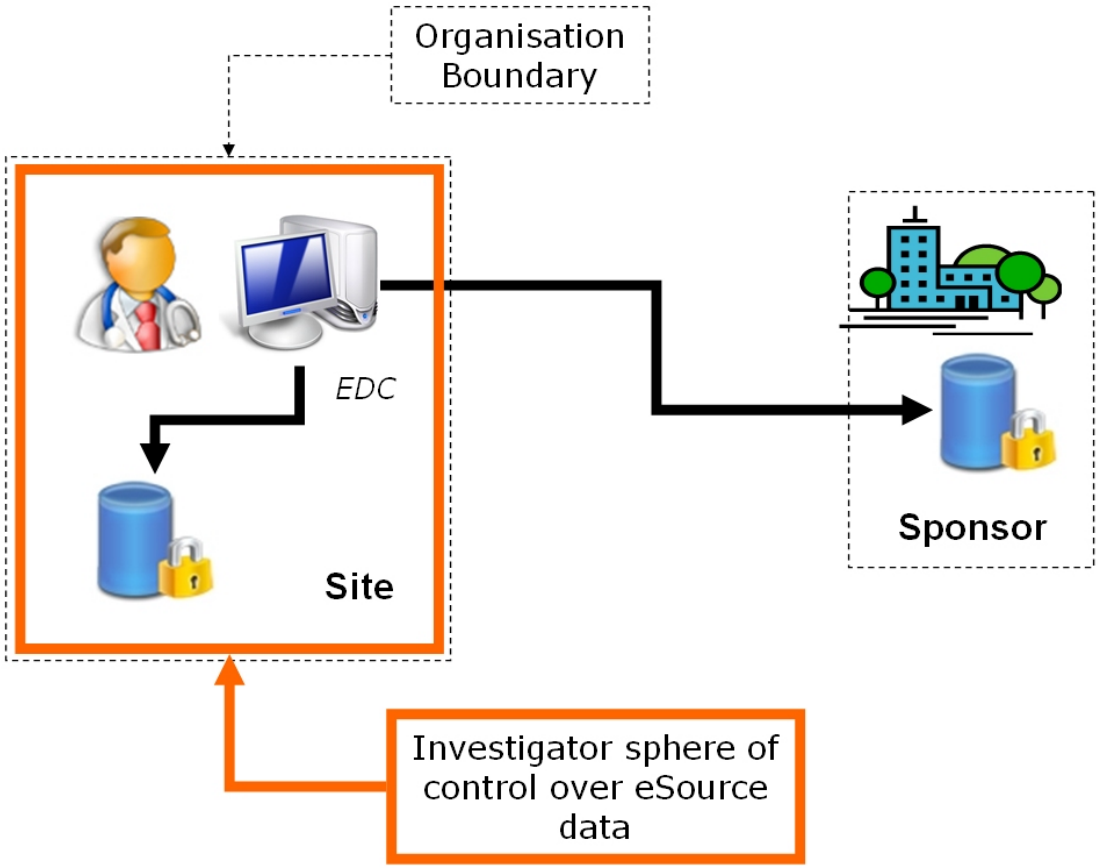
The diagrams within the following sections illustrate the scenarios and the flow of source data as described by the scenario. In particular, the diagrams indicate the investigator “sphere of control”, the control that the investigator needs to exercise over source data to meet the User Requirements, and the physical bounds of the organizations involved..

Source at Site

The first scenario is the most straight-forward, whereby source data are maintained at investigative sites under the direct control of the investigator; the sponsor not having any access. Such a solution mirrors that of the paper world but allows for the benefits of electronic technology to be leveraged.

In this scenario, the data from the eSource technology (e.g. eDiary, eCase Report Form or eData Collection Instrument) are sent directly to the principal investigator/investigative site. There can be a simultaneous feed to the trial sponsor of the specific clinical trial data (i.e. without data that are strictly for the site to retain, such as patient contact information). Alternatively, the trial-specific data can be transferred subsequently to the sponsor.

To store the data, it is recommended that the CDISC Operational Data Model (ODM) be employed for reasons given in the Benefits section of this Scenario. The data collection application can be set up as a means of storage for the investigative site and it is anticipated that commercial tools will also be available for this purpose in the future. Alternate electronic storage mechanisms can be deployed as long as they meet the requirements set forth in this document and adhere to 21 CFR11 record retention requirements. The figure below depicts this scenario.



This scenario leaves the investigator in direct control of the data and also creates two repositories (one at the site and one at the sponsor location) that can be compared at a later stage thus ensuring integrity of the data.

This scenario clearly satisfies **Requirement 10** since *the sponsor does not have exclusive control over the source document.*

This scenario also fulfils requirement 11, as follows, and 1-9, 12 if the system and processes are set up properly. This adherence should be documented for the particular processes employed in each clinical trial conducted using this scenario.

Requirement 11: The location of source documents and the associated source data at all points within the capture process shall be clearly identified.

The source documents and associate source data are at the site in this scenario.

As stated previously, sponsors should document how the processes they are following in this scenario for data flow, retention, access and archive adhere to the other requirements in this document, including authority and all appropriate regulations.

Benefits of this Approach and the Value of Standards:

1. The data are provided to the investigator at the same time or earlier than the sponsor, hence the investigator can address any safety issues promptly.
2. The sponsor can demonstrate that they did not change the data without investigator knowledge and approval. The investigator can have primary control of the data, while the sponsor retains a copy.
3. The CDISC ODM can be used to store and eventually archive electronic data at investigative sites with a standard format. This means of storage is vendor-neutral and platform-independent and does not require that the system be retained for future years in order to access the data, along with audit trail. Auditors will be able to use standard review tools.
4. The use of the ODM for storage/archive includes retention of the data management environment, edit checks and audit trail.
5. The archived trial data, complete with edit checks and audit trail, can be reviewed at a later date using off-the-shelf tools[23].

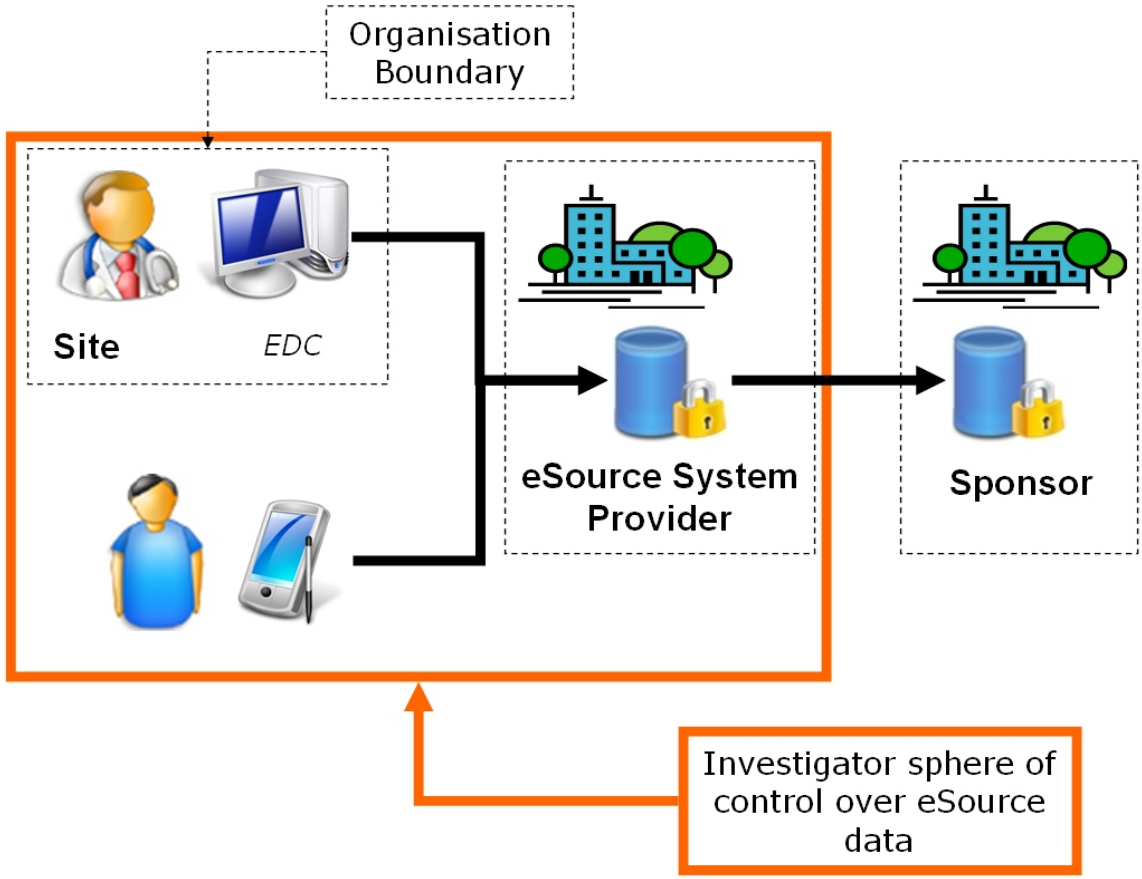
[23] One comment mentioned the use of PDF formats. ODM has the advantage of being based on XML technology and as a result is machine readable.

eSource System Provider (Contracted Supplier)

This second scenario has also been called a "trusted third party" or an ASP (application solution provider). It has frequently been employed in trials using electronic data capture (EDC), which may or may not have included eSource data collection: however, for this reason, the eSDI group felt it important to address this scenario and how it can be implemented to meet the user requirements. The primary issue, which has been raised from the regulatory perspective, with this solution in current practice is determining the accountability for the data integrity, specifically which party is responsible-- the investigator, the sponsor or the vendor. In practice today, a vendor hosts the data, but the sponsor has a contractual agreement with and pays the vendor. Per the regulations, the investigator should be (but may not be, in practice) in control of that data when the vendor is hosting that data. This is dependent on how the access controls and processes are organized. Questions arise, such as: Where does the investigator stand if the vendor company ceases operation? Where and how does the FDA auditor then access the data for an audit? These are just a couple of the many questions and regulatory concerns that arise with this scenario. Through discussions among the eSDI group members and FDA representatives, a viable solution that will meet the 12 requirements and adhere to existing regulations was sought and is offered herein.

It is common practice for the sponsor to audit the vendor against the existing regulations and to ensure that the data repository system is validated, before the sponsor contracts a trial to the vendor. In most sponsor-vendor agreements, the program code for the application is placed in escrow so as to protect the sponsor should anything happen to the vendor[24]. In addition, the sponsor ensures that there is appropriate back-up of the data that the vendor hosts and that there are processes in place such that the investigator can have continuous access to the data and be in control of the data, even if they are remotely hosted by the vendor. The figure below depicts this scenario and the appropriate sphere of control of the data.

[24] One comment received noted that not only may the software be required but the environment and personnel to operate it



This scenario can meet **Requirement 10**, if set up such that *the sponsor does not have exclusive control over the source document. Rather the investigator should have the appropriate controls such that changes can only be made with the knowledge and approval of the investigator..*

This scenario can also fulfill requirement 11, as follows

Requirement 11: The location of source documents and the associated source data at all points within the capture process shall be clearly identified.

Although the source data are not located at the site in this scenario, the investigator must have appropriate and ready access to that data and the control of the data content. The location of that data and the processes by which it is accessed, changed and protected should be clearly documented.

This scenario can fulfill the other Requirements (1-9, 12) if the system and processes are set up properly. This adherence should be documented for the particular processes employed in each clinical trial conducted using this scenario. As stated previously, sponsors should document how the processes they are following in this scenario for data flow, retention, access and archive adhere to the other requirements in this document, including authority and all appropriate regulations.

The proposed solution to ensure that the requirements are met in this scenario is that, prior to contracting with the vendor, the sponsor should undertake an evaluation of the vendor, the

processes and the system to be used to ensure that it meets the user requirements defined within this document and, of course, the existing regulations. This evaluation would be documented within a Source Data Evaluation Report, and this report would then be made available to an FDA auditor or other regulatory authority should the need arise. Alternatively, the FDA or regulatory authority must be allowed to audit the vendor.

It is important for each Investigator to be informed of the accountabilities and processes that are in place for the trial thus ensuring the system is validated and that the proper procedures are in place such that the investigator has appropriate control of the data for the subjects at their site[25]. Again, the requirements identified in this document must be met by the vendor and documented in the sponsor's report. It is recommended that the following steps be taken for this solution.

1. Vendors must agree that they could be evaluated by the Sponsor (and/or FDA or other regulatory authority) against the user requirements identified within this document.
2. Sponsors should evaluate/audit the vendor to ensure that their systems can comply with the user requirements detailed within this document.
3. Sponsors should evaluate/audit (and document this evaluation) the system, as well as the processes the vendor and investigator and sponsor are to follow for data flow, retention, access and archive ensuring adherence to the requirements in this document, including authority and all appropriate regulations. The documentary evidence would be made available to the FDA or other regulatory authority on request.[26]
4. Sponsors must explain to the Investigators in the trial that they (the Investigators) have responsibility for the data and should have ready access to the data and audit trail and the other requirements and responsibilities associated with source data. See Appendix 10 – Good Practices Checklist: Investigator Responsibilities for further information.
5. Investigators should understand the systems being provided to them, the source data controls that are in place and how they adhere to appropriate regulations.
6. Sponsors should show due diligence.

Benefits of this Approach and the Value of Standards:

1. The data are presumably hosted in a secure environment, with proper processes and procedures in place to ensure that the investigator has control of and access to the data for subjects at the clinical site.
2. The CDISC ODM can be used to store and eventually archive electronic data at investigative sites with a standard format. This means of storage is vendor-neutral and platform-independent and does not require that the system be retained for future years in

[25] The investigator may not delegate tasks, nor contract them. Note that the sponsor can contract sponsor responsibilities to a CRO (allowed by regulation, making the CRO liable for compliance with the contracted sponsor obligations) but the sponsor cannot contract investigator responsibilities.

[26] Generally, the FDA only examines QA summaries or reports at Sponsors if there is a problem or an issue with the data or the system as identified in the submission.

order to access the data, along with audit trail. Auditors will be able to use standard review tools.

3. The use of the ODM for storage/archive includes retention of the data management environment, edit checks and audit trail.
4. The archived trial data, complete with edit checks and audit trail, can be reviewed at a later date using off-the-shelf tools[27].
5. This method meets the spirit of the regulations and the accountable parties are identified as with the paper environment.

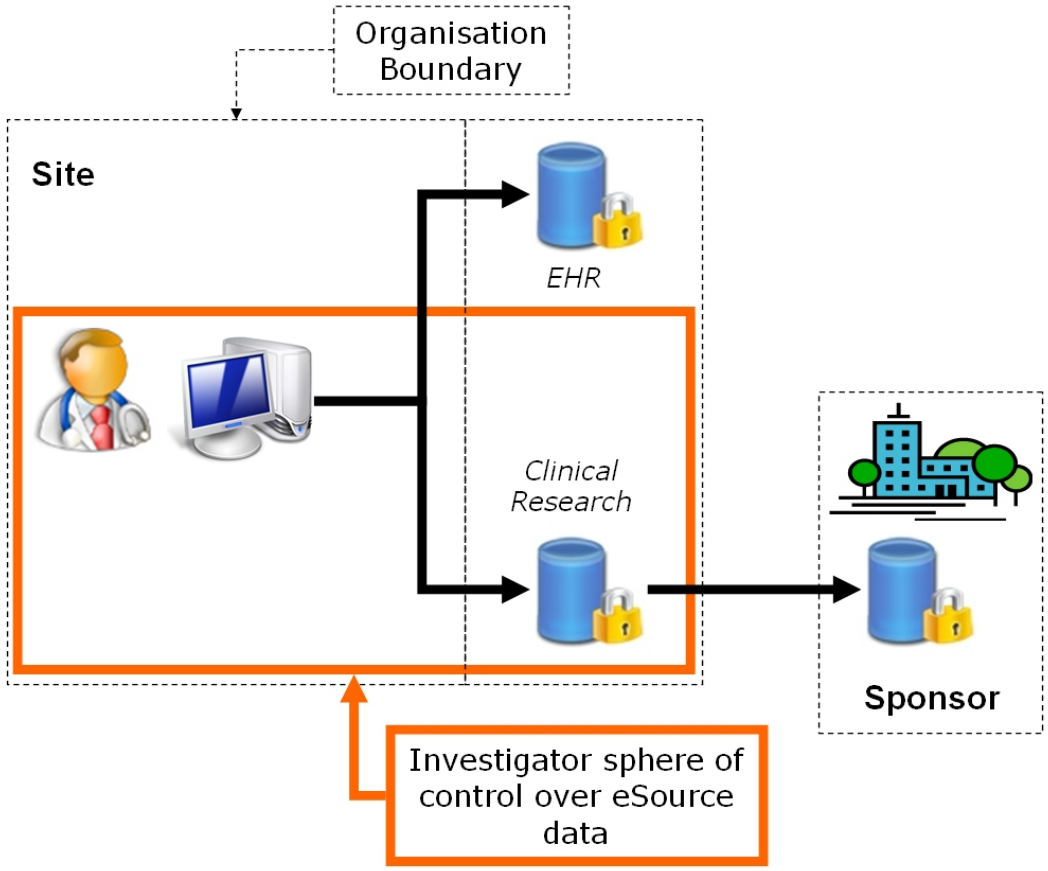
[27] One comment mentioned the use of PDF formats. ODM has the advantage of being based on XML technology and as a result is machine readable.

Single Source Concept

The third scenario is called the Single Source Concept. It is a solution that was developed and implemented (initially as a proof-of-concept) by CDISC and partners; the concept is to leverage healthcare and clinical research standards in the simultaneous population of an electronic healthcare record and a clinical trial database while adhering to existing applicable regulations for both clinical trials and healthcare. This scenario is not the ideal future methodology to facilitate clinical research by investigators; however, it does offer a viable means for data to be entered just once for multiple purposes (research, patient care, safety surveillance, etc.) within the context of existing regulations. This would presumably facilitate the processes at an investigative site and also eliminates data transcription, which is a point of potential error introduction.

In this scenario, data are entered into an electronic source document at the site (typically as an interface to the electronic health record (EHR) system but conceivably as an interface to an EDC system). All of the eSource data can flow into the EHR database, while the clinical trial data (as identified by the protocol) can be simultaneously passed into eSource repository and passed onwards to the clinical trial database. The proof-of-concept for Single Source was developed using Health Level Seven (HL7) standards for the data flow into the EHR system and the CDISC operational data model (ODM) for data flow into the clinical data management system (CDMS).

The following figure depicts a Single Source scenario.



In this particular example, the investigator enters eSource data into an EHR that has been configured to collect data for that particular trial, e.g. by accessing an eSource document or eCase report form that identifies fields required for the trial and has been integrated into the work flow for the site. The clinic notes and health record information flows in the EHR can be created using HL7 standards, if applicable for the system. The trial-specific data are simultaneously passed into the sponsor database and into a separate source data repository or data store at the site to ensure adherence to the requirements and regulations.

This scenario can meet requirement 10 as follows:

Requirement 10: The sponsor must not have exclusive control over the source document**.**

In the Single Source process, the eSource data are controlled through source data (Clinical Research) repository at the site by the investigator. The processes that are set up for the trial must ensure that the investigator has the full read, write and change access to the data. As for any trial, if changes are made in the sponsor CDMS, the investigator must be aware of these and approve them.

This scenario can also fulfil requirement 11, as follows

Requirement 11: The location of source documents and the associated source data at all points within the capture process shall be clearly identified.

The source data are located at the site in this scenario, in the source data (Clinical Research) repository. These data can also (simultaneously) be retained in the EHR as the patient record. Having a clinical-trial specific data storage mechanism at the site that is separate from the EHR can also help ensure appropriate archive of these data and availability in the case of an audit.

This scenario can fulfill the other Requirements (1-9, 12) if the system and processes are set up properly. This adherence should be documented for the particular processes employed in each clinical trial conducted using this scenario. As stated previously, sponsors should document how the processes they are following in this scenario for data flow, retention, access and archive adhere to the other requirements in this document, including authority and all appropriate regulations.

Benefits of this Approach and the Value of Standards:

1. Due to this process, where data are input (not extracted from) into the EHR system, the environment can be readily controlled in a manner that is compliant with existing regulations and the requirements in this document, and the input tool can meet 21 CFR Part 11 regulations and predicate rules without having to audit all systems coexisting with the EHR at the site, e.g. billing systems.
2. Single source facilitates investigator/site processes since data are entered once and utilized multiple times.
3. The single source process eliminates transcription and the need for source data verification, as long as the appropriate validation procedures are followed and documented to ensure data integrity.

4. Single source can be undertaken without an EHR system, if the data are entered using an eSource document within an EDC or eDiary system and the eSource data are controlled, stored and archived at the site while the clinical trial data subset (per protocol) are sent to the EDC repository and/or the clinical database at the Sponsor or CRO.
5. The clinical trial data can potentially be extracted into an EDC system or a database that subsequently populates a sponsor CDMS, which can then be used to perform edit checks. [The eSource data that are confidential to the site (e.g. patient contact information) should be stored only at the site and would not go into the sponsor clinical trial database.]
6. Due to the reliance of standards for data collection, any CDISC-HL7 standards-based EHR and/or EDC can be employed in this process and the data should be exchangeable, particularly if terminology is agreed initially.
7. Data entered as part of a clinical trial are captured to a repository under the control of the investigator. Such data could also, if desired, be retained within the EHR system. However the source data repository is considered the source data. It is recommended that this data store be in CDISC ODM format.
8. The CDISC ODM can be used to store and eventually archive electronic data at investigative sites with a standard format. This means of storage is vendor-neutral and platform-independent and does not require that the system be retained for future years in order to access the data, along with audit trail. Auditors will be able to use standard review tools.
9. The use of the ODM for storage/archive includes retention of the data management environment, edit checks and audit trail.
10. The archived trial data, complete with edit checks and audit trail, can be reviewed at a later date using off-the-shelf tools[28].
11. The SDTM can be used to identify the majority of the fields for the eSource document such that data are collected in a format consistent with how they should be submitted if these data will be part of a regulatory submission.
12. This scenario creates databases with the clinical trial specific data at both sponsor and site, enabling source data verification.

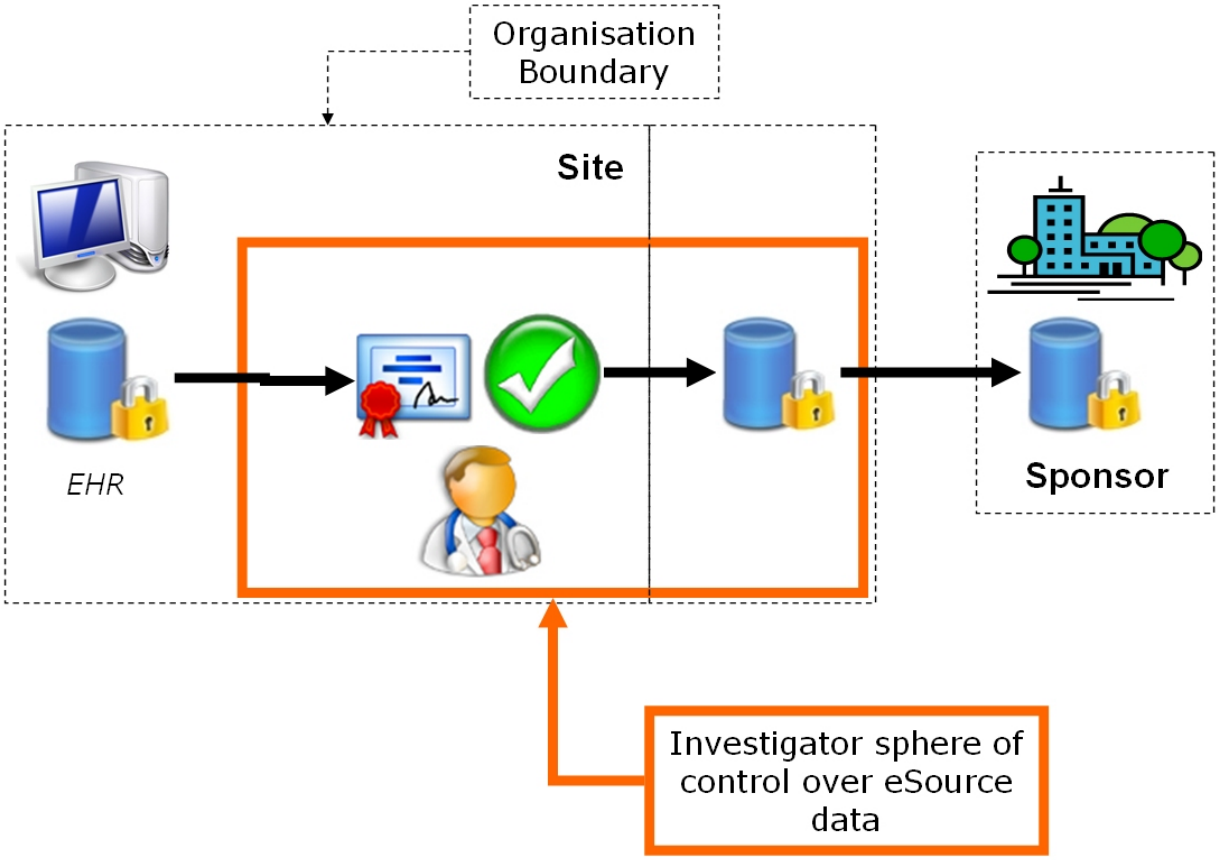
[28] One comment mentioned the use of PDF formats. ODM has the advantage of being based on XML technology and as a result is machine readable.

Extraction and Investigator Verification (Electronic Health Records)

The fourth scenario offers another approach to the use of electronic health records for clinical research while adhering to today's regulations for clinical trials. There are expressed goals for the direct use of EHR for clinical research, including to: a) facilitate clinical research for sites/investigators by enabling the entry of data once for research and healthcare; b) reduce the number of duplicate samples taken from subjects of trials who also are patients receiving healthcare; c) maximize the use of information from healthcare for the research benefit of the population as a whole and others. However, the direct extraction of EHR data for clinical research, without the process steps described in this scenario, will only be compliant with the existing regulations if the EHR can be validated in compliance with 21 CFR Part 11. (See the following fifth scenario – Direct Extraction from Electronic Health Records.) Until such time that HIPAA and 21 CFR Part 11 and related regulations are more closely aligned and can serve the needs of FDA for monitoring of trials on drugs still in development, an additional process step can be added as an interim step to allow a more direct use of EHR in clinical trials without the necessity to validate the entire EHR system. This step is for the investigator to verify that the extracted data, for clinical research use, accurately reflect the source data for that subject before it is included as part of the clinical trial data record. The following principles and process steps would apply.

1. 21 CFR Part 11 starts at the point of the creation of a clinical research record (i.e. it is not necessary to apply these regulations to the medical records at the site).
2. If an investigator requests the information and uses it as a regulated clinical research document, it is then subject to the regulations for clinical trials, including 21CFR11 and other related guidance.
3. If data are extracted from an EHR into a clinical research record, there is a need to ensure that the migration of the data from EHR to the clinical record is validated (is accurate, was not changed in the extraction process and that patient confidentiality requirements are met); the investigator should verify that the extracted data accurately reflect the patient's source data. (This can be done with an electronic signature.). The clinical research record is created at the point of signing.
4. New data for the trial can be entered into eSource documents and sent to the EHR, thus eliminating the need for the additional validation step for those data. However, the investigator must still 'sign-off' as usual on data entered for clinical trials, whether these are entered as eCRF or eSource data. The new data, specific to the clinical trial, could also potentially be entered into an EDC system and then sent to the EHR.
5. The clinical trial data can potentially be extracted into an EDC system or a database that subsequently populates a sponsor clinical data management system (CDMS) or clinical research database, which can then be used to perform edit checks. [The eSource data that are confidential to the site (e.g. patient contact information) should be stored only at the site and would not go into the sponsor clinical research database or CDMS.]

The following figure depicts the fourth scenario for extraction of data from electronic health records with verification by the investigator.



This scenario can meet requirement 10 as follows:

Requirement 10: The sponsor must not have exclusive control over the source document**.**

The eSource data in this process are extracted from the electronic health record into a second system holding the clinical research record. This then becomes the eSource and therefore needs to be under the control of the investigator. In this scenario, the investigator must review and approve the data that are extracted as eSource. These data may then go into a clinical data management system (CDMS), EDC data management system or other system that may be used for edit checks or other data management processes. As in all scenarios, the investigator must be aware of and approve any subsequent changes that are made in the source data (which should also be reflected in the EHR system).

This scenario can also fulfill requirement 11, as follows

Requirement 11: The location of source documents and the associated source data at all points within the capture process shall be clearly identified.

The source data are located within the system holding the clinical research record, not the EHR. The data can also potentially be retained in a clinical-trial specific data storage mechanism at the site that is separate from the EHR, which can help ensure appropriate archive of these data and availability in the case of an audit.

This scenario can fulfill the other Requirements (1-9, 12) if the system and processes are set up properly. This adherence should be documented for the particular processes employed in each

clinical trial conducted using this scenario. As stated previously, sponsors should document how the processes they are following in this scenario for data flow, retention, access and archive adhere to the other requirements in this document, including authority and all appropriate regulations.

Benefits of this Approach and the Value of Standards:

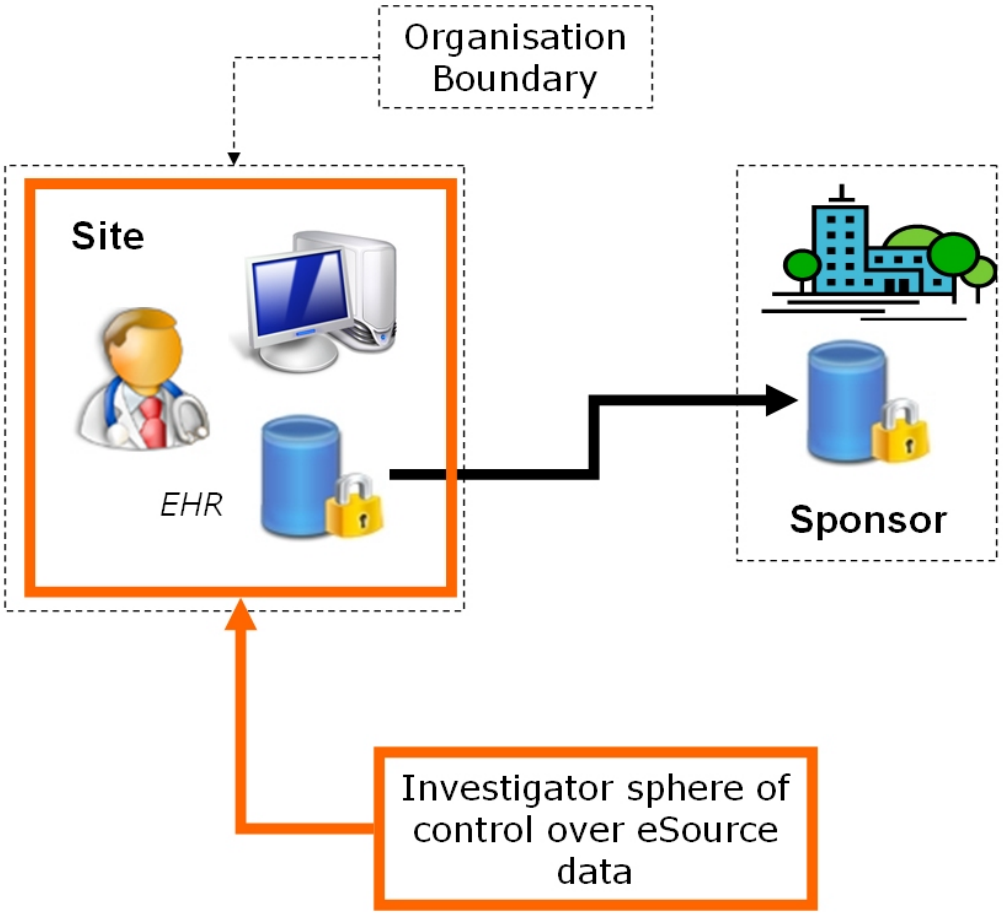
1. Source data are stored in the system holding the clinical research record. These data may be on ODM format and/or they could also be placed into an ODM repository as they are extracted into the clinical research record. This ODM repository can be archived at the site.
2. Extraction of data that are already available in an EHR at the site for use in clinical trials facilitates investigator/site processes since data are entered once and utilized again without re-entry.
3. The direct extraction eliminates transcription and the need for source data verification, as long as the appropriate validation procedures are followed and documented to ensure data integrity from EHR to clinical trial record and vice versa. Systems could also implement automated notification of changes to relevant data within the EHR system.
4. The sponsor can demonstrate that they did not change the data without investigator knowledge and approval. The investigator can have primary control of the data.
5. The clinical trial data can be extracted into an EDC system or a database that subsequently populates a sponsor CDMS, which can then be used to perform edit checks. [The eSource data that are confidential to the site (e.g. patient contact information) should be stored only at the site and would not go into the sponsor clinical trial database.]
6. The CDISC ODM can be used to store and eventually archive electronic data at investigative sites with a standard format. This means of storage is vendor-neutral and platform-independent and does not require that the system be retained for future years in order to access the data, along with audit trail. Auditors will be able to use standard review tools.
7. The use of the ODM for storage/archive includes retention of the data management environment, edit checks and audit trail.
8. Archived trial data in ODM is complete with edit checks and audit trail, which can be readily reviewed at a later date using off-the-shelf tools[29].
9. The SDTM can be used to identify the majority of the fields for the eSource document such that data are collected in a format consistent with how they should be submitted if these data will be part of a regulatory submission.
10. This scenario creates databases with the clinical trial specific data at both sponsor and site, enabling source data verification.

[29] One comment mentioned the use of PDF formats. ODM has the advantage of being based on XML technology and as a result is machine readable.

Direct Extraction from Electronic Health Records

This last scenario is included to address the interests of those who would like to implement direct extraction of data from electronic health records. It is recognized within FDA that, at some point in the future, there should be a more applicable set of regulations or regulatory guidance to facilitate this process and that these regulations should mesh better with the HIPAA regulations that are in place for EHR. However, in the context of the existing regulations (which is a requirement for the scenarios included in this document), the direct extraction of data from electronic health records for use in clinical trials requires that the EHR tool/application meet the requirements of the predicate rules and, due to the electronic nature of record storage, 21 CFR Part 11. This may become rather onerous if the EHR system is open and interfaces with other systems that may include such data as that used for billing, admissions, and insurance; the hospital will need to comply with the required validation process for the entire system. However, if an EHR can meet the existing regulations and the requirements detailed within this document (e.g. a stand-alone EHR application designed for clinical research), then it is acceptable to use that for clinical research purposes and to extract clinical research data from it.

The following figure depicts the direct extraction from electronic health records scenario.



This scenario clearly meets requirement 10 as follows:

Requirement 10: The sponsor must not have exclusive control over the source document**.**

The eSource data in this process are from the electronic health record, which is at the site and under the control of the investigator.

This scenario can also fulfill requirement 11, as follows

Requirement 11: The location of source documents and the associated source data at all points within the capture process shall be clearly identified.

The source data are located at the site in this scenario, in EHR system repository.

Depending on the processes set up for the study, the clinical trial data ca be extracted for the sponsor database and can also potentially be retained in a clinical-trial specific data storage mechanism at the site that is separate from the EHR, which can help ensure appropriate archive of these data and availability in the case of an audit.

This scenario can fulfill the other Requirements (1-9, 12) if the system and processes are set up properly. This adherence should be documented for the particular processes employed in each clinical trial conducted using this scenario. As stated previously, sponsors should document how the processes they are following in this scenario for data flow, retention, access and archive adhere to the other requirements in this document, including authority and all appropriate regulations.

Benefits of this Approach and the Value of Standards:

1. Source data are stored and archived in the EHR (which can be directly under the investigator/site control). These data could be placed into an ODM repository as they are extracted into the clinical trial record. This ODM repository can be archived at the site.
2. The sponsor can demonstrate that they did not change the data without investigator knowledge and approval. The investigator can have primary control of the data.
3. Extraction of data that are already available in an EHR at the site for use in clinical trials facilitates investigator/site processes since data are entered once and utilized again without re-entry.
4. The direct extraction eliminates transcription and the need for source data verification, as long as the appropriate validation procedures are followed and documented to ensure data integrity from EHR to clinical trial record and vice versa.
5. The clinical trial data can be extracted into an EDC system or a database that subsequently populates a sponsor CDMS, which can then be used to perform edit checks. [The eSource data that are confidential to the site (e.g. patient contact information) should be stored only at the site and would not go into the sponsor clinical trial database.]
6. The CDISC ODM can be used to store and eventually archive electronic data at investigative sites with a standard format. This means of storage is vendor-neutral and platform-independent and does not require that the system be retained for future years in

order to access the data, along with audit trail. Auditors will be able to use standard review tools.

7. The use of the ODM for storage/archive includes retention of the data management environment, edit checks and audit trail.
8. Archived trial data in ODM is complete with edit checks and audit trail, which can be readily reviewed at a later date using off-the-shelf tools[30].
9. This scenario can create databases with the clinical trial specific data at both sponsor and site, enabling source data verification.

[30] One comment mentioned the use of PDF formats. ODM has the advantage of being based on XML technology and as a result is machine readable.

Contributors

The following individuals contributed to the production of this document:

Rebecca Kush, CDISC (Group co-lead)
Dave Iberson-Hurst, Assero Limited (Group co-lead)

Ethan Basch, Memorial Sloan-Kettering Cancer Center
Peter Black, Scirex
David Detoro, Schering Plough
Hugh Donovan, Siemens
Greg Fromell, University of Pennsylvania
Ed Helton, SAS
John Jordan, Schering Plough
Suzanne Markel-Fox, GSK
Michael Noonan, Asthma Research
Lisa Olson, SEC Associates
Shaghig Palanjian, Perceptive Informatics
Jay Pearson, Merck & Co.
David Reasner, Sepracor
Dana Stone, Merck & Co.
Mark Weiner, University of Pennsylvania
Wallace Wormley, University of Pennsylvania

In addition, the following FDA Liaisons reviewed and commented upon the work presented:

Laurie Burke	Director, Study Endpoints and Label Development, Office of New Drugs CDER
Joanne Rhoads	Director, Division of Scientific Investigations, CDER
Joe Salewski	Deputy Director, Division of Scientific Investigations, CDER
Jane Scott	Study Endpoints and Label Development, Office of New Drugs, CDER
Steve Wilson	Deputy Director, Division of Biometrics II, CDER

Appendix 1 – Analysis of Paper Process

The analysis examines the data capture process that is based around a paper document and considers the life cycle of a single source document from the point when it is created or selected until that time when it can be destroyed. In order to reduce the process to its most generic form, the source document shall be considered to be a single page.

Note: A source document could be one, or more, pages each of which contain source data. Consideration may need to be given to logical collections of pages where some, but not all, contain source data items. In this situation, it would seem reasonable that the logical collection be considered a single source document.

Prior to becoming a source document, the page used to record an observation can be considered to be blank (empty), and as such, has no regulatory significance. That page could be part of the subject's medical record that already contains information none of which pertains to a clinical trial; it might be a blank CRF page waiting to be used as part of the trial. At some point, the investigator will commit the first source data item to the page. At that moment, the page becomes significant within the regulatory framework and becomes a source document. As a result, the source document needs to be managed for the duration of the trial and, upon completion of the trial, archived. At some future time, the regulations will no longer require its retention and the source document could be destroyed.

As such, a source document will pass through four states:

1. **BLANK**, where no data has been captured on the page, for example a CRF has been printed, but no data has been written on it. It may be a worksheet designed by the site for subsequent transcription, again with no data written on it. As such the page does not have significance within the regulations.
2. **POPULATED**, where the page has had one or more pertinent observations (source data items) written on it and the trial is still running. The page has thus become a source document.
3. **ARCHIVED**, where the trial has ended but the page needs to be retained under the regulations as a source document.
4. **OBSOLETE**, where the regulatory retention period is over and the page can be disposed of.

Within each of the states, those involved in the study will operate upon the source document in some way. For example, the initial capture of data would take a blank page and write (capture) some data on to it; this capture operation results in the page becoming a source document and the state of the item moving from **BLANK** to **POPULATED**. This process is applicable to both the capture of data by an investigator during a clinical encounter and by a subject as part of a Patient Reported Outcome.

The states and associated operations/events are listed in the following table, shown in the figure and described hereafter.

Note: The names of the operations/events detailed below have been selected so as to allow the operations to be referred to in the following analysis. They are not intended to be all-encompassing descriptions of the actions and process that would occur during a trial.

Current State	Operation/Event	New State
Does not exist	Select	BLANK
Does not exist	Create	BLANK
BLANK	Capture	POPULATED
POPULATED	Copy	POPULATED
POPULATED	Clarify	POPULATED
POPULATED	View / Inspect	POPULATED
POPULATED	Monitor	POPULATED
POPULATED	Store	POPULATED
POPULATED	Study Ends	ARCHIVED
ARCHIVED	View / Inspect	ARCHIVED
ARCHIVED	Store	ARCHIVED
ARCHIVED	End of Retention Period	OBSOLETE
OBSOLETE	Destroy	*No longer exists*

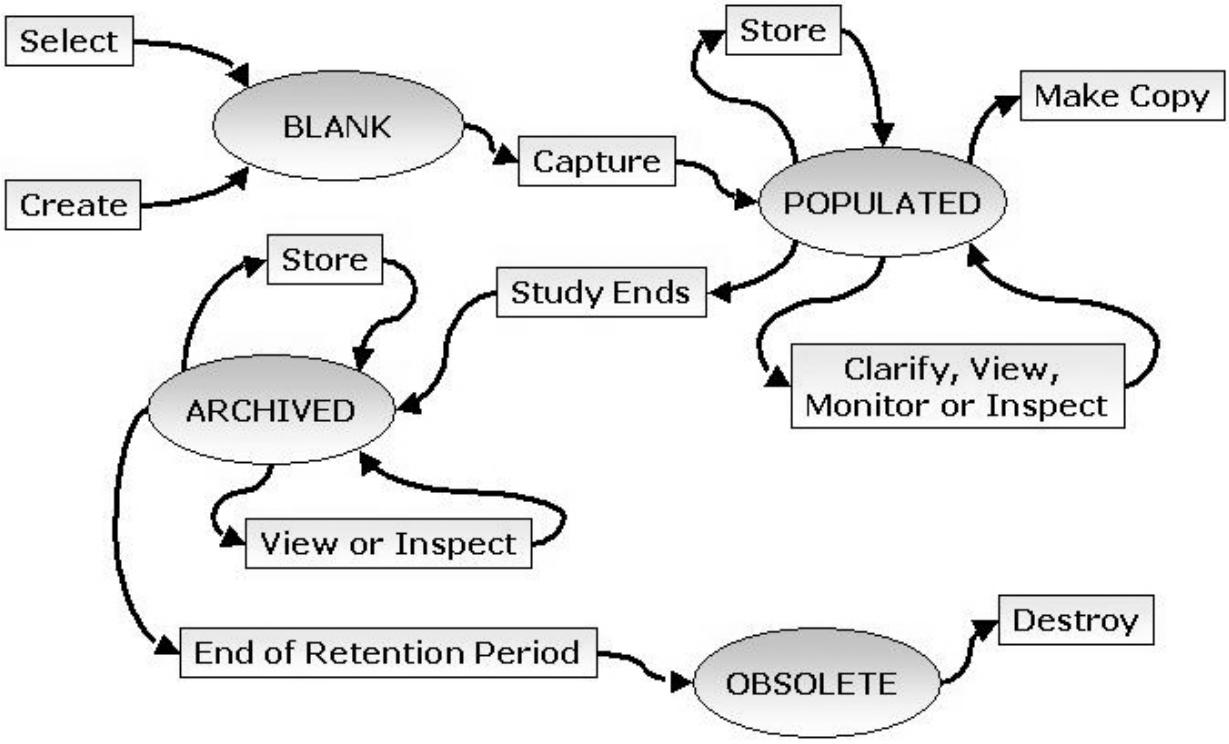


Figure 1 – State Diagram

The life cycle commences with one of two operations. A CRF page will be created as part of the process that takes the relevant information from the clinical protocol and uses it to **create** the CRF design, the resulting design being checked to ensure that it conforms to the protocol. The process for a diary page will be the same except that the psychometric aspects may also need to be evaluated to ensure that no bias has been introduced. For the CRF or diary page, the time in the **BLANK** state could be lengthy, as it reflects the time between the printing of the page and the capture of the data.

A variation on the CRF case is the site-designed worksheet. Again this will be designed against the protocol but the process of validation may be less formal.

For a page from a medical record, the investigator will **select** a page as the target for the recording for the source data. This could be the instant before the data are captured and as such the time spent in the **BLANK** state will be extremely short.

The third case is the catchall case. This is where some sheet or scrap of paper is used to capture the data. At the time of capture there is no immediate material to write on and some easily available piece of paper is used. It could be the top sheet from the printer tray, the photocopier or a page from a notebook. Whatever is used, it is not the subject's own medical record nor is it some form specifically designed for the purpose, for example the CRF.

The **capture** operation will take the page from the **BLANK** state to the **POPULATED** state and the page from being just any piece of paper to being a source document. While in this state,

someone may **view** the document, they may **copy** the document or **inspect** it. In addition, a monitor may **monitor** the document.

Two of the state transitions in the above model are based on events that are not directly related to action undertaken on source documents. The first is the Study Ends event that could be triggered by a) the site having recruited enough subjects; b) a pre-determined end date being reached; or c) the site being told not to recruit any more subjects by the sponsor. The second event, the End of Retention Period event, is triggered by a given date being reached. Because they do not directly affect the direct handling of source documents these two events will not be considered further.

For each of the events, there is a process undertaken that directly interacts with the source documents and source data and, for each, there are relevant regulations. Therefore each has been examined by analyzing:

1. What is the action being undertaken?
2. What is the purpose of the action?
3. What are the drivers for the action?
    a. What are the relevant US and ICH regulations and guidance documents?
    b. How does the action contribute to, or impact, data quality and integrity?[31]
    c. What is the impact on subject safety?

[31] The definition used for data quality and data integrity are drawn from two FDA sources

Data Quality

FDA's acceptance of data from clinical trials for decision-making purposes is dependent upon its ability to verify the quality and integrity of such data during its onsite inspections and audits. To be acceptable the data should meet certain fundamental elements of quality whether collected or recorded electronically or on paper. Data should be attributable, original, accurate, contemporaneous, and legible.

Source: Computerized Systems used in Clinical Trials. FDA, April 1999.

Data Integrity

The degree to which a collection of data are complete, consistent, and accurate.

Source: Glossary Of Computerized System And Software Development Terminology (see http://www.fda.gov/ora/inspect_ref/igs/gloss.html)

Select

What is the action being undertaken?	The determination that a page will be used at some time in the future for the capture of source data. For a paper medical record this would probably be prior to the start of the trial – a conscious decision – while for a scrap of paper this may well be at the time of data capture.	
What is the purpose of the action?	Determines where the data will be captured.	
What are the relevant regulations?	None.	
How does it contribute to data quality and integrity?	**Complete**	Does not.
	Consistent	Does not.
	Accurate	Does not.
	Attributable	Does not.
	Original	Designated as the original source.
	Contemporaneous	Does not.
	Legible	Does not.
What is the impact on subject safety?	None.	

1234 ## *Create*

1235

What is the action being undertaken?	The design and creation of the page that is to be used in the future to capture source data. This includes the design and printing of CRF pages or patient diary cards, the design being based on that documented within the clinical trial protocol. Such custom-designed pages will be considered data collection instruments.	
What is the purpose of the action?	Details the information to be captured during the clinical encounter and the associated metadata (units etc).	
What are the relevant regulations?	21 CFR 312 Section 50 ICH GCP Section 2.6 and 6.4.9	
How does it contribute to data quality and integrity?	**Complete**	The checklist nature of data collection instruments aids in collecting complete data.
	Consistent	The metadata on the instrument (fields, units etc) assist in collecting data consistently.
	Accurate	The instrument promotes accurate data collection for the reasons detailed above.
	Attributable	The design of the instrument should include the subject's identifier and the identifier of the individual collecting the data.[32]
	Original	The design of blank instruments can aid in ensuring the capture of original data.
	Contemporaneous	As for original data.
	Legible	The instrument can help by formatting responses thus aiding legibility. While not ensuring legible data, good CRF design can greatly assist.
What is the impact on subject safety?	Poor design could lead to poor data and thus impact the safety of the subject.	

[32] Note that the identity of individuals needs to be verified.

The design of the instrument used to collect data will have an impact on the quality of the data collected. If a pre-defined form is not used, then the collection process could be prone to errors, especially in regard to complete and consistent measures. It is therefore important that any instrument used to collect the data accurately reflects the clinical protocol ***[Requirement 1]***[33] so as to improve the quality of the data collected. Hence, if an instrument is employed then, by definition, it will be the source document, and that source document must accurately reflect the protocol.

In addition, the instrument should allow the data to be accurately obtained and not influence, in the case of Subject Reported data, the response. Therefore the psychometric aspects of such an instrument are important.

[33] The requirement tags have been inserted so as to allow for ease of cross-reference.

Capture

What is the action being undertaken?	The investigator captures the (CRF) data onto the source document. A subject captures the (PRO) data onto the source document.	
What is the purpose of the action?	To capture the data required by the protocol.	
What are the relevant regulations?	21 CFR 312, Sections 60 and 62 ICH GCP Sections 1.51, 1.52, 4.9.1 and 6.4.9	
How does it contribute to data quality and integrity?	**Complete**	Process needs to ensure that all data required is captured.
	Consistent	Process needs to ensure that the same data are captured in the same manner.
	Accurate	Individual capturing the data needs to be trained in the accurate use of the document.
	Attributable	Subject ID and person capturing the data need to be recorded accurately. For subject reported source data, we need to be assured that it is the subject entering the data.
	Original	By definition, this item is the original document.
	Contemporaneous	Recorded at the time of capture. With subject reported data, we need to be assured that the data are recorded at the time stated.
	Legible	Individual needs to ensure that data are legible.
What is the impact on subject safety?	Accurate and timely capture of data helps build a trail of events with the study subject, and can provide an indication of potential evolving safety concerns.	

This operation is the foundation to ensuring that high-quality data are captured as part of a trial ***[Requirement 2]***. Every measure of data quality and integrity is impacted by this operation and as such can be seen to be important to source data.

Clarify

What is the action being undertaken?	Investigator checks and amends the data originally captured within a source document (CRF)[34].	
What is the purpose of the action?	To resolve an error in data collection that has been raised as part of a clarification.	
What are the relevant regulations?	21 CFR Part 312 Section 62 ICH GCP 4.9.3 and 5.5.4	
How does it contribute to data quality and integrity?	**Complete**	Potential to improve the complete nature of the data. Clarification is part of the process to ensure complete data.
	Consistent	Same argument as above.
	Accurate	Same argument as above.
	Attributable	Corrections should be made with an identification of who made the change.
	Original	No (but should not overlay the original).
	Contemporaneous	No (but do need to capture when the changes were made – historical trail).
	Legible	Clarification may be due to data being illegible. Any changes must ensure the existing data and the new data are legible.

What is the impact on subject safety?

The clarification process drives the need for the maintenance of the audit trail ***[Requirement 3]***. Should this operation be permitted for Subject Reported data? For example a Subject makes an entry and then reports to the investigator some error within that data. As part of this process, there is a regulatory requirement to maintain an audit trail.[35] This audit trail starts with the initial entry of the data.

[34] PRO data are generally not subject to clarifications.

[35] While there are clear statements within ICH GCP (section 5.5.4 as well as section 4.9.3) regarding an audit trail, the predicate rule drivers are less clear.

View / Inspect

What is the action being undertaken?	View or inspect the source document.	
What is the purpose of the action?	Check it or undertake an audit/inspection. This can include a regulatory inspection or the sponsor's monitors.	
What are the relevant regulations?	21 CFR 312 Section 58 ICH GCP 2.11, 5.15.1 and 8.3.13	
How does it contribute to data quality and integrity?	**Complete**	The complete record (including original entries and any changes) must be available for inspection.
	Consistent	None.
	Accurate	None. *See note above.*
	Attributable	None.
	Original	None.
	Contemporaneous	None.
	Legible	Must be understandable, even at some significantly later time.
What is the impact on subject safety?	Access may be important as well as the supply of complete records rather than subsets.	

No contribution is made to data quality or integrity by this event, as the event does not result in the data being amended. However, there are data integrity issues, if the copy supplied is not the entire record, or if part of the meaning is lost. The monitoring process may result in clarification events that will result in a change. This operation requires that the source documents are readily available ***[Requirement 4]***. Those source documents must be either the original or a certified copy ***[Requirement 5]***. In addition, so as to ensure that Subject confidentiality is maintained, access to the records should be restricted to those authorized to view them ***[Requirement 9]***.

Copy

What is the action being undertaken?	Copy a source document.	
What is the purpose of the action?	Provide a copy to another organization. This includes providing the data to the sponsor or providing a copy to a regulator.	
What are the relevant regulations?	21 CFR 312 Sections 58 and 68 ICH GCP 1.51 CSUCT II, VI.B and XI.A	
How does it contribute to data quality and integrity?	**Complete**	Copies should be complete, unless only the latest version was requested.
	Consistent	No impact.
	Accurate	The copy process must be 100% accurate.
	Attributable	No impact.
	Original	No impact.
	Contemporaneous	No impact.
	Legible	Copy must be complete and accurate, and able to be read and understood by the recipient.
What is the impact on subject safety?	Important that the copy is an accurate one. If only a partial record is supplied, this could result in inappropriate decisions, which could affect patient safety.	

It is a fundamental need that the copying of a source document provides for 100% accuracy to ensure that the copy is accurate. The copy should be certified to be such ***[Requirement 8]***.

Store

What is the action being undertaken?	Place the document into storage during the trial.	
What is the purpose of the action?	To preserve the source document during the life of the project and prevent its destruction or amendment.	
What are the relevant regulations?	21 CFR 312, Section 62 ICH GCP 4.9.3 and 4.9.4 ICH GCP 8.3.13	
How does it contribute to data quality and integrity?	**Complete**	Entire record must be retained (including original entries and changes, as well as the meaning and context of the record).
	Consistent	No.
	Accurate	No. *See note above.*
	Attributable	No.
	Original	No.
	Contemporaneous	No.
	Legible	No.
What is the impact on subject safety?	Important that the source document is maintained and is accessible.	

While a trial is active, existing source data should be readily available to the investigator and others such as monitors. During this period the source data must be protected from destruction ***[Requirement 7]***, either accidental or deliberate, and amendments must be made in a controlled fashion. One crucial aspect is that, while stored, changes can only be made with the investigator's approval ***[Requirement 6]***. The corollary is that changes by unauthorized individuals, either accidentally or deliberately, must be prevented. Changes must not be made in a manner that is unknown to the Investigator.

Monitor

What is the action being undertaken?	Check the source document and the data contained therein against that held by the sponsor	
What is the purpose of the action?	To ensure that the sponsor's database contains the correct data.	
What are the relevant regulations?	21 CFR Part 312, Sections 50 and 56	
How does it contribute to data quality and integrity?	**Complete**	Checked to ensure complete (all source data matches that being checked against – CRF or database).
	Consistent	Checked to ensure consistent (units and formats are the same).
	Accurate	Checked to ensure accurate (transcription).
	Attributable	Checked to ensure that the data are attributable.
	Original	No.
	Contemporaneous	No.
	Legible	No.
What is the impact on subject safety?	Safety analyses are checked against the sponsor's copy, thus this data must be correct.	

The process of ensuring that the source data, as captured on the source document is that held by the sponsor is an important one and has important impact on the quality of the data. However, the operation itself does not change source data. The check will result in a Modify operation being undertaken that does modify the data if an error is detected. There may also be safety concerns if the source does not match that held by the sponsor.

Note: In today's environment, the sponsor's copy is almost certainly held within a database. However, given the technology-independent approach of the analysis, the means by which the sponsor undertakes the source data verification process is not considered.

Archive

What is the action being undertaken?	Store the records during the regulatory retention period.	
What is the purpose of the action?	To hold the records during the retention period in-line with the regulations	
What are the relevant regulations?	21 CFR 312, Section 62 ICH GCP 4.9.5	
How does it contribute to data quality and integrity?	**Complete**	Must retain all records and the whole record.
	Consistent	No.
	Accurate	Ensure no meaning changed, if record was moved or reformatted.
	Attributable	Ensure no loss of identification.
	Original	Be able to retrieve the original.
	Contemporaneous	No.
	Legible	Do not lose the ability to read and understand.
What is the impact on subject safety?	Access to the data, and the ease of access, may be an issue. It the data cannot be accessed then this may impact subject safety.	

The archive operation is very similar to the Store operation. However, due to the potentially lengthier retention times, consideration for adequate protection of records (including environmental controls) is important.

Destroy

What is the action being undertaken?	Source document is physically destroyed.	
What is the purpose of the action?	Eliminate retention of unnecessary documentation (save space or other such resource(s)).	
What are the relevant regulations?	Nothing requires destruction.	
How does it contribute to data quality and integrity?	**Complete**	No.
	Consistent	No.
	Accurate	No.
	Attributable	No.
	Original	No.
	Contemporaneous	No.
	Legible	No.
What is the impact on subject safety?	Should records be destroyed? Can they be, if the source is also the individual patient's medical record?	

This event is here for completeness. There does not appear to be a consistent approach across industry regarding the destruction of source documents.[36]

[36] One comment on the first draft stated "Ideally, data would be archived indefinitely to facilitate future analyses. This is a major advantage of perpetual electronic storage over space-consumptive paper."

Appendix 2 – The Electronic World and 21 CFR Part 11

The text below is that which appears in 21 CFR Part 11 containing the regulations as they relate to electronic records. The impact of these regulations on the Key Requirements is assessed.

Section 11.10 Text	Map to Existing Requirements and any New Impact
Persons who use closed systems to create, modify, maintain, or transmit electronic records shall employ procedures and controls designed to ensure the authenticity, integrity, and, when appropriate, the confidentiality of electronic records, and to ensure that the signer cannot readily repudiate the signed record as not genuine. Such procedures and controls shall include the following:	Requirements 2 and 9.
(a) Validation of systems to ensure accuracy, reliability, consistent intended performance, and the ability to discern invalid or altered records.	Requirements 1 and 3 Validation ensures that all other requirements are met.
(b) The ability to generate accurate and complete copies of records in both human readable and electronic form suitable for inspection, review, and copying by the agency. Persons should contact the agency if there are any questions regarding the ability of the agency to perform such review and copying of the electronic records.	Requirements 4 and 8. The ability to produce a readable form of an electronic record may be implied by the requirements.
(c) Protection of records to enable their accurate and ready retrieval throughout the records retention period.	Requirements 4 and 7.
(d) Limiting system access to authorized individuals.	Requirements 2, 5, 6 , 7 and 9.
(e) Use of secure, computer-generated, time-stamped audit trails to independently record the date and time of operator entries and actions that create, modify, or delete electronic records. Record changes shall not obscure previously recorded information. Such audit trail documentation shall be	Requirement 3.

retained for a period at least as long as that required for the subject electronic records and shall be available for agency review and copying.	
(f) Use of operational system checks to enforce permitted sequencing of steps and events, as appropriate.	Requirement 1; also relates to Requirement 2 and the need to produce data of a high quality.
(g) Use of authority checks to ensure that only authorized individuals can use the system, electronically sign a record, access the operation or computer system input or output device, alter a record, or perform the operation at hand.	Requirements 2, 6, 7 and 9.
(h) Use of device (e.g., terminal) checks to determine, as appropriate, the validity of the source of data input or operational instruction.	Relates to Requirement 2 and the need to for attributability.
(i) Determination that persons who develop, maintain, or use electronic record/electronic signature systems have the education, training, and experience to perform their assigned tasks.	Potentially related to requirements 2, 4, 6 and 8.
(j) The establishment of, and adherence to, written policies that hold individuals accountable and responsible for actions initiated under their electronic signatures, in order to deter record and signature falsification.	Relevant but no direct impact (indirectly relevant to requirement 2).
(k) Use of appropriate controls over systems documentation including: (1) Adequate controls over the distribution of, access to, and use of documentation for system operation and maintenance. (2) Revision and change control procedures to maintain an audit trail that documents time-sequenced development and modification of systems documentation.	No impact.

1306

Section 11.30 Text	Impact on Key Requirements
Persons who use open systems to create, modify, maintain, or transmit electronic records shall employ procedures and controls designed to ensure the authenticity, integrity, and, as appropriate, the confidentiality of electronic records from the point of their creation to the point of their receipt. Such procedures and controls shall include those identified in Sec. 11.10, as appropriate, and additional measures such as document encryption and use of appropriate digital signature standards to ensure, as necessary under the circumstances, record authenticity, integrity, and confidentiality.	Requirements 2, 6 and 9.

The statement in section 11.10 that states *"and to ensure that the signer cannot readily repudiate the signed record as not genuine"* dictates a need to consider the case where source data are maintained away from the investigative site. If data are stored away from the site for prolonged periods of time, an investigator may have concerns about the inadvertent creation, copying, amendment or destruction of the source data and thus the accuracy of the data.

Requirement 5 states that: "The investigator shall maintain the original source document or a certified copy." This means that the Investigator always has a copy of the record which in turn means that any attempt to create, copy, amend or destroy a record would need to act on both copies or else the endeavor would result in mismatching records. This discrepancy would be discovered during inspection or a recreation of trial events.

However, it may also be advisable to explicitly state the requirements that the sponsor must never have exclusive control of a source document ***[Requirement 10].*** This requirement is inherent within the regulations in that 21 CFR 312 requires that an investigator maintain accurate case histories[37]. If a sponsor had exclusive control of the source documents, then an investigator could not fulfill their obligations. ICH GCP also indicates that source documents should reside in the files of the investigator / institution not the sponsor. [38] The FDA have also made direct reference to such measures in public presentations [16].

This results in the investigator being in a position where the source data cannot be repudiated. The ability for an investigator to be in a position of repudiation should be seen by a sponsor as a significant business risk.

[37] See 21 CFR 312.62(b).

[38] See ICH GCP E6 8.1 and 8.3.13.

Appendix 3 – Mapping to Technology

The development of the user requirements in Appendix 2, in a form that is independent of technology, permits common electronic solutions, as well as that of paper, to be assessed to judge their compliance. This assessment is useful in that it permits those areas where compliance is not clear cut to be identified and, as a consequence, allows an insight into where changes to the regulations may need to be made in the future.

It should be borne in mind that, when checking against the user requirements, a technology is neither compliant nor non-compliant. It is the combination of process and the technology that is important; the same system could be deployed in a compliant and non-compliant fashion depending on the processes used to operate it

The following technologies and implementations have been evaluated to see if they can meet the regulatory needs, using a strict interpretation as a means of undertaking a gap analysis to assist in developing the overall recommendations contained within this document.

1. Paper Medical Record
2. Case Report Form
    a. Paper CRF
    b. eCRF
        i. Thin (Web Browser) Client
        ii. Thick Client
3. Diaries
    a. Paper Diary
    b. eDiary
        i. Connected System
        ii. Semi Connected System
        iii. Disconnected System
4. Electronic Health Records
    a. Printed Records from an Electronic Health Record (EHR) system
    b. EHR System Used to Capture CRF Data

Site Data Collection

Paper Medical Record

Source data are captured and entered directly onto the subject's medical record. The source data will be copied to a CRF for transmission to the sponsor.

Requirement	Met By
An instrument used to capture source data shall ensure that the data are captured as specified within the protocol.	The subject's medical record is used and as such is not a purpose-built collection instrument. *Note: The medical record is not an instrument designed for the collection of the trial data. As such it may be prone to error.*
Source data shall be Accurate, Legible, Contemporaneous, Original, Attributable, Complete and Consistent.	The process used in capturing the data and mechanisms to identify who recorded it. There is no guarantee of legibility or attributability.
An audit trail shall be maintained as part of the source documents for the original creation and subsequent modification of all source data.	The initial use of the medical record initiates the audit trail. *Note: It is considered that an audit trail consists of both the original data and any subsequent modifications.* Amendment of the original data could be made on the medical record, the CRF page or both. The process needs to be clear such that the actual source is modified and consistency maintained between the medical record and that reported to the sponsor via the CRF.
The storage of source documents shall provide for their ready retrieval.	The investigator makes arrangement for the storage of the documents on-site.[39]
The investigator shall maintain the original source document or a certified copy.	The subject's medical record and CRF page are both stored at the site.
Source data shall only be modified with the knowledge or approval of the investigator.	The Investigator maintains control of the medical records at the site and needs to

[39] Investigators actually are allowed to store source documents "off-site" (e.g. for paper documents, in a storage facility.) However, they are responsible for maintaining control over the documents and who can access them.

	ensure that the medical record cannot be modified without appropriate controls.
Source documents and data shall be protected from destruction.	The Investigator needs to ensure that the medical record cannot be destroyed.
The source document shall allow for accurate copies to be made.	The subject's medical record is a paper document. Thus, it can be photocopied or information copied by hand. Verification needs to be undertaken to ensure that the copy is accurate (certified copy).
Source documents shall be protected against unauthorized access.	The Investigator needs to ensure that the medical record is protected against unauthorized access.
The sponsor shall not have exclusive control of a source document.	The medical record is in the control of the Investigator.

1362
1363

Paper CRF

Source data are captured and entered directly onto a three part, No Carbon Required (NCR), paper CRF. Site retains one of the three parts from the CRF.

Requirement	Met By
An instrument used to capture source data shall ensure that the data are captured as specified within the protocol.	The design of the CRF itself.
Source data shall be Accurate, Legible, Contemporaneous, Original, Attributable, Complete and Consistent.	The process used to fill in the CRF. *Note: Depending on the quality of the process employed, good documentation practices etc.*
An audit trail shall be maintained as part of the source documents for the original creation and subsequent modification of all source data.	Initial use of CRF creates the audit trail. Amendment of the original CRF form with capture of who, when, and what changed information. Use of good documentation practices (single line strikeout, initials, date) provides the ability to see a trail of changes.
The storage of source documents shall provide for their ready retrieval.	Investigator makes arrangement for the storage of the CRF pages on-site or off-site *Note: the CRF is the source document in this case. Also note however, a source document may contain a number of data items, some of which are source data and some of which are not.*
The investigator shall maintain the original source document or a certified copy.	The CRF page retained by the site is stored at the site.
Source data shall only be modified with the knowledge or approval of the investigator.	The investigator holds the CRF page at the site. Investigator should take steps to ensure that these pages cannot be modified without approval. However, this currently happens, as sponsors do change CRF data. In most instances, this is handled via a formal query (data clarification form) back to the Investigator, with some kind of written authorization to make the change. However, certain types of

	corrections are made by sponsors without authorization each time (these are usually identified at the beginning of the trial).
Source documents and data shall be protected from destruction.	The investigator stores and protects the documents. The process at the site is required to ensure that they cannot be destroyed. *Note: NCR copies removed from the site provide some of the best protection.*
The source document shall allow for accurate copies to be made.	Being paper, photocopies can be made. The investigator needs to sign these copies to state that the copy is accurate.
Source documents shall be protected against unauthorized access.	Investigator needs to takes steps to ensure that the CRF pages are stored such that unauthorized access is not possible.
The sponsor shall not have exclusive control of a source document.	Investigator maintains one part of the NCR CRF.

1367
1368

electronic CRF: Thin (Web Browser) Client

Source data are captured and entered directly into a web-based system without first being captured to paper. All data are stored on a central server that is located at sponsor.

Requirement	Met By
An instrument used to capture source data shall ensure that the data are captured as specified within the protocol.	The design of the eCRF forms and the system processing.
Source data shall be Accurate, Legible, Contemporaneous, Original, Attributable, Complete and Consistent.	The use of electronic capture should result in an improvement over a paper-based capture process. Attributable needs to be assured by the system (login, username, password etc.). Electronic entry eliminates problems with legibility. Use of identification mechanisms leads to attributable data. Completeness and consistency are advanced through the use of features such as drop-down lists of choices, online edits, check boxes, and branching based on entries. Use of automatic system date/time stamps yields the ability to determine if entries were contemporaneous.
An audit trail shall be maintained as part of the source documents for the original creation and subsequent modification of all source data.	The system needs to implement the audit trail requirement for the source data. This will be as part of the central database.
The storage of source documents shall provide for their ready retrieval.	The central server allows for ready retrieval. This requires assuring the server is available during times when all sites may need to access records. Records would need to be maintained on the central server for the regulatory retention period (and accessible by sites during this time) or archived and access provided to the sites.
The investigator shall maintain the original source document or a certified copy.	Only a single copy is stored on the central server. Therefore this arrangement cannot meet the requirements.
Source data shall only be modified with the knowledge or approval of the investigator.	Within this arrangement, fraudulent or accidental amendment is possible since the investigator does not have a copy of the

	source data/documents. Changes can be made without the approval of the investigator, but by having an audit trail immediately and readily available with the record, the investigator could become aware of changes, if periodic review is completed. However, administrative rights for a system may allow the audit trail to be circumvented.
Source documents and data shall be protected from destruction.	Steps can be taken at the central database to prevent destruction. However, fraudulent or accidental destruction is possible, due to the storage at a single location that is not the site. The audit trail may provide evidence of record deletion but administrative rights for a system may allow the audit trail to be circumvented.
The source document shall allow for accurate copies to be made.	Copies can be made from the central database. There is a need to define what is an accurate copy in an electronic sense. Accurate copies must include the meaning of the data (for example, date formats), as well as the full audit trail. The site would need to have the capability to review and generate copies.
Source documents shall be protected against unauthorized access.	Sponsor can take steps to ensure that the contents of the central database are protected against unauthorized access. However, this should be under the Investigator's control.
The sponsor shall not have exclusive control of a source document.	With this arrangement, the sponsor has exclusive control of the source data/documents.

1372

1373

electronic CRF: Thick Client

Source data are captured and entered directly into a thick-client application running at the site without being captured to paper. The data are stored at the site (typically a laptop is provided by the sponsor and the data reside on the laptop) prior and subsequent to being transmitted to the sponsor and stored on a central server. The source data remain on the laptop until the end of the site's participation in the study. Actions required to ensure compliance from the end of the study until the end of retention period are discussed below.

Requirement	Met By
An instrument used to capture source data shall ensure that the data are captured as specified within the protocol.	The design of the eCRF forms and the system processing.
Source data shall be Accurate, Legible, Contemporaneous, Original, Attributable, Complete and Consistent.	The use of electronic capture should result in an improvement over a paper-based capture process. Attributability needs to be assured by the system (login, username, password etc.). The system would need to ensure user identification is unique, since entries would be later consolidated into a central database.
An audit trail shall be maintained as part of the source documents for the original creation and subsequent modification of all source data.	The system needs to implement the audit trail requirement for the source data. This will be as part of the local client. Timestamps need to be defined, so that an understanding (and overall sequence of events) is maintained.
The storage of source documents shall provide for their ready retrieval.	The local client allows for ready retrieval. This is dependent on maintenance of the laptop (and software) at the site. *Note: This is analogous to paper, the laptop/software has to be under the control of the investigator.* It is also possible to archive the data off the laptop for long-term storage at the site and return the laptop to the sponsor. However, the source data should not be out of the control of the investigator at any point in this process.
The investigator shall maintain the original source document or a certified copy.	The original stored on the local client needs to be preserved.
Source data shall only be modified with the	The source data stored on the local client

knowledge or approval of the investigator.	allows for the investigator to be able to meet this requirement. However, the sponsor copy can be changed without the investigator's knowledge. This would only be evident upon inspection of both sets of records.
Source documents and data shall be protected from destruction.	Steps can be taken at the local client to prevent destruction. The site needs to ensure that the local copy can be read over time (logical "destruction" can occur if the source becomes unreadable due to hardware failure or software obsolescence). Regular backups should be taken, to protect against hardware failure.
The source document shall allow for accurate copies to be made.	Copies can be made from the local client. Accurate copies must include the meaning of data (for example, date formats), as well as the full audit trail. The site would need to have the capability to review and generate copies.
Source documents shall be protected against unauthorized access.	The investigator can take steps to ensure that the contents of the local client are protected against unauthorized access. However, access to the copy is uncontrolled by the investigator.
The sponsor shall not have exclusive control of a source document.	As long as the source data are kept within the control of the investigator until the end of the regulatory retention period, then the requirement can be met.

1381
1382

Summary – Site Data Collection

The table below provides a summary of the analysis presented above.

Requirement	Medical Record	Paper CRF	Web Client	Thick Client
An instrument used to capture source data shall ensure that the data are captured as specified within the protocol.	NA	✓	✓	✓
Source data shall be Accurate, Legible, Contemporaneous, Original, Attributable, Complete and Consistent.	✓	✓	✓	✓
An audit trail shall be maintained as part of the source documents for the original creation and subsequent modification of all source data.	✓	✓	✓	✓
The storage of source documents shall provide for their ready retrieval.	✓	✓	✓	✓
The investigator shall maintain the original source document or a certified copy.	✓	✓		✓
Source data shall only be modified with the knowledge or approval of the investigator.	✓	✓		✓
Source documents and data shall be protected from destruction.	✓	✓		✓
The source document shall allow for accurate copies to be made.	✓	✓	✓	✓
Source documents shall be protected against unauthorized access.	✓	✓	✓	✓
The sponsor shall not have exclusive control of a source document.	✓	✓		✓

Several issues arise within the above analysis

1. Web-based CRF systems, where no paper source documents are maintained[40], are unable to meet the key requirements due to the fact that the investigator does not store the source data on-site and, as a consequence, that data are open to modification without the investigator's knowledge.
2. The requirements of making a "certified copy" of an electronic record need to be fully understood. In the paper world, a document can be photocopied, a check made to ensure that the photocopy is accurate in the sense that no data have been obscured or lost and the paper signed to indicate such. A similar "process" is required for electronic records. This may mean manual review and electronic or digital signature, although that process would be labor intensive for large volumes and would mean significant application changes for existing technology. Therefore, some form of electronic check may need to be considered. A validated process for generating complete and accurate copies might be considered sufficient.
3. There is an issue with thick client systems in the exclusive control of source data. The source data needs to be retained by the investigator through the collection phase of a study and through the retention period. Sponsors using such systems may collect client systems, such as laptops, at the end of the trial for the reasons of efficiency. Unless a certified copy of the source data are made prior to the removal of the client from investigator control, and this copy remains in the control of the investigator, the system would fail to meet the key requirements.

[40] Note: Where a web-based system is used in conjunction with a medical record or CRF approach, then the analysis that applies is that for paper source documents, the use of the technology is not relevant to the regulatory discussion.

Subject Data Collection

Paper Diary

Source data are entered directly onto a paper diary by the subject. The subject passes the diary to the investigator. The data are then passed to the sponsor with a copy of the diary being kept at the site[41].

Requirement	Met By
An instrument used to capture source data shall ensure that the data are captured as specified within the protocol.	The design of the diary cards and training of the study subject.
Source data shall be Accurate, Legible, Contemporaneous, Original, Attributable, Complete and Consistent.	The process used to fill in the diary. Research has indicated that diary data fails to meet a number of the ALCOA requirements due to the methods employed by subjects. Handwritten paper diaries lead to legibility problems, although attributability of all entries to the subject is more evident. Studies have shown that in many cases, the entries are not contemporaneous. Design of the diary form can contribute to consistency, however, handwritten entries may result in inconsistent entries, multiple choices selected, etc.
An audit trail shall be maintained as part of the source documents for the original creation and subsequent modification of all source data.	Initial use of diary by the subject creates the audit trail. Study subjects are not likely to use good documentation practices for making corrections to diaries, thus changed entries may result in obscuring the original data.[42] Amendment of the original diary form may be made by the investigator, as long as information is captured as to who made the

[41] It is understood that many studies do not follow this practice in that no copy of the diary is maintained at the site. However, such a practice would be in contravention of the existing predicate rules. It is also noted that some protocols explicitly state that the investigator should not review. There is a difference however, between storage and reviewing the data.

[42] It is probably worth creating the distinction between amendments made by the subject prior to passing the diary to the site and those changes made post that time.

	change and when. *Note: There have been cases where a subject has reported, after entering diary data, that they have entered it incorrectly. This raises an interesting question, should the investigator amend it in line with the subject's wishes?*
The storage of source documents shall provide for their ready retrieval.	Investigator makes arrangement for the storage of the diary on-site.
The investigator shall maintain the original source document or a certified copy.	The diary is retained by the site and is stored at the site.
Source data shall only be modified with the knowledge or approval of the investigator.	The investigator holds the diary at the site. Investigator should take steps to ensure that these pages cannot be modified without approval.
Source documents and data shall be protected from destruction.	The investigator stores and protects the documents. A process at the site is required to ensure that they cannot be destroyed.
The source document shall allow for accurate copies to be made.	Being paper, photocopies can be made. The investigator needs to sign these copies to state that the copy is accurate.
Source documents shall be protected against unauthorized access.	The investigator needs to takes steps to ensure that the diary is stored such that unauthorized access is not possible.
The sponsor shall not have exclusive control of a source document.	The investigator maintains the original diary. *Note: It is noted that this is not always existing practice.*

1413
1414

electronic Diary: Connected System

Source data are captured and entered directly into a diary application without being captured to paper. The application communicates with a central server located at the sponsor and has to be connected for the entire duration of the data entry session. Examples of such systems are a web site accessed using a browser running on a PDA or PC or an Interactive Voice Response System (IVRS) patient diary system. All source data are stored on a central server.

Requirement	Met By
An instrument used to capture source data shall ensure that the data are captured as specified within the protocol.	The design of the eDiary form or IVRS script.
Source data shall be Accurate, Legible, Contemporaneous, Original, Attributable, Complete and Consistent.	The use of electronic capture should result in an improvement over a paper-based capture process. Attributability needs to be assured by the system (login, username, password etc.). Contemporaneous, legibility, completeness, and consistency are facilitated by proper design and use of the system.
An audit trail shall be maintained as part of the source documents for the original creation and subsequent modification of all source data.	The system needs to implement the audit trail requirement for the source data. This will be part of the central database application. It should be determined who is allowed to change entries.
The storage of source documents shall provide for their ready retrieval.	The central server allows for ready retrieval. This requires assuring the server is available during times when all sites may need to access records. Records would need to be maintained on the central server for the regulatory retention period (and accessible by sites during this time).
The investigator shall maintain the original source document or a certified copy.	Only a single copy is stored on the central server. Therefore this arrangement cannot meet the requirements.
Source data shall only be modified with the knowledge or approval of the investigator.	Within this arrangement, fraudulent or accidental amendments are possible since the investigator does not have a copy of the source data/documents.
Source documents and data shall be protected from destruction.	Steps can be taken at the central database to prevent destruction. However, fraudulent or

	accidental destruction is possible due to storage at a single location that is not the site.
The source document shall allow for accurate copies to be made.	Copies can be made from the central database. Need to define what an accurate copy is in an electronic sense. Accurate copies must include the meaning of the data (for example, date formats) as well as the full audit trail. The site would need to have the capability to review and generate copies.
Source documents shall be protected against unauthorized access.	Sponsor can take steps to ensure that the contents of the central database are protected against unauthorized access. However, this should be under investigator control.
The sponsor shall not have exclusive control of a source document.	With this arrangement, the sponsor has exclusive control of the source data/documents. Therefore, this arrangement cannot meet the requirements.

1421
1422

electronic Diary: Semi-Connected System

Source data are entered directly into a thick-client diary application, generally running on a PDA, without first being captured to paper. The PDA has the ability to communicate with a central server using some form of wireless technologies (mobile cellular technology, WiFi or similar). The data are stored on the PDA until such time as they can be transmitted to the sponsor and stored on a central server. The data are not preserved on the PDA post transmission.

Requirement	Met By
An instrument used to capture source data shall ensure that the data are captured as specified within the protocol.	The design of the eDiary application and the configuration of the diary within it.
Source data shall be Accurate, Legible, Contemporaneous, Original, Attributable, Complete and Consistent.	The use of electronic capture should result in an improvement over a paper-based capture process. Attributability needs to be assured by the system (login, username, password etc.). Contemporaneous, legibility, completeness, and consistency are facilitated by proper design and use of the system.
An audit trail shall be maintained as part of the source documents for the original creation and subsequent modification of all source data.	The system as a whole (central server and local application) needs to implement the audit trail requirement for the source data. This will be as part of the central database. If changes are made on the local device, the audit trail needs to be transmitted also to the central database.
The storage of source documents shall provide for their ready retrieval.	The central server allows for ready retrieval. This requires assuring the server is available during times when all sites may need to access records. Records would need to be maintained on the central server for the regulatory retention period (and accessible by sites during this time).
The investigator shall maintain the original source document or a certified copy.	Only a single copy is stored on the central server. Therefore this arrangement cannot meet the requirements.
Source data shall only be modified with the knowledge or approval of the investigator.	Within this arrangement, fraudulent or accidental amendment is possible since the investigator does not have a copy of the source data/documents.

Source documents and data shall be protected from destruction.	Steps can be taken at the central database to prevent destruction. However, fraudulent or accidental destruction is possible due to storage at a single location that is not the site.
The source document shall allow for accurate copies to be made.	Copies can be made from the central database. Need to define what an accurate copy is in an electronic sense. Accurate copies must include the meaning of the data (for example, date formats) as well as the full audit trail. The site would need to have the capability to review and generate copies.
Source documents shall be protected against unauthorized access.	Sponsor can take steps to ensure that the contents of the central database are protected against unauthorised access. However, this should be under investigator control.
The sponsor shall not have exclusive control of a source document.	With this arrangement, the sponsor has exclusive control of the source data/documents. Therefore, this arrangement cannot meet the requirements.

Note: A central concern about this model, like paper, is that the intermediate source of data (the PDA, for example), can be lost, corrupted, dropped in the toilet, etc., resulting in lost data.

Electronic Diary: Disconnected System

Source data are captured and entered directly into an application running on a PDA or similar device that has no communication capability other than that provided when the device is "docked" with a PC. The data are stored on the device until such time as they can be copied off. This copying of the source data is achieved by "docking" the device with a PC. The data are stored at the site prior to being transmitted to the sponsor and stored on a central server. The data are preserved at the site on the PC that the PDA was docked with.

Requirement	Met By
An instrument used to capture source data shall ensure that the data are captured as specified within the protocol.	The design of the eDiary application and the configuration of the diary within it.
Source data shall be Accurate, Legible, Contemporaneous, Original, Attributable, Complete and Consistent.	The use of electronic capture should result in an improvement over a paper-based capture process. Attributability needs to be assured by the system (login, username, password etc.). Contemporaneous, legibility, completeness, and consistency are facilitated by proper design and use of the system.
An audit trail shall be maintained as part of the source documents for the original creation and subsequent modification of all source data.	The local application needs to implement the audit trail requirement for the source data. If changes are made at the PDA or site PC, the changes need to be transmitted with the data. Timezone and meaning of timestamps would need to be defined, for accurate understanding and reconstruction.
The storage of source documents shall provide for their ready retrieval.	Once the source data are stored on the site PC, they can be readily retrieved, as long as the hardware and software are maintained.
The investigator shall maintain the original source document or a certified copy.	The source data are stored on the device and then transferred to the site's PC. Validation of the transfer can assure an accurate and complete copy.
Source data shall only be modified with the knowledge or approval of the investigator.	The investigator can take steps to protect the source data when at the site.
Source documents and data shall be protected from destruction.	The local PC allows for ready retrieval. This is dependent on maintenance of the PC (and software) at the site.

The source document shall allow for accurate copies to be made.	Copies can be made from the local data source. Need to define what an accurate copy is in an electronic sense.
Source documents shall be protected against unauthorized access.	The investigator can takes steps to protect the site system.
The sponsor shall not have exclusive control of a source document.	Investigator holds the source data and the sponsor has a copy.

1440
1441

Summary – Subject Data Collection

The table below provides a summary of the analysis presented above.

Requirement	Paper Diary	Connected	Semi-Connected	Disconnected
An instrument used to capture source data shall ensure that the data are captured as specified within the protocol.	✓	✓	✓	✓
Source data shall be Accurate, Legible, Contemporaneous, Original, Attributable, Complete and Consistent.		✓	✓	✓
An audit trail shall be maintained as part of the source documents for the original creation and subsequent modification of all source data.	✓	✓	✓	✓
The storage of source documents shall provide for their ready retrieval.	✓	✓	✓	✓
The investigator shall maintain the original source document or a certified copy.	✓			✓
Source data shall only be modified with the knowledge or approval of the investigator.	✓			✓
Source documents and data shall be protected from destruction.	✓			✓
The source document shall allow for accurate copies to be made.	✓	✓	✓	✓
Source documents shall be protected against unauthorized access.	✓			✓
The sponsor shall not have exclusive control of a source document.	✓			✓

The analysis raises similar issues to those raised for CRF data with respect to the source data copy process. What is more noticeable with a diary system than with a CRF system is the potential for progressive movement of source data. This can be seen with the disconnected

example above where the source data resides on a device until that device is docked. It may then be copied to a PC to be prepared for transmission to a sponsor. Here we see the data copied twice, once from the device to the PC and then again when it is passed to the sponsor. But the source data could still be considered to be on the device, or if deleted, on the PC. It is therefore important to consider the location of the source data as well as the mechanism for copying it.

Electronic Health Records

Printed Records From an EHR System

Data are entered into an EHR system without being captured to paper and then the relevant data are printed, the printed record checked and signed to indicate that the data are accurate and the printed record then used for regulatory purposes. In this situation the situation would be that as per paper medical records as described above and 21 CFR Part 11 would not apply. This is the application of the "typewriter" rule as the paper record is the source document and it is that paper record that is used for regulatory purpose.[43]

EHR System Used to Capture CRF Data

Source data are captured and entered directly into a site-based EHR system without being captured to paper. The data are stored at the site within the normal EHR data repository prior to being sent, electronically, to the sponsor.

Requirement	Met By
An instrument used to capture source data shall ensure that the data are captured as specified within the protocol.	EHRs, in general, would not have specific functionality designed for the capture of research data. This may be an issue.
Source data shall be Accurate, Legible, Contemporaneous, Original, Attributable, Complete and Consistent.	The use of electronic capture should result in an improvement over a paper-based capture process. Attributability needs to be assured by the system (login, username, password etc.). The system would need to ensure user identification is unique, since entries would be later consolidated into a central database.
An audit trail shall be maintained as part of the source documents for the original creation and subsequent modification of all source data.	The system needs to implement the audit trail requirement for the source data. This will be as part of the EHR. Timestamps need to be defined, so that an understanding (and overall sequence of events) is maintained.
The storage of source documents shall	The EHR system allows for ready retrieval.

[43] "On the other hand, when persons use computers to generate paper printouts of electronic records, and those paper records meet all the requirements of the applicable predicate rules and persons rely on the paper records to perform their regulated activities, FDA would generally not consider persons to be "using electronic records in lieu of paper records" under §§ 11.2(a) and 11.2(b). In these instances, the use of computer systems in the generation of paper records would not trigger part 11."

***Source**: Guidance for Industry. Part 11, Electronic Records; Electronic Signatures — Scope and Application, Lines 166 to 171.*

provide for their ready retrieval.	*Note: This is analogous to paper, the EHR being under the control of the investigator/site.*
The investigator shall maintain the original source document or a certified copy.	The original is stored within the EHR.
Source data shall only be modified with the knowledge or approval of the investigator.	The source data stored in the EHR allows for the investigator to be able to meet this requirement.
Source documents and data shall be protected from destruction.	Steps can be taken at the site to prevent destruction. The site needs to ensure that the local copy can be read over time.
The source document shall allow for accurate copies to be made.	Copies can be made from the EHR. Formats, ease of use and other such considerations come to mind.
Source documents shall be protected against unauthorized access.	The investigator/site can take steps to ensure that the contents of the EHR are protected against unauthorized access.
The sponsor shall not have exclusive control of a source document.	As long as the source data are kept within the control of the investigator until the end of the regulatory retention period, then the requirement can be met.

1467

1468 The EHR can meet the user requirements for the maintenance of source data. However, the
1469 implication for health care providers is that such a system would come under the provisions of 21
1470 CFR Part 11 as records defined under the predicate rules are being held in electronic form. This
1471 is something that an organization may not wish to contemplate, as this would affect the entire
1472 institution and all users of the system.

1473

Appendix 4 – Regulatory Text

ICH GCP 1.51 source data

All information in original records and certified copies of original records of clinical findings, observations, or other activities in a clinical trial necessary for the reconstruction and evaluation of the trial. Source data are contained in source documents (original records or certified copies).

ICH GCP 1.52 source documents

Original documents, data, and records (e.g., hospital records, clinical and office charts, laboratory notes, memoranda, subjects' diaries or evaluation checklists, pharmacy dispensing records, recorded data from automated instruments, copies or transcriptions certified after verification as being accurate copies, microfiches, photographic negatives, microfilm or magnetic media, x-rays, subject files, and records kept at the pharmacy, at the laboratories and at medico-technical departments involved in the clinical trial).

ICH GCP 2.6

A trial should be conducted in compliance with the protocol that has received prior institutional review board (IRB)/independent ethics committee (IEC) approval/favourable opinion.

ICH GCP 2.10

All clinical trial information should be recorded, handled, and stored in a way that allows its accurate reporting, interpretation and verification.

ICH GCP 2.11

The confidentiality of records that could identify subjects should be protected, respecting the privacy and confidentiality rules in accordance with the applicable regulatory requirement(s).

ICH GCP 4.9.1

The investigator should ensure the accuracy, completeness, legibility, and timeliness of the data reported to the sponsor in the CRFs and in all required reports

ICH GCP 4.9.3

Any change or correction to a CRF should be dated, initialled, and explained (if necessary) and should not obscure the original entry (i.e. an audit trail should be maintained); this applies to both written and electronic changes or corrections (see 5.18.4(n)). Sponsors should provide guidance to investigators and/or the investigators' designated representatives on making such corrections. Sponsors should have written procedures to assure that changes or corrections in CRFs made by sponsor's designated representatives are documented, are necessary, and are endorsed by the investigator. The investigator should retain records of the changes and corrections.

ICH GCP 4.9.4

The investigator/institution should maintain the trial documents as specified in Essential Documents for the Conduct of a Clinical Trial (see 8.) and as required by the applicable regulatory requirement(s). The investigator/institution should take measures to prevent accidental or premature destruction of these documents.

ICH GCP 4.9.5

Essential documents should be retained until at least 2 years after the last approval of a marketing application in an ICH region and until there are no pending or contemplated marketing applications in an ICH region or at least 2 years have elapsed since the formal discontinuation of clinical development of the investigational product. These documents should be retained for a longer period however if required by the applicable regulatory requirements or by an agreement with the sponsor. It is the responsibility of the sponsor to inform the investigator/institution as to when these documents no longer need to be retained (see 5.5.12).

ICH GCP 5.5.4

If data are transformed during processing, it should always be possible to compare the original data and observations with the processed data.

ICH GCP 5.15.1

The sponsor should ensure that it is specified in the protocol or other written agreement that the investigator(s)/institution(s) provide direct access to source.

ICH GCP 6.4 and 6.4.9

The scientific integrity of the trial and the credibility of the data from the trial depend substantially on the trial design. A description of the trial design, should include:

…

The identification of any data to be recorded directly on the CRFs (i.e. no prior written or electronic record of data), and to be considered to be source data.

ICH GCP 8.3.13

Title of Document	Purpose	Located in Files of	
		Investigator / Institution	Sponsor
SOURCE DOCUMENTS	To document the existence of the subject and substantiate integrity of trial data collected. To include original documents related to the trial, to medical treatment, and history of subject	X	

21 CFR 312.50

Sponsors are responsibile [SIC] for selecting qualified investigators, providing them with the information they need to conduct an investigation properly, 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, maintaining an effective IND with respect to the investigations, and ensuring that FDA and all participating investigators are promptly informed of significant new adverse effects or risks with respect to the drug. Additional specific responsibilities of sponsors are described elsewhere in this part.

21 CFR 312.60

An investigator is responsible for ensuring that an investigation is conducted according to the signed investigator statement, the investigational plan, and applicable regulations; for protecting the rights, safety, and welfare of subjects under the investigator`s care; and for the control of drugs under investigation. An investigator shall, in accordance with the provisions of part 50 of this chapter, obtain the informed consent of each human subject to whom the drug is administered, except as provided in Secs. 50.23 or 50.24 of this chapter. Additional specific responsibilities of clinical investigators are set forth in this part and in parts 50 and 56 of this chapter.

21 CFR 312.62

(b) Case histories. An investigator is required to prepare and maintain adequate and accurate case histories that record all observations and other data pertinent to the investigation on each individual administered the investigational drug or employed as a control in the investigation. Case histories include the case report forms and supporting data including, for example, signed and dated consent forms and medical records including, for example, progress notes of the physician, the individual`s hospital chart(s), and the nurses` notes. The case history for each individual shall document that informed consent was obtained prior to participation in the study.(c) Record retention. An investigator shall retain records required to be maintained under this part for a period of 2 years following the date a marketing application is approved for the

drug for the indication for which it is being investigated; or, if no application is to be filed or if the application is not approved for such indication, until 2 years after the investigation is discontinued and FDA is notified.

CSUCT II – Definitions

Certified Copy means a copy of original information that has been verified, as indicated by dated signature, as an exact copy having all of the same attributes and information as the original.

CSUCT VI – System Features

B. Systems used for direct entry of data should be designed to include features that will facilitate the inspection and review of data. Data tags (e.g., different color, different font, flags) should be used to indicate which data have been changed or deleted, as documented in the audit trail.

CSUCT XI – Records Inspection

A. FDA may inspect all records that are intended to support submissions to the Agency, regardless of how they were created or maintained. Therefore, systems should be able to generate accurate and complete copies of records in both human readable and electronic form suitable for inspection, review, and copying by the Agency. Persons should contact the Agency if there is any doubt about what file formats and media the Agency can read and copy.

Appendix 5 – Mapping of User Requirements to Regulatory Text

A mapping between the user requirements and the regulatory text is provided in the table below.

User Requirement	21 CFR 312 (section)	ICH GCP (section)	CSUCT (section)
1	50	2.6 6.4.9	
2	60 62	1.51 1.52 4.9.1 6.4.9	
3	62	4.9.3 5.5.4	
4	58	2.11 5.15.1	
5	58	2.11 5.15.1	
6	62	4.9.3 4.9.4 Chapter 8	
7	62	4.9.3 4.9.4 Chapter 8	
8	58 68	1.51	II VI.B XI.A
9	58	2.11 5.15.1	
10	62	8.3.13	

Note: Requirement 4 could also be mapped to 21 CFR Part 312, Sections 50, 56 as a result of the monitoring function.

Appendix 6 – Process for the Development of this Document

The initial eSDI document, produced in March 2005, was incomplete and not intended for broad external review. However, because there were no vendors in the eSDI Group and CDISC is an open organization, the rationale for the work of the eSDI group was presented, along with an overview of the initial analysis of the existing regulatory environment for eSource data and how it relates to today's technologies, to open groups of interested parties during February-April. These groups included sponsors, technology vendors, contract research organizations (CROs) and investigator site representatives; they were convened at four pre-specified meetings that occurred in conjunction with other industry conferences between 23 February and 6 April 2005 (Philadelphia – Clinical Trials Congress; Lisbon – DIA EuroMeeting, Arlington-DIA ePRO Workshop and Philadelphia - SUGI). Attendees from these meetings who expressed interest in reviewing the initial draft of the document were sent a copy via e-mail (beginning mid-March), and comments were requested. The first draft was distributed more broadly than anticipated, hence comments were also received from individuals and groups who did not attend the meetings to hear the rationale and analysis presented. By early April, written comments had been received from five vendor companies, three pharmaceutical companies, the Pharmaceutical Research and Manufacturers of America (PhRMA) Electronic Data Capture (EDC) task force, one CRO, one investigative site and two individuals from the Food and Drug Administration. Others reviewed the initial draft and offered verbal comments.

The second draft was developed by May 2005 and addressed comments received on the initial draft. The second differed, in particular, from the initial draft as follows:

1. The sections on psychometrics and validation were extracted and placed 'on hold' at least until the FDA Guidance on Patient Reported Outcomes (PRO) is released;
2. The sections from the first draft on the analysis of current regulations and mapping to existing technologies were placed in appendices – these are considered exercises that form the basis for the user requirements section;
3. The second draft includes recommended best practices such as the leveraging of standards for the use of technology in the context of existing regulations and considerations for potential future changes in existing regulations to pave the way for research at the point of care and care at the point of research.
4. Throughout the text of the second draft and in the appendices, the comments received from reviewers (other than those on the psychometrics and validation sections) have been addressed and/or incorporated as appropriate. Comments and responses to these will be made available.

The eSDI group that contributed to the initial draft included two vendors, three site representatives, two sponsors, a validation expert, a CDISC representative and five liaisons ("observers") from FDA. This has now been expanded to include three more sponsor representative, another site representative, a CRO representative and a vendor/CRO representative.

The second draft of the document was prepared as detailed above and upon input from the minutes of the 6 April face-to-face meeting of the eSDI group and the comments received on the first draft. This was done by a small set of representatives of the eSDI Group. On 24 May, this second draft was distributed to the others in the eSDI group simultaneously with the CDISC Board of Directors, the CDISC Director of Technical Coordination, CDISC Operations representatives, and a small group of CDISC Industry Advisory Board representatives who agreed to provide 'preview' comments with a 48-hour turnaround. The eSDI Group met on 1 June to discuss and revise the second draft. The process for the release and review of the third draft was discussed at the 1 June face-to-face meeting of the eSDI Group and by the CDISC Board of Directors. It was decided that comments received on the second draft should be incorporated into a third draft, which should be again reviewed by the eSDI group and the CDISC Board prior to broader distribution.

An eSDI summary document was prepared prior to the DIA Annual meeting in June to give an overview of the purpose, scope and recent progress of the group.

This third draft document differs from the former two in the following primary ways:

1. The Requirements and Recommendations sections were discussed at the 1 June meeting of the eSDI group and the content was revised significantly, particularly in the Recommendations section.
2. The initial Important Background section and the Future Vision sections were deleted and sections on Purpose and Scope were added.
3. The Rationale and Introduction Section was revised and an Executive Summary was added.
4. New appendices were added.
5. The comments received on draft version 1 and 2 were either incorporated into the version 3 document or will be addressed individually in a spreadsheet.

The eSDI Group and Board have reviewed the third and fourth draft versions to determine if and when a broader view could occur. The fifth draft version was distributed broadly for an open comment period of at least 30 days. Numerous comments were received. These were each addressed in a separate document.

The sixth version was distributed for review by a focused group, specifically those who provided comments to the fifth version and the eSDI Group. Additional comments were received and addressed. A teleconference was held with the eSDI group to discuss any that were unclear or required further discussion for resolution. In addition, one system was used as a test case for completion of the Source Data Evaluation Report. It was determined that Scenario 4 was not clearly described. Following a consultation meeting with FDA representatives, this scenario was further clarified.

This represents the seventh version and will be posted as Version 1.0.

Appendix 7 – Responsibilities

Pharmaceutical clinical research involves multiple parties – the sponsor, clinical investigator, clinical lab, perhaps a contract research organization (CRO), and of course the FDA. And even more organizations may be involved when computer systems are used to capture data, with the generation of eSource. These include the vendor of the software, and potentially also a data hosting service. Some of these parties have responsibilities per regulations (related to interaction with subjects, following the protocol, generation and retention of accurate data). Each of these parties have responsibilities regarding implementation of systems and ensuring they are used correctly. Let's take a closer look at the responsibilities and expectations of each party.

Investigational Site

The clinical investigator has primary responsibility for ensuring the investigation is conducted according to the protocol. This includes accurate use of systems provided for data collection. The investigator is responsible to ensure the accuracy and completeness of data collected for clinical trials. This means that mechanisms used to access systems for the purpose of creating, modifying, and/or viewing data must be protected and used only by the authorized individuals. In other words, each individual has a private access mechanism (such as a password) which is not shared. Sharing access mechanisms can allow unauthorized individuals to have access to systems and data and defeats the ability to determine who created or modified data. The investigator is responsible for the accuracy of the source data collected at his/her site. It is important to protect source data to allow for subsequent record verification against it. The sponsor is ultimately responsible for the accuracy of the data sent to the FDA in support of the sponsor's marketing application. The FDA must be able to determine whether the information in the submission accurately reflects what happened at the clinical sites. The investigator is responsible to ensure records represent the measurements and assessments that were taken at the site. If data are to be changed from what was originally recorded, this must be at the authorization of the clinical investigator. Although there are other staff at the site who are involved with seeing subjects, administering study drug, taking samples, and filling out study paperwork, the investigator cannot abdicate his/her responsibility for the accuracy of study data. There needs to be a way to tell if records have changed from the original recording. Many systems include an automatic audit trail feature, which will capture who made a change to a record, when the change was made, and what the change was. These audit trails should also capture if records were deleted. The investigator should regularly review these "record histories," to ensure that nothing unintended is happening to the data.

The investigator also must, per regulatory requirements, ensure that clinical trial data and records are available for inspection and copying at any point by the FDA. Although more commonly this occurs after the study is completed, a regulatory inspection may occur during the course of a study. In this instance, the expectations are the same – that the data are capable of being retained accurately, retrieved in a timely manner, and copies can be supplied. Along with this concept of availability, the investigator is responsible to ensure the records are protected against loss. In an electronic world, this takes on new meaning beyond fire or water damage. Computer hard drive failure, loss or compatibility of the software needed to read a record, the inability to remember which directory something was filed in, the lost ability to read/recover a backup from media all

lead essentially to "loss" of a record. Thus, additional protections must be taken, to guard against loss of electronic source.

The investigator is responsible to ensure that staff involved in the study conduct are qualified to perform their functions. This includes being trained on how to use data collection and retrieval systems, as well as the procedures for use within the organization and the mechanisms for protection of data. This training should address security policies, both internal and external (at the site and at the sponsor). In most instances, systems being used to collect electronic source for trials are being supplied by sponsors. However, the investigator can still expect to receive assurances of the accurate functioning of the software, the security mechanisms of the software, and the ability to retrieve records and make copies available. Systems may be implemented solely by the site for regulatory activities such as investigational drug accountability, storage of regulatory records or adverse event tracking. If the site has implemented their own systems, the investigator is responsible to ensure these are accurately validated - defined, specified, verified, tested, installed, made secure, maintained under control, and document that all these processes and activities have been completed. This is true even if the software is purchased. Control should be maintained over local PCs that have study-related software installed and/or are holding study data. Introduction of other software can have an unintended adverse effect on the accuracy of the software or the integrity of the data. The investigator is also responsible to ensure accurate receipt of electronic results from other parties such as clinical labs, imaging services, specialty test providers, and image reading services. If data from these sources is coming in electronically, there need to be accurate specifications for how to receive it and load it into any receiving systems. The sponsor may address this for the sites. However, if sites are receiving electronic lab data into their own systems, the investigator is then responsible to specify and verify the accurate mapping and loading of the data.

The investigator also has primary responsibility to protect the rights, safety, and welfare of subjects. This means that each subject's results and reports should be regularly reviewed to detect any safety concerns which should be arising. Investigators cannot do this if they do not have immediate access to and control over the source data.

As can be seen by the above discussion, even when systems are supplied by the sponsor, investigators still maintain several responsibilities regarding the use of those systems and cannot assume or defer that these are the responsibilities of the sponsor or the vendor.

Sponsor

In many instances, sponsors are supplying the sites with the systems that are to be used for collection of clinical trial data. Even though system validation has traditionally required the end user of the system to be responsible for its accuracy, the level of knowledge on the part of clinical sites related to electronic record systems issues is still not widespread. So sponsors find themselves working on behalf of the sites to address accurate delivery of systems. Thus, sponsors take on responsibility to define the requirements for the system, ensure the system is accurately designed and built to meet those requirements, verify the functions are delivered correctly, and ensure the system can be installed consistently and correctly (especially where the system may be deployed at multiple locations, such as multiple global servers or on individual workstations). The sponsor also must help assure that changes made to the system or its environment do not impact the accurate functioning of the system or the integrity of the data

already collected. The sponsor also has an interest in ensuring the study is conducted according to the protocol. Thus, the configuration of the study for each particular study must be verified to be correct. In addition, mechanisms to be supplied to study subjects to obtain patient reported outcomes must also be proven to be set up to collect the data that is intended to be collected. As part of the specification and verification of the system, the sponsor should address how data changes will be known and how these can be reviewed. Without this, data integrity can be in question.

The sponsor is responsible to select qualified investigators and to provide them with the information needed to conduct the trial. This includes ensuring investigators receive adequate training in the use and purpose of the system, how secure access is maintained, how to retrieve and copy records in the event of an inspection, and how records are protected against loss.

The lure of electronic systems is partially that data are available very quickly and can be made available to a large number of individuals. This can result in inappropriate "looks" at the data continuously, which can then lead to changes in study conduct without related protocol changes or statistical penalties. It is critical for the sponsor to ensure that ongoing data analysis is not being conducted by those who do not understand the statistical implications. This may involve limiting who has access to retrieve data or restricting access to reporting tools which can be used to do quick adhoc analyses.

For systems where the sponsor is maintaining a physically separate copy of data that was collected at the site, the sponsor is responsible to ensure it was accurately copied and/or transcribed, and that it is protected against change without authorization from the site. The sponsor is also responsible to ensure that output generated by the system is accurate and reflects what occurred in the study. This is true even if the data "moves" between systems (from initial collection, to a repository, to data manipulation datasets, to final generation of analysis output). In these instances, the sponsor is also responsible to ensure their own staff receives training in the use of the system, as well as the support functions that address continuing system accuracy and data protection. In the end, it is in the sponsor's best interests to be able to demonstrate that the data submitted to the FDA is the same as that collected at the clinical sites. It must be possible to perform this independent verification.

Sponsors are also responsible for adequate monitoring of studies. Study monitors should receive targeted training in the use of electronic source and how this affects traditional monitoring activities and techniques. Although the generation and protection of data is mainly the investigator's responsibility, it is also advisable for the sponsor (likely the monitor) to regularly review the system audit trail, to be aware of data changes or deletions which are occurring.

If CROs are to be used to assist with data collection activities, study monitoring, or supplying system solutions, the sponsor must define the regulatory obligations that are being transferred to the CRO in writing. If CROs are supplying systems, it is key that sponsors remain involved in the process of specifying the system requirements (especially for the specific study) and verifying that the delivered solution is accurate.

CRO

CROs essentially have the same responsibilities as the sponsor organization, for the tasks that they were hired by the sponsor to perform. It is their responsibility to ensure that the transferred

obligations are clearly identified in writing. Related to the use of computer systems, some of the areas which are frequently not well defined include who will perform system testing and who will train clinical sites in the use of systems. When CROs provide system solutions, they are responsible to ensure the systems are specified and function correctly, that the system functions are protected against unimpacted impacts by changes, and that the systems are configured correctly, per the needs of the protocol and other study documents. Although CROs are very customer-oriented, it is their responsibility to ensure that clear specifications and requirements have been received from the sponsor. CROs may also find themselves in a coordinating role between parties, including the sponsor, sites, clinical labs, and other contract organizations that may be involved in data analysis.

Clinical Lab

Although clinical laboratories do not technically fall under FDA regulations, they still maintain a responsibility, when involved in clinical research, to ensure they are generating accurate results. This means not only accurate functioning of laboratory instrumentation and systems used for data collection and analysis, but also protection of data against unauthorized change. Results must be traceable to the individual subjects, and data must be protected against change during transmission to the recipient. The standards that are used in medical practice are not necessarily the same that are expected under regulated research.

Software Vendor

Software vendors are not regulated entities (unless the software is a medical device), but they do have responsibilities to their regulated customers. They are expected to provide accurate software, which has been developed according to a structured methodology, and provide documentation surrounding the verification (testing) of the software. Software vendors are expected to control the final version of the software and to accurately prepare media and installation instructions for distribution to their customers. This includes clearly defining for the customer the environment in which the software should be installed. The vendor should ensure that documentation provided to customers related to software use and support is accurate and matches the current release of the software. Documentation and procedures at the software vendor should be sufficient to ensure the software can be adequately supported, even if the original development staff is no longer there. The software vendor should have processes in place to ensure that changes to the software do not adversely impact correct functioning of the software. The vendor should provide sufficient information to the customer as to the extent of changes, so that the customer can determine any necessary testing that they feel is necessary. The vendor should also communicate to the customer any potential impacts to existing data, due to the structural changes of any upgrade.

Data Hosting Service Provider

Due to the proliferation of electronic data capture (EDC) and electronic patient reported outcome (ePRO) systems, another entity has entered the picture – that of the data hosting service provider. These entities provide a secure environment to store the data collected in clinical trial systems. These entities may be CROs or software vendors, or sometimes even independent entities supplying only this function. Unless the entity is a CRO, these organizations are typically not

covered by regulations. However, there is still a business responsibility to their customers to ensure that data are protected against unauthorized access, change, or loss. These responsibilities are typically demonstrated by technical, physical, and procedural controls to authorize access to the facility, access to servers, ensure equipment is protected against adverse environmental effects, and that data are backed up in case of loss in the primary facility. If copies of records are to be sent offsite, it is expected that the data hosting service provider ensures the data are protected in transit, as well as at the third party location. This may involve additional contractual arrangements. In the event of disruptions or potential data loss, it is necessary for the service provider to notify its customers of the extent of the "damage." Although most of these types of service providers are focused on data protection, data availability is also key in a regulated environment. If a clinical site or sponsor needs to be able to access records quickly in the event of an inspection, availability must be guaranteed, even if this is off-hours for the hosting facility. Although direct inspection of these entities by a regulatory authority may not be currently possible, the regulated entity (the clinical site, in the case of storing eSource) will be held responsible to demonstrate the data's protection and integrity.

FDA

As the regulatory enforcement body, the FDA also carries some responsibility in the equation of clinical research. As technology, approaches, and possibilities evolve, industry looks to the FDA to provide clear and updated guidance on acceptable approaches and exposures of concern. It will be key for FDA to keep up with technology possibilities, in order to develop its own position on what is reasonable, in light of current regulations. Delivering a consistent message during inspections (regardless of the individuals involved) helps industry develop an understanding of expectations and be able to define acceptable solutions.

Appendix 8 – Validation of the Electronic Portion

There are issues of validity and reliability for the eSource data collection mechanism itself. In the world of eSource, we must not forget about the "e" component. Designing, building, delivering, and using an electronic data collection mechanism (EDC) or assessment instrument (ePRO) carries its own concerns in the areas of validity, reliability, and assurances of data integrity. In addition to selecting the correct assessment instrument or system for the needs of the study, mix and wording of questions, screen layout and entry options, and ability to solicit consistent answers, the electronic component must also be designed to meet the requirements and delivered accurately and consistently, or the validity of the instrument or system itself may be in question. The activities of system validation include:

1. Defining study data and/or assessment requirements
2. Designing the technical solution
3. Purchasing, configuring, or building the electronic components
4. Determining how the delivered solution will be verified
5. Conducting the verification activities and resolving discrepancies in expected results
6. Ensuring a controlled deployment
7. Providing end user training and documentation
8. Protecting data against loss or inappropriate changes
9. Ensuring that changes or updates to the software do not impact the correct functioning of the system

Without proper attention, both industry and regulatory agencies may fall into the trap of making assumptions regarding the use of electronic systems. As was found during the several years following the implementation of 21 CFR Part 11 (Electronic Records, Electronic Signatures), assumptions had been made by FDA that industry was validating computer systems as general good practice and had incorporated these processes almost "without thinking" and no longer as an add-on activity. In reality, what was found was that industry had in fact not across the board adopted computer system validation nor was performing it in a general sense. This was evidenced by the significant resistance to Part 11 and the claims of the cost of its implementation. Much of this resistance was a result of needing to find resources to validate systems that were already in place and should have been validated to begin with, but never had been. Since we are still evolving in terms of ensuring adequately validated systems are put in place as regular practice, there needs to be continued attention (both from an industry practice and quality assurance audit, and from an FDA inspection and review perspective) to ensure these activities are occurring.

Another assumption that is often made is that vendor software has been built with a defined methodology, adequate design, thorough testing, and impact assessments and re-testing performed when changes to the software occur. Software of this type is not regulated, thus there are no mechanisms in place to guarantee that these activities are occurring across vendors, other than market evaluation and customer economic pressure. ePRO and EDC vendors are becoming

more aware of the regulatory considerations for this type of software. However, it is still prudent for customers to pre-evaluate these vendors, their practices, support processes, level of testing, and controlled release of software, before the software is purchased. Without this level of attention, the customer is not fully aware of the risk they are assuming by purchasing a particular piece of software. This is not to imply that marketed software is of inferior quality; merely that it is an unknown, and needs customer assessment in each instance.

Looking at the areas of concern with eSource, the following areas are relevant for the computer system component:

Validity

In terms of an ePRO or EDC system, validity can be construed to mean that the system supplied works for its intended purpose. "Intended purpose" is a key concept and means that the goals and requirements for the system and its users must be defined up front, before development and certainly before implementation begin. Characteristics of the study subject population, the clinical site staff, and data to be collected can (and should) be included in the requirements for the system and will impact the validation process.

Electronic patient diaries can invoke an additional set of validation concerns over and above basic software that is used within the confines of a pharmaceutical sponsor company. For instance, is eyesight or finger dexterity a concern for the subject population? If so, a tiny PDA type device may be in appropriate. Is language of concern, where allowances for multiple translations of the instrument must be included? Study subjects in the general population (or even study site staff, for EDC) may not be completely computer literate and may not be patient with slow response time in a web-based application. Is a cell phone type instrument chosen for a diary, and then, during deployment, it is found that the geographic area's cell service is unreliable or non-existent? Is a web-based interface chosen, and then it is found that the system users do not have access to high-speed Internet capabilities? Accessibility may be a factor of the instrument type and the frequency of desired measures. What type of error messages will be most helpful to assist users who have entered incorrect choices? Will error messages force a correct answer to be made, or do they server as warnings, with overrides allowed? These are all important items to define up front, to ensure that the system design addresses these points.

Validity also includes correct implementation of the protocol – are the correct questions asked to solicit the measures or data to be collected? Is the frequency of measurement implemented correctly? For example, if measures are to be taken post-dose every 30 minutes for 3 hours, does the instrument correctly implement the alarms and timepoints? Organizations must ensure that these agreed-to requirements actually do get implemented correctly in the software component of the eSource system.

Correct design of the system becomes even more key for this type of system perhaps than for some systems used in-house at the sponsor organization. Since the user base is wide and unknown at the time of design, it becomes even more key to identify an individual or group within the delivering organization who can accurately represent the user in terms of defining requirements and judging the solution. Oftentimes, clinical data collection systems are designed solely from the sponsor's point of view, and with the end result in mind. While it is still important to define the type of data that must be collected to facilitate analysis, for ePRO it is

even more critical to take user needs and characteristics into account up front, or the use of the system (and thus the accuracy of the data) may be too variable (or not accepted). Planning for how the technical solution will be verified is also a key activity. The validity of the system is based on how it meets its defined requirements. Testing (operating and running through the program code as it is delivered in the user interface, as well as behind the scenes) gives confidence that the solution is correct. Tests should be planned, with defined test data to thoroughly exercise the system, including sample answers and conditions. For example, if there is a user login, attempts should be made to bypass it, as well as entering incorrect passwords multiple times. Testing should include scenarios such as attempting to access prior answers, leaving drop down box fields empty, selecting more than one checkbox, entering an incorrect choice in an IVR system, closing down a web-based session before completion, etc. "Correct" (expected) functioning should work as expected, but the system must be robust enough to handle unexpected or missing entries. As errors in the system are found (e.g. functionality not delivered correctly), these should be tracked and re-tested after resolution. In addition, an important activity is to assess where in the system the fix occurred, as it may be necessary to re-test other previously "working" areas of the code, to ensure no unintended impact occurred. Upon completion of thorough testing, the program code must be protected against change. This typically occurs by securing it (read-only) in some kind of program code repository or directory.

Reliability

A system is reliable if it operates the same all the time and provides consistent results across multiple users. It is also reliable if it is available when needed. For instance, a web-based or IVR system must be available to study subjects or clinical sites when they wish to enter data. If data are to be entered at frequent intervals, down-time of several hours is not acceptable. For these types of implementations, some level of redundant hardware and network capabilities may be necessary to protect against downtime due to failures. If subjects are using a handheld device, they will be asked to periodically upload data, usually via phone lines. Again, the system must be available to respond, or subjects will become frustrated. Or, they may forget to try again later, potentially putting the unprotected data at risk while it still resides only on the remote device. Uploads should work consistently and provide some feedback to the study subject of successful completion (or alternatively of a problem, if this means that the subject will have to re-try the upload). Similarly, sites have a heavy load, with regular patient care as well as potentially multiple clinical trials. Availability of the system is key to its adoption and acceptance.

To be able to be relied on, the system should operate the same at all times, for all users. Controlled installations should be defined, with clear instructions for the installation process and steps to execute to verify the installation worked. For a web system based on one server, this is less of an issue. However, if program code is to be downloaded to handheld devices or site laptops, the same installation must occur on all devices, to ensure the system is the same for all users. If a defective device is returned several months later for replacement, the same program code must be loaded and verified in the same manner. For a web-based system, reliability equates to the system running in the same manner, no matter which browser software the end user has. This compatibility is something that can be addressed during system testing, before deployment.

One of the best documented benefits of ePRO is the ability to determine when the subject made entries into the instrument, to prevent or identify the "parking lot syndrome" – patients completing many or all entries in a diary assessment in the clinic parking lot just before their study visit. In general, contemporaneous entries are considered a necessary quality for clinical trial data. However, the reliability of these entries is only as good as the reliability of the system date/time that is being used to "stamp" the entries. It is critical to ensure that the system date/ and time are synchronized to a recognized defined time source, and that the end user does not have the ability to change the system date/time. There should be a means to determine if the date/time were changed at the system level, and track to when this occurred, who changed it, and why.

A system may be reliable on day one when it is rolled out in a study. However, if changes are made to the protocol, which necessitate changes to the system, or if an upgrade is necessary to the operating environment or security patches must be applied, the system must remain reliable as these changes occur. This speaks to the importance of change control – defining a desired change, carefully assessing the impact of the change on the system, end user, and previously-collected data, and then defining appropriate testing to demonstrate that no unintended impact is introduced by the change. This testing must occur before the changes are rolled out to the operational environment. Controlled implementation is the final stage of change control. Some changes, such as security patches, may need to occur across the entire study all at once. Others, such as changes from protocol amendments, may be rolled out on a site by site basis, as IRB approvals are received. This requires careful tracking and the ability to segregate changes at the site level. Where an IVR or web-based system is used, changes may be made at a central server. But for ePRO systems that use a remote device, there is the added complication of how the software changes will be downloaded to each patient's device. Detailed tracking is required to maintain the configuration management records.

Data Integrity

System validation contributes to the integrity of the data (especially processed data which is presented to reviewers) by ensuring the accuracy of collection (branching of questions, storage of entries, error conditions, and calculations). In addition, other system features can provide strong assistance in ensuring the integrity and trustworthiness of the data.

Clinical study staff must have accurate information, in order to be able to make determinations about continuation of the study, potential arising safety concerns, and perhaps to adjust other study procedures accordingly. This is somewhat of a parallel to the traditional case report form paradigm. Paper case report forms are retained and can always be reviewed to determine what was initially recorded. However, going forward, it is the clinical database or even the derived statistical analysis datasets which are used for decisions and analysis. Thus it becomes critical that these repositories and calculations for derived information are correct. The same is true for ePRO. The original patient data that is collected must not be lost, but the derived information that the clinical investigator reviews must also be proven to have been calculated accurately. System validation should include tests and data representing a variety of conditions to ensure the system results in accurate data.

It has been mentioned that a clinical investigator can review a handwritten diary, with entries over time, and see whether the entries are likely to have been made by the same person (the

study subject), by looking at the handwriting. A similar concept arises with signed CRFs. With an electronic instrument, this concept of visual attribution is lost. Thus it becomes important to have another mechanism to identify and attribute the entries to the study subject. A typical implementation (but perhaps not the only?) is by assigning the subject some type of identification mechanism which can be used to identify themselves to the assessment system. The most common mechanism is likely a unique password or id and password that can only be executed or applied by this subject. By assigning and delivering these in a controlled manner, training subjects to secure these mechanisms and not share them, and implementing procedures to identify the subject and assist when passwords are forgotten, this level of attributability can be addressed. Similarly, biometrics can be used for subject identification. Biometrics can include technologies such as iris scans, fingerprint scans, registering typing speed and cadence, and can prove to be much less likely to be falsified, although probably not a common implementation for ePRO or EDC yet. The goal is to provide a level of comfort that the entries made really were from that study subject or that site staff member.

Another concern with electronic records is providing the assurance that the record today looks as it did at the time it was first recorded. If changes have been made, the reviewer should be able to see what these changes were, when they happened, and who made them (and probably for what reason). There are multiple options for approaching this with ePRO. One might be to not allow changes after the initial recording, even by the study subject themselves. However, it must be recognized that often the original data records are captured in one system (perhaps a PDA type device) and then transferred electronically to another repository (perhaps the clinical data management system), and sometimes even beyond that (statistical analysis datasets). Even though an original or certified copy is retained, these subsequent copies are the ones used for data analysis and submission claims. Thus the protection of the "record" must be maintained, even as the record moves between systems. Another mechanism to help maintain record integrity is to implement an automated audit trail which will record changes as they are made by system users, whether this is the subject, clinical investigator, or other individual. It should be noted that there is regulatory concern over other individuals changing ePRO data, thus systems should be designed to protect against this possibility. The automated audit trail provides a mechanism similar to the single-line crossout, initial, and dating that is done for corrections to paper documents. This provides a trail of what was originally recorded, as well as any changes that occurred.

It should be remembered that delivering an audit trail behind the scenes is not sufficient without the means to review a human understandable representation of the record and all its changes. Vendors are more frequently now providing audit trails with clinical data management systems, but the ability to clearly see all changes to a record via display or report is still extremely variable. Even more rare is the industry practice to periodically review the audit trail, to look for data changes which might be suspect. All three components (capturing a trail of changes or deletions, providing a means to see the trail, and actually reviewing record history) are key provisions to demonstrating record integrity.

As mentioned above, typically clinical data "moves" between multiple systems. This may be true for some ePRO implementations, especially those with direct patient interaction. If data are collected in one system, then transferred to a more secured or centralized repository, not only must the data be protected in each system, but there should be validation of the move, to ensure

accurate mapping of fields between the systems, the same meaning (e.g. is the date yymmdd in one system, and mmddyy in another – leading to misinterpretation of the date?), and no loss of records during the transfer. This validated move can be argued to be providing the certified copy. Typically, other transfers or copies are made at later stages, for the analysis that will be performed for the regulatory submission. In each instance, the same issues must be addressed. In the end, it must be possible to trace the data from its original collection through to the conclusions reached in final analyses. In many instances, this situation of multiple systems is further complicated by the introduction of third parties that the sponsor may have outsourced part of the operation to – clinical research organizations, software vendors, data hosting service providers. The potential for data integrity problems rise when more parties are in the mix, and communication of requirements and validation of processes and systems become even more critical. Development of standards for specific types of data will help a great deal in addressing this problem, as the variety of possibilities for storage would decrease, if a standard approach (such as that provided by CDISC) was adhered to.

Use Validation

Although technically part of system validation, this section more specifically focuses on validation of the use of the system[44]. This means ensuring correct and consistent use and support, in essence the "people" element.

In addition to the stability of the software, reliability also means that the system will work with a variety of end users. This means that the interface should be clear and forgiving, and that appropriate training or end user documentation is available. In many types of ePRO implementations, the instrument "administrator" is not a trained clinical professional, but rather the patient themselves, thus they must fully understand what is expected of them. There are potentially three levels of understanding that must be achieved:

1. Understanding the questions and measures, including any scales which must be marked or selected from.
2. Understanding the expected use, including frequency of entries, whether multiple answers are allowed, what to do if a timepoint is missed.
3. Understanding the instrument itself, including how to initiate a session, how to logon, what to do if a password is forgotten, responding to errors, how to perform a system upload, and how to close down a session.

All three levels of understanding can be impacted by the electronic environment and must be tested even when reliability has previously been tested in a paper format. End user training is often neglected with these types of systems, because they are thought to be "intuitive." Shortcutting the end user training will almost always lead to unexpected entries, frustrated users, and a failure of the overall implementation. A balance needs to be struck between providing some written instructions that the patient can periodically refer to, and providing such a hefty user manual that patients won't read it. A combination can perhaps be provided of a short normal case "cheatsheet" and more detailed addendums for problems or less frequent scenarios.

[44] Use Validation is also referred to as User Acceptance Testing.

For some instruments, the clinical professional may be involved in administering the assessment (or at least with training the study subjects). Site staff are the end users for EDC systems. Although initial training for the first deployment of the system is usually handled in a careful and formal manner, study staff may change over the life of a lengthy study. In this instance, who is responsible to ensure that the new staff receives the same level of training on the system? Is this a controlled consistent process, or does it become a "hand-me-down" word-of-mouth training that, over time, becomes less than what it should be?

Standard operating procedures at the clinical site should be in place to ensure consistency of use across various staff members. Depending on the type of system, in addition to normal use, these procedures should also address device deployment, administration of user access mechanisms (how are ids and passwords assigned and delivered), user training (including patients), and receipt, tracking, and resolution of problems. It is advisable that site staff be the front line to receive and assist patients with any problems encountered, but then also have a defined process if these need to be referred on to the sponsor or software vendor.

Individuals providing system support also need to be appropriately qualified to address support issues for the particular type of system. If technical staff are not the same as the original software developers, they need the appropriate technical skills and knowledge of the software to be able to make changes without incurring unintended adverse effects. This speaks to the need to have accurate current system design documentation, so the support individual can understand the internal structure and interaction of the software components.

Another support function is that of assisting system users with problems. When a system is used internally, a sponsor can have users call the IT Help Desk. But for external users (essentially members of the public), this may not be appropriate. Sometimes a patient or site staff member is calling about a topic that may really be a study issue and not a technology problem. Thus, those with initial contact with the subject should have the expertise to be able to either resolve problems immediately or refer the problem to the appropriate party. Resolution right at the time of first patient contact is preferable. If a patient has to wait for an answer, gets transferred to three other people, or is told "we'll get back to you," this will definitely impact the acceptance of use of the ePRO. One other consideration is the situation when a user needs help with an access mechanism, such as forgetting their password. Support staff must be trained in how to definitively identify the caller and securely deliver the new components, balancing privacy concerns with the need to ensure access is not granted to an imposter. This builds a strong case for the clinical site staff being able to administer access controls for their patients.

As can be seen from the above discussion, the adoption of standards can help facilitate the quick definition and accurate delivery of data collection and storage systems. Eliminating varying ways of storing data improves the ability to accurately copy, represent, and move data between clinical research parties and systems.

Appendix 9 – Source Data Evaluation Report

The Source Data Evaluation Report as recommended within scenario 2 is intended as a vehicle by which the agency can quickly establish the process and controls around source data employed within a given clinical trial. The report need not be a large document, but it does need to provide the necessary information such that the agency can quickly establish trust in the data collected using the processes and systems described.

The report should comprise the following three elements.

1. A description of the system or systems being deployed, the roles played, the interactions between systems to a level of detail sufficient to understand where source data are stored and how source data flows around the system as described in the following section.
2. A description of where source data reside and flows during the process of collecting the data from sites and/or subjects. This section should indicate where the source data are stored and under what circumstances source data migrate from source data store to source data store.
3. An analysis of how the system and processes meet the 12 user requirements. Each user requirement should be addressed in turn, each analysis addressing the technology and processes used to meet the user requirement. Where a user requirement is deemed to be not applicable, the analysis should state this and the reasons why it is not applicable.

Appendix 10 – Good Practices Checklist: Investigator Responsibilities

Based on an analysis of FDA predicate rules, regulations, and ICH Good Clinical Practices, the CDISC eSDI Working Group has identified the following user requirements that an eSource system must fulfill. Study investigators are expected to understand how these requirements are fulfilled and their roles and responsibilities in meeting these requirements:

User Requirement	Met By
The system shall enable the investigator to protect the rights, safety, and welfare of subjects.	The eSource system must be designed to allow the investigator to monitor safety data collected using the system. Confidentiality of patient data must be maintained. ***Investigator responsibility***: 1) Monitor any safety data collected using the eSource system. 2) Maintain security and oversight of pass-codes for devices and web reports in order to maintain confidentiality of patient information.
An instrument used to capture source data shall ensure that the data is captured as specified within the protocol.	The sponsor and vendor designed the eSource application and the configuration of the diary within it to match the protocol specifications. Testing was performed to insure the functionality matched the protocol. ***Investigator responsibility***: Notify the sponsor of any deviations in the eSource design from protocol requirements.
Source data shall be Accurate, Legible, Contemporaneous, Original, Attributable, Complete and Consistent.	The eSource system is programmed and validated to increase accuracy, legibility, completeness, and timeliness of data collection. Attributability needs to be assured by the system (login, username, password etc.). ***Investigator responsibility***: 1) Monitor compliance of patients in completing eSource as scheduled. Monitor completeness of data in a timely fashion. 2) Maintain and oversee security of pass-codes for devices and web reports for their studies. 3) Verify accuracy of date and time stamps applied by the system.
An audit trail shall be maintained as part of the source documents for the original creation and	The eSource system was designed with a full audit trail. ***Investigator responsibility***: 1) Any data changes must be made according to established

subsequent modification of all source data.	SOP(s) using the eSource system software. 2) Review the audit trail.
The storage of source documents shall provide for their ready retrieval.	The source data are stored on a centralized server with retrieval access provided through the web reports. ***Investigator responsibility***: The investigator must understand how to access the source data.
The investigator shall maintain the original source document or a certified copy.	The source data are stored on the device and then certified copy is transferred to server. ***Investigator responsibility***: 1) The investigator must understand how to access the certified copy on the server. 2) The investigator must approve any changes to source data. 3) The investigator must store and maintain the final archival study file with the certified copy of their site's source data.
Source data shall only be modified with the knowledge or approval of the investigator.	Data management SOPs require investigator approval of any data changes. A full audit trail is available to document changes. ***Investigator responsibility***: 1) The investigator must approve any changes to source data. 2) Review the audit trail.
Source documents and data shall be protected from destruction.	The eSource system keeps a duplicate certified copy of the source data in a separate location to protect physical destruction. Backup and security systems are also in place. ***Investigator responsibility***: 1) The investigator must protect the final archival media from destruction for the time period specified by the sponsor and **existing** regulations.
The source document shall allow for accurate copies to be made.	Procedures for producing certified copies have been validated. ***Investigator responsibility***: 1) Ensure the investigator knows how to generate copies.
Source documents shall be protected against unauthorized access.	The eSource system has pass-code security controls to protect against unauthorized access. ***Investigator responsibility***: 1) Maintain and oversee security of pass-codes for devices and web reports.

	2) The investigator must protect the final archival media in their possession from unauthorized access.
The sponsor shall not have exclusive control of a source document.	eSource vendor holds multiple copies of the source data. Any data changes must be approved by the investigator. ***Investigator responsibility***: 1) The investigator must approve any changes to the source data.
The location of source documents and the associated source data shall be clearly identified at all points within the capture process.	***Investigator responsibility***: 1) The investigator should be aware of where the source data are being held during the life of the trial and during the period of source data retention.
When source data are copied, the process used shall ensure that the copy is an exact copy preserving all of the data and metadata of the original.	***Investigator responsibility***: 1) In the electronic world the investigator should be aware that the system in use has been validated for the purposes of clinical research..

End of Document