# Data Storage

May 2007

***Abstract***

The storage of data that is collected during a clinical trial must be carefully planned. This chapter discusses issues that should be considered whether a study's data is stored electronically or on paper. Guidelines for securely storing data are provided, with an emphasis on preventing unauthorized access that could detract from the integrity of a study. Issues concerning passwords, access controls, electronic signatures (including 21 CFR 11), and audit trails are considered. Recommendations for the locking and archival of data at the conclusion of a study are detailed.

## *Introduction*

The secure, efficient and accessible storage of clinical trial data is central to the success of clinical research. Whether collected using validated electronic tools or traditional paper forms, data are often transferred many times during the course of a clinical trial. These transfers occur between an organization's functional groups as well as between companies, contract research organizations (CROs), and regulatory agencies. Hence, the potential for data corruption and version control errors during data storage and transfer is significant and must be minimized to ensure consistency of results and data quality.

## *Scope*

This chapter provides key considerations for the data storage and archival of clinical trial data.

## Minimum Standards

- During the conduct of a clinical trial, store all original data collected (e.g., case report forms and electronic laboratory data) in secured areas such as rooms or file cabinets with controlled access (e.g., locks). These original documents are to be considered part of the audit trail for tracing back to the source data and should be protected and controlled as rigorously as the electronic audit trail of database modifications or backup procedures.

- Document the procedures for granting access to database servers, establishing system controls, and assigning passwords. This process is especially important in a trial where the original data collection is done electronically and no paper backups will exist.

## Best Practices

- Store clinical data in such a way that backup copies can be easily and frequently made. For example, paper documents should be scanned and archived electronically.

- Use open formats for archival, storage, and transport of data (e.g., ASCII, SAS Transport, Portable Document Format (PDF), and the CDISC ODM Model) whenever possible. Adherence to this practice enables current and future access to the data by multiple systems or reviewers.

## Physical Storage

The physical security of original data sources (e.g., case report forms, electronic data files, and other original data documents) should be maintained carefully. Original paper and electronic documents should be warehoused in secure rooms or file cabinets with controlled access. Whenever possible, paper documents should be scanned soon after receipt and archived electronically so that they are included with the backup of other electronic files.

Database servers can be the primary warehouse of clinical data and should be physically secured, and appropriate standard operating procedures (SOPs) should exist to regulate access to them. Direct access to database servers should be restricted to individuals who are responsible for monitoring and backing up the system. All other access to database servers should be

controlled by logical security and should occur across a secure network protected by password access and appropriate system controls.

Special considerations must be given to the physical security of computers that are used for electronic data collection during a clinical trail. Whenever data are entered into a central database using a network connection, the physical security of the central server, most likely hosted by the sponsor or a vendor, is a primary consideration. If any data are stored locally at the study site before being sent to a central server (as is the case with a hybrid or "offline" system), the physical security of the system at the source of data entry is more critical. In either case, care must be taken to ensure the physical and logical security of computers that are used to store clinical data for any period of time.

Passwords and access-permission controls are vital to ensure that only authorized personnel may access study data. An administrator designated by company policy should assign permissions on an as-needed basis. Mechanisms should be implemented to capture and prevent unauthorized attempts to access the system. If such an attempt takes place, the administrator should be notified immediately. A procedure should be established describing best practices for the selection of passwords and the frequency that passwords should be changed. Passwords should never be shared among individuals or study teams. These operating procedures are designed to minimize the opportunity for data corruption via accidental or intentional manipulation of the electronic raw data.

To maintain compliance with Code of Federal Regulations Title 21 Part 11, trials that use electronic data collection and management will necessarily regard a user's authentication (i.e., user name and password) as the user's electronic signature. All data entry and modification should be captured and stored in an audit trail (user name, date and time stamps) that regards the electronic signature as evidence of the user's authority to alter information in the clinical database.

## *Electronic Storage and Warehousing*

In addition to access controls, database design and organization are important considerations for a thorough data storage system. Database validation and associated documentation is the cornerstone of a secure and reliable system.

All database validation and user acceptance documentation should be readily available to the study personnel to ensure that all database functions being performed on a study have been validated for quality and reliability. Additionally, consideration should be given to ensure that project team access to clinical data is sufficient to expedite efficient and high-quality interim reporting, data-metrics evaluation, and safety-reporting requirements (see the Safety Data Management and Reporting chapter and the Measuring Data Quality chapter).

The need for thorough validation and trial database design is even more critical for trials utilizing electronic data collection. If a software malfunction or unintended loss of information occurs, data collected on paper CRFs can be re-entered. Because electronic data collection eliminates paper documents as an intermediary step between the study observations and the database, it is critical that database validation and reliability issues are resolved on the validated system prior to the entry of any study data. As electronic data entry moves closer to the point of patient care, electronic information more often will be the source data and, as such, require protection and secure warehousing.

## Data Archival

Several archival procedures should be followed to ensure that the data are preserved in their raw format. Most importantly upon completion of a study, the database itself should be locked. That is, permissions to further modify the data should be removed from all except the most critical study personnel. A thorough study archive includes all of the following:

- Database design specifications: Documentation of the table definitions used to build the study database and file structure.
- Raw data: The final raw data files preserved within the study database format, and all original data transfers in their raw format.
- Audit trail: A complete electronic audit trail documenting all modifications to a database by date, time, and user identification.
- Final data: Preserved in a standard file format (e.g., ASCII, SAS transport) that can be easily accessed, reviewed, or migrated to another system.

- Original study documents: The original and/or scanned images of all original documents, which may be archived separately in a central records facility if necessary.

- Procedural variation documentation: Memos and relevant information about any variations from standard operating procedures or working practices that occurred during the trial.

- Database closure: Documentation of each database-lock and unlock, describing the time and conditions surrounding those procedures (for additional information, see the Database Closure chapter).

- Site copies of data: May be required for audit purposes; if needed, these copies should be locked, read-only datasets delivered on CD-ROM or a similar storage medium.

## *Recommended Standard Operating Procedures*

In addition to SOPs, please also reference the chapters on Database Validation, Data Entry and Data Processing, and Database Closure. The following SOPs are recommended as a standard for controlling database storage and archival:

- Database validation.

- Database design.

- Database closure (including procedures for unlocking a locked database).

- Storage of original documents both during and after the trial.

- Forms management and e-data management—this procedure should cover shipping and handling of original and/or working copies of relevant study documents; If electronic documents and files are used, the SOP should specifically address file transfer specifications and storage for those files.

- Version/change control for revisions to software.

- Version/change control for revisions to hardware.

- Version/change control for revisions to data.

- Version/change control for revisions to documentation.
- Disaster recovery.
- System controls and security.

It is also advisable for investigational sites to maintain their own SOPs related to physical and logical computer security.

## References

1. US Food and Drug Administration. *Code of Federal Regulations,* Title 21, Volume 1, Part 11. Rockville, MD: US Department of Health and Human Services; 1998.
2. US Food and Drug Administration. *Guidance for Industry: Computerized Systems Used in Clinical Trials.* Rockville, MD: US Department of Health and Human Services; April 1999.
3. US Food and Drug Administration. *Guidance for Industry: Electronic Submissions of Case Report Forms (CRFs), Case Report Tabulations (CRTs) and Data to the Center for Biologics Evaluation and Research.* Rockville, MD: US Department of Health and Human Services; 1998.

## Further Reading

Not applicable.

## Chapter Revision History

| Publication Date | Comments |
|---|---|
| September 2000 | Initial publication. |
| May 2007 | Revised for style, grammar, and clarity. Substance of chapter content unchanged. |