护理杂志
*Journal of Nursing*
Volume 70 Number 12
ISSN: 0047-262X
DOI: 10.5281/zenodo.10301397

# Clinical Data Management Activities During Clinical Trials

**Chaithanya Kolluri [1], Maheshwar Kolluri [2] and Venkat Ramaiyan [3]**

*[1, 2, 3] Department of Clinical Operations, Skill Buddy Inc, USA, 1135 Kildaire Farm Road, Suite 200, Cary, NC 27511. Email: chaithanyak87@gmail.com*

***Abstract***

*Establishing the most powerful treatment methods and developing new drugs is not easy. Most processes and procedures in clinical trials have faced challenges since they need high-quality data for reporting to get approval from regulatory authorities. Equally, there is a need for high accuracy in clinical trials data management procedures to ensure high efficacy in newly developed drugs. The growth of clinical data management in clinical trials is playing its role in clinical development, leading to the parallel growth of startups and procedures related to clinical data management in most clinical trials. In this part, clinical data management plays a crucial role. Start-up, Conduct and Close-out activities help in the smooth processing of data collection, accuracy, and quality.*

***Keywords:*** *Clinical Data Management, Clinical Trials, Study Startups, Procedures, and Conducts Close-Out Activities.*

## I. INTRODUCTION

Clinical trials can be regarded as an important phase of new drug development as they investigate the safety and efficiency of new therapeutics using study-specific protocols. They are conducted in several phases that require highest quality management for success with the use of data to improve the overall outcome of the clinical trials [1]. Hence, clinical trials are highly associated and related or interlinked to data management systems leading to the development of a multidisciplinary area of study-clinical data management (CDM, as an aid to clinical trials. According to Lu and Su [2], CDM is a cross-functional department that works on data collection, cleaning, and management of subjects and study data according to standard procedures to ensure the validity of clinical research data and development. As a result, clinical trial management systems, including software tools are used to manage the clinical trials, technique development, place selections, patient enrollment, data gathering, analysis, and study close-out [2].

### A. Clinical Trials and its Processes

Discovery and drug development take a long duration of research and analysis in clinical trials. Often, clinical trial processes aim to establish whether the investigational new drug (IND) meets the threshold efficacy without severe side effects to the subjects before being released to the market [1,3]. Thus, thousands of molecules are screened during drug development in clinical trials, with only a few approved in the markets. For a new product before releasing it into the market, Murugesan et al. [3] highlight four major phases in clinical trials to ensure maximum efficacy and safety. The first phase, Phase 0, which is also called a human micro-dosing study is carried out on less than 20 patients and aims at the drug's safety. It provides micro-level testing scales before the onset of phase I. In phase I, 20-100 volunteers are involved in investigating the efficacy and safety of the IND. At this stage, 30% of the drugs are dropped, with the remainder moving into phase II [3].

Phase II involves more testing on the efficacy and side effects of the IND. In this phase, the investigators study the maximum tolerated dose and its side effects, tolerance, pharmacodynamics, and pharmacokinetics an IND can work optimally with low side effects. The second phase has randomized studies in clinical trials staking three months to 2 years with volunteers between 100-300.

Volume 70 Number 12
ISSN: 0047-262X
DOI: 10.5281/zenodo.10301397

This extensive study and rigorous checks result in low-quality output as only 33% of total drugs make it to phase III. The third phase will see an exponential rise in volunteer needs as it requires 300-3000 volunteers testing over a period of close to 4 years. It is the lengthiest and most important phase of the clinical trial as it aims at evaluating the efficacy and monitoring the adverse effects of the IND. Equally, it's a pre-marketing stage as it aims to reveal the new performance of the standard drug. Phase II is also randomized and presents a comparative analysis of the drug performance between the new and control drugs and provides a statistical inference of clinical data analysis [4]. The randomization provides an equal chance of understanding the drug amongst the test population – a factor that makes it a costly and rigorous phase, with only 25-30% of drugs making it into Phase IV. In phase IV, the clinical trials teams are given tasks with post-market surveillance studies with several thousands of volunteers (with the condition under study) to understand how the drug works in practical scenarios.

**B. Processes in CDM**

The processes involved in CDM for clinical trials are focused on collecting and validating the research data to store it for further use by researchers, patients, or medical practitioners. Hence, a CDM process enables researchers to evaluate the effectiveness, safety, benefits, and possible risks that may arise from data collection and management. Nouraniet al. [6] identified five common themes in data management in clinical trials: data gathering, technologies used in data safety and privacy management, data quality management, and data management standards, which are crucial in data management. Equally, Krishnankutty et al. [4] outline nine key steps in Table 1.

**C. CDM in Clinical Trials: The role**

Figure 1 shows the clinical Data Management activities during clinical trials. Statistical analysis outputs are generated by Statistical Team and Submissions by Regulatory Affairs team. Usually, clinical trials generate huge data from several processes, such as preparing peer-reviewed journals and approving new treatments and drug analysis. Consequently, the integrity of these data is critical for various stakeholders such as patients, academic researchers, medical practitioners, drug regulatory agencies, and the government. However, the validity of trial data is threatened by several factors, such as alterations or loss, publication bias, incomplete data collection, and the exclusion of some data. With the advancement of scientific research in clinical trials, the demand for quality data has been increasing. As a result, increased demand for better handling of data acquired from clinical trials, especially data from publicly funded research. For instance, phase III clinical trials alone will generate around 3.6 million data sets, which have tripled the dataset generated over the last decade [5]. This fact will call for a very elaborate CDM system.

As with any study, clinical trials financial and human resources are limited. Having a strong data management plan is crucial. It should include an outline for the research personnel, resources, and storage. A clinical trial is a large scale investment of time, people, and money. It requires subject matter experts (SMEs) management from the first level of study. Some key contributors to this growth are intensive drug development projects targeting rare diseases, using biomarkers and genetic data, and the growing number of patient data sources, from web-based questionnaires to paper-based accessible questionnaires. No matter the cause, clinical researchers need to use this information and make the most of it as an area that CDM is brought in to address.

The CDM, being an essential part of clinical trials, begins with identifying the objectives, and hence every step should be planned, keeping each deliverable at every stage well-defined. However, a slight variation in final deliverables exists between the clinical trials and CDMs [4]. While CDMs deliver valid, statistically sound, and error-free databases, clinical trials are designed to answer a specific research question. Hence, these slight variations make them very different according to the procedure and become a subset of each other.

护理杂志
*Journal of Nursing*
Volume 70 Number 12
ISSN: 0047-262X
DOI: 10.5281/zenodo.10301397

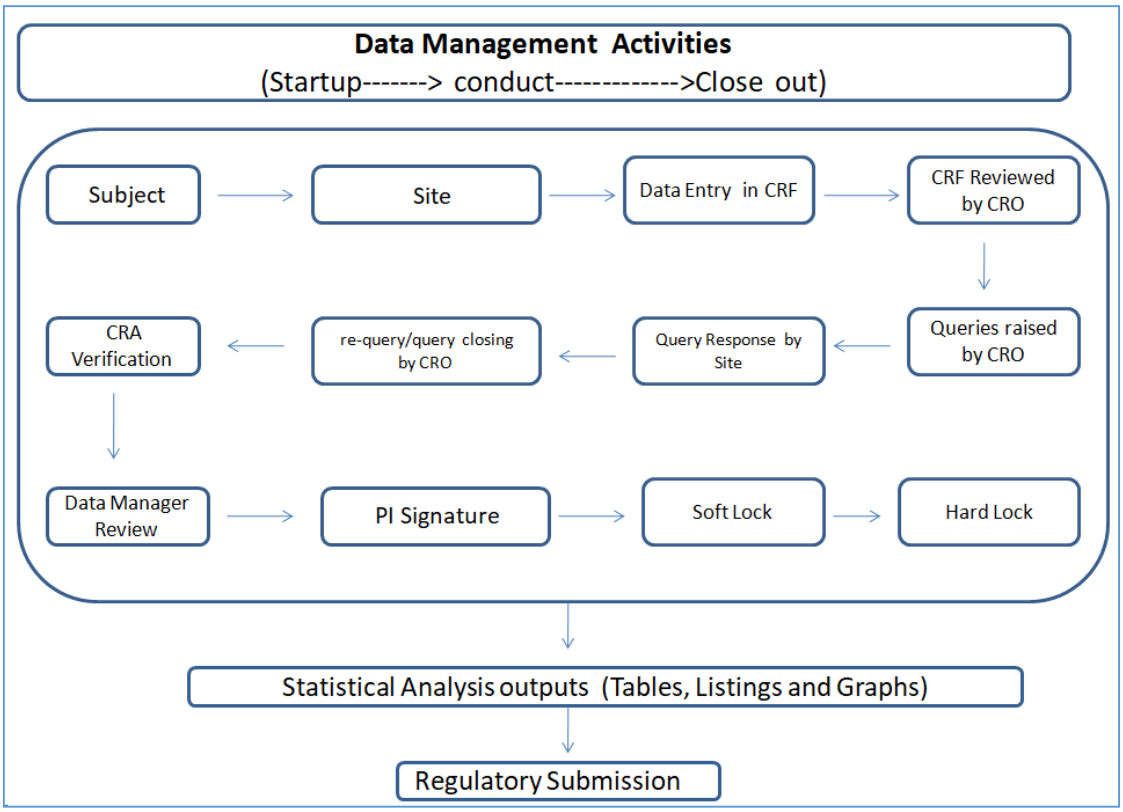


**Figure1: CDM as Part of Clinical Trials**

**Table1: Activities Involved in Clinical CDMs**

| Process /Steps | Activities and Descriptions |
|---|---|
| **Data Collections** | Extensive use of data response to build data points through in-house or electronic data entry. In modern world drug development, e-CRF is deployed to speed up operations [4]. |
| **Case Report Form (CRF) tracking** | Monitoring of CRF or eCRF by clinical Research associates for completeness and follow-ups. |
| **CRF Annotation** | Identification of dataset, variables names, and notes in CRFs |
| **Database Designs** | Defining the objectives, intervals, visits, investigators, sites, patients, CRF layouts, syntax, security codes, and standards to govern information storage, writing, and retrieval in CDM units [1, 3, and 5]. |
| **Data entry** | It's governed by the data management plan (DMP). A double data entry by two operators is performed to help in verifications and reconciliation by identifying possible errors in paper CRFs [4]. |
| **Medical coding** | Assigning codes to variables and parameters for better and standardized operation across the CDM teams and research groups. |
| **Data validation** | Testing the validity of data according to the protocol specification. Such protocols include; Good Clinical Data Management Practices (GCDMP) guidelines, Study Data Tabulation Model Implementation Guide for Human Clinical Trials (SDTMIG), and the Clinical Data Acquisition Standards Harmonization (CDASH) standards [4]. |
| **Discrepancy Management** | It entails cleaning the data to provide concrete proof of concepts from data deviation. The basic process includes discrepancy review, reason investigation, and declaring the documentary irresolvable [3, 5]. |
| **Database lock** | Database locks to reduce data alteration after discrepancy management to allow statistical analysis. |

护理杂志
*Journal of Nursing*
Volume 70 Number 12
ISSN: 0047-262X
DOI: 10.5281/zenodo.10301397

However, clinical data should be specially prepared before sharing due to privacy concerns. Clinical data sharing acts as a means to honor the kindness of clinical trial participants through the increased utilization of the information they make available.

## II. THE STUDY START-UP OVERVIEW

A study startup activity in clinical data management consists of crucial tasks that set the foundation for the entire clinical trial. These activities include planning, preparing, and establishing processes to ensure smooth data management throughout the study.

### A. Database Designs

The database architecture establishes a structured and organized system for storing, managing, and retrieving clinical trial data. This process includes establishing the study's aims and deadlines, choosing when and where data will be gathered, identifying who will be involved (such as investigators and patients), and designing the form and structure of the Case Report Forms (CRFs) used for data input [7,8]. A key recent trend is the growing use of cloud-based databases, which provide scalability, ease of access, and cost-effectiveness. A well-designed database provides good data quality and compliance with regulatory criteria throughout the trial [9]. Data managers and database administrators are critical in this activity, taking responsibility for database setup and maintenance to ensure trial success.

### B. CRF Annotation

CRF annotation is an essential task in clinical data management. It involves identifying and labeling different sections of information in Case Report Forms (CRFs). Electronic CRFs (e-CRFs) are becoming more popular as they offer advantages like real-time data entry and remote access [10]. Proper CRF annotation is crucial because it ensures that data is collected consistently and accurately, making it easier for data entry and validation processes. Clinical research associates and data managers are the key players responsible for performing these activities and ensuring the CRFs are adequately annotated and organized for a successful clinical trial.

### C. Data mapping

Data mapping is a critical aspect of clinical data management that involves assigning standardized codes to variables and parameters to ensure consistent data representation in the database. Using uniform codes allows data from various sources to be integrated seamlessly, enhancing data interoperability [11]. A significant trend in data mapping is the adoption of CDISC (Clinical Data Interchange Standards Consortium) standards, such as SDTM (Study Data Tabulation Model), which further improves data integration and exchange. The importance of data mapping lies in its ability to maintain consistency and standardization, enabling smooth data communication between different systems and stakeholders involved in the clinical trial[7,9,11]. Data managers and clinical data coordinators play critical roles in mapping, ensuring that data is accurately and uniformly represented for successful clinical data management. For similar studies, data collection should be similar to avoid data mapping issues and save time.

### D. Study Site Selection and Training

Study site selection and training are crucial steps in the preparation and execution of clinical trials. The activities involved in this process include identifying and carefully choosing suitable clinical trial sites and training investigators and site personnel on data collection and management procedures. To keep up with the changing sites, virtual site selection visits and remote training sessions have become more prevalent, particularly in global trials where physical visits may only sometimes be

Volume 70 Number 12
ISSN: 0047-262X
DOI: 10.5281/zenodo.10301397

feasible [12]. The importance of these activities cannot be overstated, as they play a significant role in ensuring consistent data collection and adherence to study protocols across all selected sites. Clinical Project Managers and Clinical Research Associates play important role in the site selection and training, overseeing and coordinating these efforts to ensure the smooth conduct of the clinical trial and the collection of high-quality data from each participating site.

**E. Data Management Planning**

Creating a detailed document outlining the strategies, methods, and responsibilities for different data-related activities during the clinical trial is a part of the strategy. The rising emphasis on risk-based data management measures, including risk-based monitoring and data quality management to proactively detect and resolve possible concerns, is an emerging trend in DMP development [13]. A well-defined DMP is critical because it ensures uniform data management methods, regulatory compliance, and high-quality data for reliable analysis and reporting. Data Managers and Clinical Research Coordinators are the primary personnel in establishing and executing the DMP and ensuring that data is efficiently handled and that defined criteria are followed throughout the clinical trial. Their contributions are critical in preserving data integrity and dependability.

**F. Data Privacy and Security Measures**

Data privacy and security measures in clinical data management include deploying processes and protections to secure patient data while following data protection legislation such as GDPR and HIPAA. One of the rising trends in this field is a greater emphasis on cybersecurity and data encryption, which is critical in protecting sensitive clinical trial data from possible attackers [14]. Data privacy and security are critical for maintaining participant identity and preventing unauthorized access or breaches that might compromise the study's integrity. Data managers and security professionals create and maintain these safeguards throughout the clinical trial process.

**G. Data Collection**

Data collection is fundamental to clinical trials, gathering essential data from multiple sources. These sources include Case Report Forms (CRFs), electronic data capture (EDC) systems, medical devices, laboratories, and patient-reported outcomes (PROs). A recent growing trend in data collection involves the adoption of handy and mobile health technologies. This innovative approach allows for remote data collection, facilitating real-time monitoring of patient’s health status and treatment response while reducing the burden on participants [15].

Accurate and timely data collection is essential in the clinical trial process. The data collected is the foundation for generating reliable study results and is crucial in making informed decisions throughout the trial duration. Properly managed and meticulously documented data ensures the integrity and validity of study results.

To carry out an effective data collection, various roles come into play. Clinical Research Coordinators oversee the entire data collection process, coordinate activities, and ensure adherence to study protocols. Simultaneously, data entry Operators play a vital role in accurately inputting the collected data into designated databases or CRFs, providing the data's precision and reliability. A clinical data manager ensures the collected data is as per the protocol.

**H. Technologies, Standards, and Software for Study Startup Activities**

In the study startup activities in clinical data management, various technologies and software solutions have emerged to enhance efficiency, data consistency, and collaboration:

护理杂志
Journal of Nursing
Volume 70 Number 12
ISSN: 0047-262X
DOI: 10.5281/zenodo.10301397

***Electronic Data Capture (EDC) Systems***: Leading EDC platforms like Medidata Rave, Inform, Medrio, IBM clinical development, and Veeva Vault have revolutionized data collection by offering streamlined processes and real-time data entry capabilities. These systems enable errorless data capture and monitoring, reducing data entry errors, and speeding up study timelines.

***CDISC Standards*:** CDISC standards guide data consistency and interoperability, including the Study Data Tabulation Model (SDTM) and Clinical Data Acquisition Standards Harmonization (CDASH). By adopting these standards, clinical trials can ensure that data is collected, stored, and analyzed in a standardized format, enhancing data quality and facilitating data transfer between stakeholders.

*Cloud-based Databases*: Cloud-based solutions, such as Amazon Web Services (AWS) and Microsoft Azure, provide secure and scalable data storage options for clinical trials. Cloud-based databases offer the advantage of accessibility from anywhere with an internet connection, allowing global teams to collaborate seamlessly and facilitating data sharing while maintaining data security.

***Remote Training and Collaboration Tools***: With the rise of remote and decentralized clinical trials, virtual meeting platforms and e-learning tools have become unavoidable for study startup activities. These tools enable remote training for investigators and site staff, ensuring a proper understanding of study protocols and data collection procedures. Even more than that, virtual collaboration tools increase communication and coordination among study stakeholders across different geographic locations.

## III. STUDY CONDUCT

The study conduct improves procedure efficiency and maximizes data validation with the aim of improving clinical trial output. Most of the study conduct activities overlap with those involved in start-up, though they are more oriented to serious adverse events (SAE), data cleaning controlled by clinical trial data managers.

### A. CRF Tracking

CRF tracking involves closely monitoring the progress and completeness of Case Report Forms (CRFs) throughout the study. This activity ensures that all required data points are recorded and that any missing data is promptly addressed. The significance of CRF tracking lies in its crucial role in maintaining data integrity and completeness, enabling accurate analysis and reporting of study results. The trend toward using electronic data capture (EDC) systems and mobile-enabled solutions facilitates real-time tracking and data collection, improving efficiency and data accuracy. Clinical Research Associates (CRAs) and Data Managers oversee CRF tracking [16]**.**

Electronic Data Capture (EDC) Systems are crucial in modern clinical data management. These systems facilitate electronic data entry, tracking, and management of data collected during the trial. Prominent EDC systems include Medidata Rave, Phase Forward inform, Medrio, IBM Clinical Development, and Veeva Vault EDC. Data Managers and Clinical Research Associates (CRAs) use EDC systems effectively to ensure accurate and efficient data collection and monitoring.

### B. Data Entry or Data Transfer from the CRF to the CDMS

Data entry involves accurately transferring data from the CRF to the Clinical Data Management System (CDMS) to create a digital repository of study data. Accurate data entry is crucial as it ensures that the data in the CDMS accurately reflects the information collected on the CRFs, supporting reliable analysis and reporting [17]. The increasing adoption of EDC systems and automated data entry tools streamlines the data transfer process, resulting in faster and more accurate data transfer. Data

护理杂志
*Journal of Nursing*
Volume 70 Number 12
ISSN: 0047-262X
DOI: 10.5281/zenodo.10301397

Managers and Data Entry Specialists play crucial roles in ensuring the accuracy and completeness of data in the CDMS.

### C. Discrepancy Management

Discrepancy management involves identifying and resolving data discrepancies that may arise between the CRF and the CDMS. This activity is essential to ensure data accuracy and consistency, enhancing the overall data quality for analysis and reporting [18]. The trend of using advanced query management systems and automated discrepancy resolution tools enhances the efficiency of discrepancy management. Data Managers and Clinical Research Associates, collaborate to identify and resolve discrepancies effectively**.**

Noted tools in this category are IBM Clinical Development and Medrio. Data Managers and Clinical Research Associates collaborate to use these tools for an efficient and timely resolution of data discrepancies.

### D. Data Coding

Data coding involves assigning standardized codes to variables and parameters, ensuring uniform data representation for analysis and reporting. Proper data coding is crucial as it facilitates data analysis and enables easy comparison across different studies or databases. Adopting automated coding tools and natural language processing (NLP) streamlines and enhances the coding process [19]. Medical Coders and Data Managers play pivotal roles in data coding to ensure consistency and accuracy.

Data coding is made more efficient and consistent through automated coding tools. Examples include ClinPlus Coding and OpenClinica Code Book. Medical Coders and Data Managers use these tools to assign standardized codes to variables and parameters, ensuring uniform data representation for analysis and reporting. Coding dictionaries used in medical coding are MedDRA for adverse events and medical history and WHODD for concomitant medications and procedures.

### E. Data Review and Ongoing Quality Control:

Data review and ongoing quality control include regular data checks for accuracy, completeness, and consistency throughout the study. This ongoing quality control ensures data accuracy and adherence to study protocols, enhancing data reliability and validity. Data visualization and analytics tools enable real-time monitoring and automated quality control checks [20]. Data Managers and Quality Control Specialists collaborate to conduct thorough data reviews and quality control checks.

Data visualization and analytics tools are essential for exploring and interpreting study data. Spotfire, Qlik, R Shiny, and Python with data visualization libraries are popular choices in this category. Data Managers and Analysts leverage these tools to create interactive and informative visualizations, aiding in real-time monitoring and decision-making.

### F. Data Transfer

Data transfer refers to the secure movement of data between different systems or locations, such as from clinical sites to the central database. This process is essential to preserve data integrity and prevent data loss during the transfer. The utilization of secure and encrypted data transfer protocols ensures enhanced data security. Data Managers and IT Specialists oversee data transfer activities [21].

Secure data transfer tools are critical for moving data between different systems or locations while ensuring data integrity and confidentiality [22]. Noted tools include Aspera, FileCatalyst, and Accellion. Data Managers and IT Specialists, use these tools to securely transfer study data, especially when exchanging sensitive or large datasets.

护理杂志
*Journal of Nursing*
Volume 70 Number 12
ISSN: 0047-262X
DOI: 10.5281/zenodo.10301397

### G. Data Import Protocols

Data import protocols outline the procedures and guidelines for transferring data from external sources (e.g., laboratories and imaging centers) into the study database. Proper data import ensures the accurate and efficient integration of external data into the study database for analysis. [22]. Automated data import tools like SAS Data Integration Studio, Talend, and Informatica enable seamless data integration into the study database from external sources. Data Managers and IT Specialists implement these tools to ensure the accurate and efficient integration of external data for analysis.

### H. Sponsor Submissions

Sponsor submissions include preparing and submitting study data and safety reports to regulatory authorities or ethics committees. Accurate and timely sponsor submissions are crucial for regulatory compliance and the progression of the clinical trial [23]. The use of electronic submission tools expedites the submission process. Data Managers and Regulatory Affairs Specialists are essential in preparing and submitting study data to regulatory authorities. DSUR (Development Safety Update Report), PSUR (Periodic Safety Update Report) help in ongoing safety reports.

Regulatory submission software is crucial for preparing and submitting study data and safety reports to regulatory authorities or ethics committees. Prominent examples include Liquent Insight, EXTEDO eCTD Manager, and ISI Toolbox. Data Managers and Regulatory Affairs Specialists use these tools to ensure accurate and timely sponsor submissions for regulatory compliance.

## IV. CLOSE UP ACTIVITIES

Close-up activities are after a clinical trial process that ensures data is ready for statistical analysis. This activity reduces alteration or changes during analysis.

### A. Database Lock

Database lock is a pivotal step in the study closeout process, involving the finalization of the database once all data collection, cleaning, and validation activities are completed. Once the database is locked, no further changes can be made to the data, ensuring its integrity and consistency. A notable trend in this process is the adoption of electronic signatures and blockchain technology to enhance data security, providing an extra layer of protection against unauthorized alterations [24].

Database lock holds immense importance as it marks the official end of data collection and allows for the commencement of data analysis, the generation of comprehensive study reports, and the preparation for regulatory submissions. Data Managers and Clinical Research Associates are actively involved in managing the database lock process, ensuring that all data is accurately and securely locked for future analysis and reporting.

### B. Data Archiving

Data archiving is a critical aspect of the study closeout activity, encompassing the secure storage of all study-related documents and data to ensure long-term preservation and compliance with regulatory requirements. Cloud-based data storage solutions are gaining popularity as they offer convenient and secure archiving options, enabling easy access to archived data when needed [25]. Proper data archiving is of utmost importance as it guarantees data integrity, traceability, and accessibility for potential future audits or inspections. Archiving Specialists and Data Managers work collaboratively to organize and preserve the study data and documents systematically, ensuring that valuable information remains accessible and well-protected for future reference or regulatory purposes.

护理杂志
*Journal of Nursing*
Volume 70 Number 12
ISSN: 0047-262X
DOI: 10.5281/zenodo.10301397

Cloud-based archiving services, such as Amazon S3 Glacier and Microsoft Azure Archive Storage, offer secure, scalable, and cost-effective options for data storage during study closeout. These solutions provide reliable long-term preservation of study-related documents and data while ensuring data integrity and accessibility. Cloud-based archiving also enables easy retrieval of archived data whenever needed for audits, regulatory submissions, or future reference. Organizations can optimize their data storage and management processes during study closeout by leveraging cloud-based data archiving solutions.

### C. Study Report Preparation

Generating comprehensive study reports is crucial during the study closeout phase, summarizing the trial's results, findings, and conclusions. A prominent trend in this process is using data visualization and interactive reporting tools, enhancing the impact and user-friendliness of the study reports. The study reports are essential documentation for regulatory submissions, scientific publications, and references in future research endeavors [26]. Medical Writers and Data Managers collaborate to create detailed and accurate study reports, presenting the trial's outcomes effectively and transparently.

Study reporting tools, such as SAS or R for statistical analysis and report generation, are essential to analyze and summarize the trial's results and findings during the closeout phase. These tools allow data managers and medical writers to conduct robust statistical analyses, generate comprehensive study reports, and present the data clearly and concisely. Additionally, data visualization tools like Tableau or Power BI are utilized to create interactive and informative reports, enhancing the visualization and understanding of study data for various stakeholders. These reporting tools significantly contribute to the effective communication of study outcomes and support regulatory submissions and publication of study results.

### D. Audit and Inspection Readiness

Audit and inspection readiness is vital in the study closeout activity, ensuring that all study-related documents and data is meticulously organized, complete, and readily available for potential audits or regulatory inspections [27]. There is an increasing focus on continuous audit readiness throughout the trial's duration, requiring ongoing diligence in maintaining the study's documentation and adherence to Good Clinical Practice (GCP) guidelines. Being audit-ready is paramount, as it demonstrates the trial's compliance with ethical and regulatory standards, enhancing its credibility and reliability. Quality Assurance Specialists and Data Managers collaborate to achieve and maintain audit readiness, ensuring that the study is conducted compliant and ethically.

Electronic Trial Master File (eTMF) systems like Veeva Vault TMF or Phlexglobal's eTMF Connect are invaluable platforms that streamline document management and ensure inspection readiness during the study closeout activity. These systems enable the organization and maintenance of study-related documents, ensuring that all essential documents are complete, well-organized, and readily accessible for potential audits or regulatory inspections. eTMF platforms also facilitate collaboration between teams involved in the closeout process, allowing them to work efficiently and proactively in achieving audit readiness.

### E. Data Sharing and Disclosure

In the study closeout activity, data sharing and disclosure are carried out if applicable, preparing to share study data with relevant stakeholders, including regulatory authorities, study sponsors, and the scientific community. An emerging trend is a growing emphasis on their transparency and open data sharing in clinical research, fostering scientific advancements resulting from reproducibility and further research opportunities [28]. Proper data sharing ensures that the valuable insights gained from

Volume 70 Number 12
ISSN: 0047-262X
DOI: 10.5281/zenodo.10301397

the trial are disseminated to contribute to the broader scientific knowledge. Data Managers and Clinical Research Associates work together to prepare and coordinate the data-sharing process in compliance with relevant regulations and agreements, promoting transparency and supporting scientific progress.

## V. CONCLUSION

The clinical trials rely on CDM for clinical data management activities. In the modern world of clinical research, CDM proves invaluable and hence, more technology investment should be carried out targeting CDM to improve the accuracy and efficiency of the new drugs. However, to improve the development path, it's important to review individual processes of clinical trials to enable researchers to develop process-specific solutions in clinical trials. One of the most intriguing trends is the growth of databases for data management and cloud computing for use in CDM. Of these two, there is a need for improved security in data handling especially in close-out activities where data locks and discrepancy management become critical. Here, security is of great importance in protecting data to increase the credibility of their research, improve data integrity, and safeguard the patients. Hence, there is a need for technology restructuring in start-up, conducts, and close-out activities.

## VI. Abbreviations

| | |
|---|---|
| CDM: | Clinical Data Management |
| CRF : | Case Report Form |
| IND : | Investigational New Drug |
| SME : | Subject Matter Expert |
| DMP : | Data Management Plan |
| SDTM : | Study Data Tabulation Model |
| CDASH: | Clinical Data Acquisition Standards Harmonization |
| CDISC: | Clinical Data Interchange Standards Consortium |
| PRO : | Patient-Reported Outcomes |
| EDC : | Electronic Data Capture |
| AWS : | Amazon Web Services |
| SAE : | Serious Adverse Events |
| MedDRA: | Medical Dictionary for Regulatory Activities |
| WHODD: | World Health Organization Drug Dictionary |
| SAS : | Statistical Analysis System |
| ETMF : | Electronic Trial Master File. |

Volume 70 Number 12
ISSN: 0047-262X
DOI: 10.5281/zenodo.10301397

## References

1) Y. Zhuang, L. Zhang, X. GAO, Z.-Y. Shae, J. J. P. Tsai, P. Li, and C.-R. Shyu, "Re-engineering a Clinical Trial Management System Using Blockchain Technology: System Design, Development, and Case Studies," Journal of Medical Internet Research, vol. 26, no. 4, p. e36774, 2022.

2) Z. Lu and Su, "Clinical Data Management: Current status, challenges, and future directions from industry perspectives," *Open Access Journal of Clinical Trials*, p. 93, 2010. doi:10.2147/object.s8172

3) D. Murugesan, R. Banerjee, and G. Ramesh Kumar, "Clinical Data Management Importance in Clinical Research," Asian Journal of Pharmaceutical and Clinical Research, p. 59, 2016. doi:10.22159/ajpcr.2016.v9s213940

4) B.Krishnankutty, B. Naveen Kumar, L. Moodahadu, and S. Bellary, "Data Management in Clinical Research: An overview," Indian Journal of Pharmacology, vol. 44, no. 2, p. 168, 2012. doi:10.4103/0253-7613.93842

5) Tufts CSDD Impact Report, "Rising protocol design complexity is driving rapid growth in clinical trial data volume.," Tufts Center for Study on Drug Development, vol. 23, no. 1, 2021

6) A.Nourani, H. Ayatollahi, and M. S. Dodaran, "Data management in diabetes clinical trials: a qualitative study," Trials, vol. 23, no. 187, 2022.

7) Society for Clinical Data Management, Good Clinical Data Management Practices, 2013

8) R. Deari, X. Zenuni, J. Ajdari, F. Ismaili, and B. Raufi, "Analysis and comparison of document-based databases with relational databases: MongoDB VS MySQL," 2018 International Conference on Information Technologies (InfoTech), 2018. doi:10.1109/infotech.2018.8510719

9) J.M. Larson, "Account-and Database-Level Privileges," in Snowflake Access Control: Mastering the Features for Data Privacy and Regulatory Compliance, pp. 123-127, Berkeley, CA: Apress, 2022.

10) A. Pandey, Surachna, and S. S. Mishra, "Applications for data collection: Appdatcol," International Journal Of Community Medicine And Public Health, vol. 8, no. 11, p. 5551, 2021. doi:10.18203/2394-6040.ijcmph20214296

11) E. Scheurwegs, K. Luyckx, L. Luyten, B. Goethals, and W. Daelemans, "Assigning clinical codes with data-driven concept representation on Dutch clinical free text," Journal of Biomedical Informatics, vol. 69, pp. 118–127, 2017. doi:10.1016/j.jbi.2017.04.007

12) B. Wright, "Setting up of the site, site assessment visits, and selection," A Comprehensive and Practical Guide to Clinical Trials, pp. 17–26, 2017. doi:10.1016/b978-0-12-804729-3.00003-1

13) P. G. Stokman, "Risk-based quality management in CDMan inquiry into the value of generalized query-based data cleaning," Journal of the Society for Clinical Data Management, vol. 1, no. 1, 2021. doi:10.47912/jscdm.20

14) D. Hallinan, "Testing the GDPR about Biobanking," Protecting Genetic Privacy in Biobanking through Data Protection Law, pp. 129–147, 2021. doi:10.1093/oso/9780192896476.003.0007

Volume 70 Number 12
ISSN: 0047-262X
DOI: 10.5281/zenodo.10301397

15) P.Dias, M. Cardoso, F. Guede-Fernandez, A. Martins, and A. Londral, "Remote Patient Monitoring Systems based on conversational agents for Health Data Collection," Proceedings of the 15th International Joint Conference on Biomedical Engineering Systems and Technologies, 2022. doi:10.5220/0011011000003123

16) H. Donnelly et al., "Streamlining Safety Data Collection in hospital-acquired bacterial Pneumonia and ventilator-associated bacterial pneumonia trials: Recommendations of the clinical trials transformation initiative antibacterial drug development project team," Clinical Infectious Diseases, vol. 63, no. Suppl 2, 2016. doi:10.1093/cid/ciw316

17) Y.Bochud et al., Clinical Data Management System (CDMS) evaluation, 2021. doi:10.54920/site.2021.01

18) T.Koltay, "Data quality, the essential 'ingredient,'" Research Data Management and Data Literacies, pp. 49–75, 2022. doi:10.1016/b978-0-12-824475-3.00004-7

19) P. Adu, "Manually assigning codes to data," A Step-by-Step Guide to Qualitative Data Coding, pp. 89–120, 2019. doi:10.4324/9781351044516-7

20) S. Gupta, "Overview of data quality and compliance checks," Clinical Data Quality Checks for CDISC Compliance Using SAS, pp. 1–16, 2019. doi:10.1201/9780429345043-1

21) Y. Yu, "Data integrity checking with reliable data transfer for secure cloud storage," International Journal of Web and Grid Services, vol. 14, no. 1, p. 1, 2018. doi:10.1504/jugs.2018.10006680

22) Q. Li, "Get data ready: Import, inspect, and prepare data," Using R for Data Analysis in Social Sciences, pp. 43–93, 2018. doi:10.1093/oso/9780190656218.003.0002

23) M.A. Malikova, "Practical applications of regulatory requirements for signal detection and communications in Pharmacovigilance," Therapeutic Advances in Drug Safety, vol. 11, p. 204209862090961, 2020. doi:10.1177/2042098620909614

24) M.A. Malikova, "Practical applications of regulatory requirements for signal detection and communications in Pharmacovigilance," Therapeutic Advances in Drug Safety, vol. 11, p. 204209862090961, 2020. doi:10.1177/2042098620909614

25) S. Wu and Y. Zhang, "Secure cloud storage using Anonymous and black box traceable Data Access Control," Security and Communication Networks, vol. 8, no. 18, pp. 4308–4318, 2015. doi:10.1002/sec.1343

26) R. Nayak, N. Nawar, J. Webb, M. Fatemi, and A. Alizad, "Impact of imaging cross-section on visualization of thyroid microvessels using ultrasound: Pilot study," Scientific Reports, vol. 10, no. 1, 2020. doi:10.1038/s41598-019-57330-w

27) A. Sheth and J. V. Sinfield, Synthesis study: Overview of readily available culvert inspection technologies, 2019. doi:10.5703/1288284316789

28) D.R. Donaldson and J. W. Koepke, "A focus groups study on data sharing and Research Data Management," Scientific Data, vol. 9, no. 1, 2022. Doi: 10.1038/s41597-022-01428-w.