The Clinical Data Management Process

Efficient collecting, transferring, and managing of data during a clinical trial is important and helps advance the drug development process.

BY ARON SHAPIRO AND DALE USNER, PHD

Approval of new retina drugs by the US Food and Drug Administration (FDA) is contingent upon trust and belief that clinical trial data are of sufficient integrity.[1] From database build to database lock, clinical data management (CDM), the process of collecting, cleaning, and managing subject and/or trial data in compliance with regulatory standards,[2] is a crucial component of any clinical trial. When done efficiently, CDM of retina trials leads to the generation of statistically sound data and accelerates the drug development process.

KEY TERMS	
CDM	Clinical Data Management
CRF	Case Report Form
EDC	Electronic Data Capture
DMP	Data Management Plan
DVM	Data Validation Manual
CCG	CRF Completion Guideline

CDM FROM THE BEGINNING

From a business perspective, drug developers want to ensure that the data delivered to regulatory bodies is reliable; from an ethical perspective, clinical data inform treatment decisions and ultimately affect patient health. For both of these reasons, clinical data quality and integrity are crucial. Although it seems obvious that data management only happens after the data are collected, the process actually starts before the study protocol is finalized.

Previous columns in this section have reviewed guidelines for drawing up a protocol (see "Guidelines for Writing Protocols" in the October 2014 issue of *Retina Today*). Because CDM is best viewed as a process that weaves throughout the entire clinical study, it's important to involve the CDM team from the beginning, starting with protocol development.

THE CDM PROCESS

A Case Report Form (CRF) is designed by the CDM team for data collection from protocol-specific activities. The CRF may exist in either a paper version or as electronic data capture (EDC). The CRF will be annotated with coded terms to communicate where the data collected for each question is to be stored in the database.

The next item developed is the Data Management Plan (DMP), which details how the data are to be handled according to how the study is anticipated to be run. A DMP describes the CDM activities to be followed in the trial, including trial master file maintenance, CRF specification, database design, data collection, CRF tracking, data entry, data storage and privacy, medical coding, data reconciliation (eg, serious adverse events and central laboratory data), data review, discrepancy management, data extraction, and database lock. The DMP is intended to standardize procedures and ensure that all CDM personnel understand the plan.

Next, a Data Validation Manual (DVM) is developed. This document contains the edit check programs for discrepancy identification.[2] The edit check programs in the DVM help the site, monitoring, and CDM staff identify any discrepancies during data cleanup. Entry of data that are not validated may prompt a request for clarification. This process, called *discrepancy management* or *query resolution*, is put into place to investigate the reason for the discrepancies; ideally, discrepancies should be resolved quickly.[2] Queries slow the progress of data analysis, so closing queries promptly helps to ensure that the trial continues to hit timelines without ignoring data points or losing subjects. EDC systems must capture any change to data after it has been saved and all discrepancies that are generated in an audit trail. Discrepancies should be reviewed at regular intervals by the CDM team to ensure that they are being resolved in a timely manner.

Coding of all medical terms reported allows standard-

ization when it comes time for the data to be analyzed and reviewed.[3] Information on adverse events, medical history, and concomitant medications must be coded in a uniform manner. This is especially important in multicenter trials in which multiple investigators collect and report information.

Finally, a database lock is put in place after all data management activities are complete to ensure there was no manipulation of study data after unmasking of the treatment groups and during the final analysis. A prelock checklist should be used to confirm that all necessary activities have been completed.

CDM LEADERSHIP

As with any team, there is always a leader. The CDM team is led by a data manager who is responsible for supervising the entire CDM process and coordinating data management activities.[2] The data manager is a key player in early discussions about data collection options. The data manager may also supervise the application of quality control procedures and take responsibility for database locks.

A clinical data analyst generally designs the CRF and prepares the CRF completion instructions, also known as CRF Completion Guidelines (CCGs). A database programmer-designer is in charge of creating the study database, performing the CRF annotation, programming edit checks for data validation, and subsequently validating those edit checks with dummy data.[2]

Although most clinical trial data are now entered directly into the database via EDC by site personnel, data entry personnel, if needed, will be assigned the task of entering the data into the database following receipt of paper CRF pages. Although the roles and responsibilities may vary slightly for each retina study, the aforementioned team members are considered the minimum requirements for a CDM team.

TIPS FOR EFFICIENT DATA MANAGEMENT

Consistency is key, especially across multiple sites with multiple staff members. It is imperative that every study staff member follows to a T the examinations or procedures as they are listed in the protocol from 1 visit to the next. If, for example, a query is issued because a cataract was missed at baseline and then is documented at the second visit, this only adds more work for all study staff and creates the potential for the false reporting of an adverse event. After a query is issued, the study staff will have to go back and address the query, which takes significantly more time than simply being thorough and consistent in what is being graded from visit to visit. It seems like common sense, but the fewer queries, the better.

Another thing to consider is the true meaning of "relevant" medical history. As far as data management in retina trials is concerned, medical history occupies a significant amount of time and is often a low lying endpoint, insofar as it does not impact the key primary and secondary endpoints. A lot of time is spent cleaning medical histories that have little relevance to the study instead of recent or more pertinent medical histories given the disease state. Surely the importance of medical histories cannot be understated, but a tonsillectomy that occurred 15 years ago may not be as relevant as LASIK surgery the year prior.

Lastly, adverse events occurring in the eye are coded differently from those occurring anywhere else in the body. Ensure that adverse events or medical history terms used in recording are complete and specific. For example, if hyperemia is observed, be sure to consider whether it should be coded as a vascular disorder or ocular hyperemia. When recording a finding, be sure that it is anatomically specific, as this will result in fewer queries for clarification.

CONCLUSION

Careful clinical data management is essential to the integrity of a clinical trial. Involving the CDM team early on ensures that a concrete data management plan is set forth from the start. ■

Aron Shapiro is vice president of retina at Ora in Andover, Massachusetts.

Dale Usner, PhD, is president of Statistics and Data Corporation in Tempe, AZ.

1. Dixon JR Jr. The International Conference on Harmonization Good Clinical Practice guideline. *Qual Assur*. 1998;6(2):65-74.
2. Krishnankutty B, Bellary S, Kumar NB, Moodahadu LS. Data management in clinical research: An overview. *Indian J Pharmacol*. 2012;44(2):168-72.
3. Babre D. Medical coding in clinical trials. *Perspect Clin Res*. 2010;1(1):29-32.