Ensuring Data Quality in the Pharmaceutical Industry

by Christine M. Collins, MA, RRA

Education and training play a considerable part in efforts to enhance data quality in healthcare. How do other fields work toward data consistency? The author describes how professionals in the pharmaceutical industry manage their data.

Today healthcare data serve a variety of significant functions, such as patient care evaluation; making judgements relating to policy issues, delivery of care, research, funding, and growth; supporting information for insurance and benefit claims; aiding in defending patients, healthcare providers, and their employers in legal affairs; disease classification and progress in issues of public health; educating healthcare professionals; and expanding medical knowledge.

Training and education of those involved in data collection are key components in ensuring the quality, integrity, and reliability of healthcare data for its various uses. Just as this is true for the healthcare industry, the same can be said for the pharmaceutical industry.

This article will concentrate on study design and planning, with a major emphasis on how training and education will help achieve data consistency. Finally, the article will provide the reader with a general understanding of data management within the pharmaceutical industry.

Laying the Foundation
Research data provides information that is statistically analyzed to answer the questions being posed by the study objectives as listed in the protocol. The final compilation of such data and submission to the necessary regulatory authority (i.e., the Food and Drug Administration) significantly impacts a company's chance of gaining marketing approval.

A clinical study begins after development, review, finalization, and approval of a protocol by the sponsoring company. A protocol is a document that describes a clinical trial in detail and provides information and rules for the conduct of the trial to all those involved. In parallel to protocol development, it is recommended that the statistician and clinical project manager prepare a preliminary analysis plan outlining the information that will be included in the final report of the trial. The other members of the clinical research team (such as the medical director, clinical research associates, representatives from regulatory affairs and data management, programmers, and medical report writers) review this outline and provide input. Once this is accomplished, it is easier to identify the data items that need to be collected for eventual analysis.

Case report forms (CRFs) are tools designed to collect the necessary safety/efficacy data consistent with the protocol. Measuring and recording clinical trial data are perhaps the most critical steps in the overall data management process. Therefore, it is important for CRFs to be designed with clarity and ease of use in mind. CRF design has a direct impact on the quality of the data collected for a trial, so it is worthwhile to take time in the design and development of the forms and to develop a layout that is user friendly (see Exhibits 1-2). The clinical research team drafts and reviews the forms. The number of personnel involved depends on the size and organization of a company. Industry experience has demonstrated that CRFs subjected to an interdepartmental review yield a superior final product (e.g., fewer data corrections are needed).

exhibit 1—Sample Case Report Form

Study Sponsor: ABCDE Laboratories, Inc. Subject initials: |____|____|____|
Sponsor's Study No.: 12345-XYZ Subject Number: |____|____|____|
Principal Investigator: John Q. Doctor, M.D. Study Number: 100001X

PRE-TREATMENT PHYSICAL EXAMINATION

Date of Evaluation (mm/dd/yy) |____|____|____| Weight (kg) |__|__|.|_| Height (cm) |__|__|.|_| Frame Size: Small ❒ Medium ❒ Large ❒

VITAL SIGNS

Vitals performed after 2 minutes supine

Pulse (/min) |__|__|__| Resp. Rate (/min) |__|__| Supine Blood Pressure (mmHg) |__|__|__| / |__|__|__| Temperature (°C) |__|__|.|_|

Vitals performed after 2 minutes standing

Pulse (/min) |__|__|__| Standing Blood Pressure (mmHg) |__|__|__| / |__|__|__|

PHYSICAL EXAMINATION

System/Body Part	*Not Done*	*Normal*	*Abnormal*	*If "Abnormal," specify abnormality*
General Appearance	❒	❒	❒	
Head	❒	❒	❒	
Eyes	❒	❒	❒	
Ears, Nose, Throat	❒	❒	❒	
Neck	❒	❒	❒	
Lymph Nodes	❒	❒	❒	
Skin	❒	❒	❒	
Lungs	❒	❒	❒	
Heart	❒	❒	❒	
Abdomen	❒	❒	❒	
Skeletal	❒	❒	❒	
Neurological	❒	❒	❒	
Genitourinary	❒	❒	❒	

Comments _____

Typically, the draft CRFs are distributed to all reviewers one to two weeks prior to a review meeting. At the meeting, all participants review the CRFs page by page for logic, consistency to the protocol, grammar, spelling, and data flow, noting corrections, changes, or modifications. It is best to set a time frame for this meeting and stay on schedule. After review, revisions to the case report forms are distributed, and a second meeting is scheduled for final review.

All departments can become more involved and familiar with the details of the protocol when they meet to discuss CRFs. The resulting data will have an impact on several different departments, and personnel can enhance data quality prior to the start of the study.

A common pitfall in CRF design is the collection of data outside the scope of the protocol (i.e., too much data). This can be detrimental because as the volume of data required increases, the quality of data recorded decreases. Therefore, it is important to limit data collection to items that are truly necessary to answer the trial objectives. In assessing data requirements, the review team should distinguish between data needed for clinical care of the patient and data necessary to answer the research objectives. Data collection for the trial should be limited to research data. In many trials, only a small portion of the clinical information will end up as part of the trial database.

A Meeting of Minds

Prior to the start of a clinical trial, pharmaceutical companies typically host an investigators' meeting. The purpose of this meeting is to specifically review protocol objectives, tests, data collection, and patient safety. While they are not required by regulation, the pharmaceutical industry has found investigators' meetings to be a useful review mechanism for investigators as well as their staff.

exhibit 2—Sample Case Report Form

Study Sponsor: ABCDE Laboratories Inc. Subject initials: |____|____|____|
Sponsor's Study No.: 12345-XYZ Subject Number: |____|____|____|
Principal Investigator: John Q. Doctor, M.D. Study Number: 100001X

PRE-TREATMENT CLINICAL LABORATORY RESULTS

Please record and note significance of **abnormal** values **only**:

Clinically Significant (CS); Not Clinically Significant (NCS); Lab Repeated (RPT)

All Lab Values within Normal Limits: ☐

Attach a copy of the clinical laboratory report to this form.

Date Sample Collected (mm/dd/yy): |____|____|____|

Hematology	Value	CS, NCS, or RPT
WBC		
RBC		
Hemoglobin		
Hematocrit		
Platelet Count		

Microscopic Urinalysis	Value	CS, NCS, or RPT
RBC/HPF		
WBC/HPF		
Casts/LPF		

WBC Differential	Value	CS, NCS, or RPT
Neutrophils		
Lymphocytes		
Eosinophils		
Basophils		
Monocytes		
Bands		

Clinical Chemistry	Value	CS, NCS, or RPT
Albumin		
Alk. Phos.		
ALT (SGPT)		
AST (SGOT)		
Bilirubin, Total		
BUN		
Calcium		
Cholesterol		
Creatinine		
Glucose		
Inorganic Phos.		
LDH		
Total Protein		
Triglycerides		
Uric Acid		
Chloride		
Potassium		
Sodium		
GGT		

Urinalysis	Value	CS, NCS, or RPT
Specific Grav.		
pH		
Glucose		
Ketones		
Albumin		
Acetone		

Other Tests	Value
Barbiturates	
Morphine	
Amphetamines	
Benzodiazepine	
Alcohol	
Cannabinoids	
HBsAg	
RPR/VDRL	

Comments: _____

Following internal review of CRFs by the sponsor's clinical research team (and preferably before final printing), copies of the case report forms are usually taken to the investigators' meeting. Here, a significant part of the agenda is allocated for page-by-page CRF review, and appropriate written completion of required data elements data is discussed. These standardizations may include examples of date format (mm/dd/yy or dd/mm/yy) and which format to use when no data is available (e.g., dash, nd, na).

This meeting is typically provided for the clinical research coordinators who will be completing the CRFs by extracting information from the patient chart or source documents. Study investigators are strongly encouraged to attend this session, since they are ultimately responsible for the conduct of the trial at their site.

At the investigators' meeting, the review or case report form training session is an opportunity for site personnel to hear a step-by-step plan for data transcription and standardization that will improve data quality. Frequently, clinical research coordinators will identify inconsistencies within the system of CRFs or perceptual differences (e.g., the data field question is understood one way by the sponsor and interpreted another way by the clinical research coordinator).

Clinical research associates (CRAs) attend the investigators' meeting, and they note questions or feedback from clinical research coordinators. This review process can be reinforced by creative methods such as games, quizzes, and so on.

Another recommended approach is to pilot or test the forms at one or more of the sites that will participate in the trial. Providing as much prospective training and communication regarding CRF completion as possible enables the sponsor to gain valuable feedback on potential problems prior to activation of the trial -- a major advantage.

An additional review of CRF completion instructions may be provided at initiation of the site, especially if multiple clinical research coordinators will be completing the forms and were not in attendance at the investigators' meeting.

After study initiation, upon enrollment of the first few subjects as well as throughout the course of the trial, it is highly recommended that the CRA visit the site to further ensure data accuracy. At that time, the CRA will review case report form completion with the clinical research coordinator. Once again, this reinforces definitions of the various data elements to be captured in accordance with the study protocol. Between each monitoring visit, the CRA is responsible for using various means of communication to strengthen the site personnel compliance with data retrieval and documentation.

Managing the Data

The quantity of data collected during a study is much too great for manual analysis. Therefore, almost all clinical data are analyzed by computer. Data management is the process of translating data from CRFs into a format that can be used by a computer and making sure that the data is completely and accurately entered into the computer.

The data management process begins when CRFs are checked against source documents and reviewed for logical errors at the site ("data cleanup"). This step is usually the responsibility of the clinical research associate. Data entered on a CRF are usually coded before they are entered into the computer. Coding often takes place after another visual check for logical errors and missing data. Some companies, however, prefer to enter the data first and follow up with a computer check for errors and missing data.

The most common method of data entry is the double-entry method. Data are entered from the coded CRFs twice, by two different data entry operators. The computer compares the input, either while a second operator enters data or after the second entry is complete. In either scenario, the two data sets are compared with the CRF, and one data set is corrected to agree with the CRF. Some companies use a single-entry method with a visual comparison of the computer printout and the CRF.

Despite the most careful monitoring procedures and visual checks of the CRFs, it is still possible that there will be logical errors and missing data. Clinical programmers write computer programs to check the database for such errors.

Once the computer has identified the errors, clinical data associates or clinical research associates check the CRF for the source of the errors. If the errors cannot be resolved from information on the CRFs, the investigative site must be contacted. Most often information on a source document can resolve the error. The clinical data associate prepares a "query" which is sent to the site (see Exhibit 3). The majority of companies consider the database to be "clean" when all outstanding errors and discrepancies are resolved. The statisticians then receive the database and conduct the statistical analysis.

There is no correct way to conduct a clinical trial. There are many different ways to organize a trial, and choices must be made based on the environment and available resources. It is important to realize that while everyone involved in clinical trials may believe their way of doing things is best, in reality a data management system is successful if, using available resources, it results in the collection of complete, timely, and accurate data that answer the scientific questions.

exhibit 3—Sample Computer Generated Data Query

REQUEST FOR INFORMATION

960020

Patient ID: 190003 **Site Name: Allegheny University Hospital**

Seq # / Date	Field Name	Reported Value	Condition(s) Checked for	Message	RFI #	Corrections
Page Number: 27.0 Table: Assessment						
10/6/97	Questions 8 and 10	1-No	Questions 8 and 10 on ASSESSMENT p. 27 are listed '1-No', but pt has discontinued DRUG XYZ and RAND STUDY DRUGS	Although pt did not discontinue DRUG XYZ and Rand Study Drugs prior to last assessment ASSESSMENT p. 27 must reflect discontinuation of the study drugs; please correct Questions 8 and 10 to '2-Yes'	37	
Page Number: 28.0 Table: Hematology						
1.0 7/10/1997	BANDS		Not given	Please provide value or indicate 'Not Done'	38	
Page Number: 39.1 Table: DRUG XYZ Dosing						
	PATIENT NUMBER	190014	Questionable PATIENT NUMBER (S/B 190003?)	Please correct Patient Number to 190003	39	
Page Number: 39.1 Table: Dose						
1.0 5/23/1997	END DATE	10/20/1997	END DATE and CONT cannot both be given	Please correct	40	
Page Number: 90.0 Table: Rand Drug Account						
			END of STUDY CRF received but Randomized Drug Accountability and Compliance not received	Please submit copy of RAND DRUG ACCOUNTABILITY and COMPLIANCE	41	

FOR DATA MANAGEMENT USE ONLY
Date Database Updated ___ / ___ / ___ Initials ____
Date Database Vented ___ / ___ / ___ Initials ____

Site Signature_____ Date _____
CRA _____

NC = No Correction Required
V= Data Item Verified

Created By_____

Opportunities for HIM Professionals

HIM experience lends itself to various career opportunities in pharmaceutical research. One area of opportunity is data management. First and foremost, in clinical trials it is beneficial that the data managers understand the clinical issues surrounding the trial and are able to interpret medical record data. Other prerequisites of equal importance are organizational skills and attention to detail. Individuals must be knowledgeable about the trial data to be able to intelligently perform the job. As trials become increasingly computerized, it is also important that these individuals be familiar with computer systems to the extent required for the job. In addition, given the nature of data management and the fact that there are frequent deadlines and urgent and sometimes unexpected requests by the sponsor, it is also advised that these individuals be able to work under pressure.

References

"Glossary of Clinical Research Terminology." *Applied Clinical Trials* 12 (1997): 18-34.

McFadden, Eleanor. *Management of Data in Clinical Trials.* New York: John Wiley & Sons, Inc., 1998.

"Quality Healthcare Data and Information," *Journal of AHIMA* 68, no. 7 (1997): insert.

Rondel, R.K., S.A. Varley, and C.F. Webb. *Clinical Data Management.* Chichester: John Wiley & Sons Ltd., 1993.

Clinical Research Definitions

Case report form (CRF) -- A printed, optical, or electronic document designated to record all of the protocol-required information to be reported to the sponsor on each trial subject

Clean database -- One from which errors have been eliminated and in which measurements and other values are provided in the same units

Clinical research associate (CRA) -- A person employed by a sponsor (or by a contract research organization acting on a sponsor's behalf) who monitors the progress of investigator sites participating in a clinical trial. This person is sometimes referred to as a monitor

Clinical research coordinator -- A person who handles most of the administrative responsibilities of a clinical trial, acts as liaison between investigative site and sponsor, and reviews all data and records before a monitor's visit (also called a trial coordinator, study coordinator, research coordinator, clinical coordinator, or research nurse)

Clinical trial/study -- Any investigation in human subjects intended to discover or verify the clinical, pharmacological, and/or other pharmacodynamic effects of an investigational product(s), and/or to identify any adverse reactions to investigational product(s), and/or to study absorption, distribution, metabolism, and excretion of an investigational product(s), with the object of ascertaining its safety and/or efficacy

Good clinical practice (GCP) -- A standard for the design, conduct, performance, monitoring, auditing, recording, analyses, and reporting of clinical trials that provide assurance that the data and reported results are credible and accurate, and that the rights, integrity, and confidentiality of trial subjects are protected

Investigator -- A person responsible for the conduct of the clinical trial at a trial site. If a trial is conducted by a team of individuals at a trial site, the investigator is the responsible leader of the team and may be called the principal investigator

Monitoring -- The act of overseeing the progress of a clinical trial and of ensuring that it is conducted in accordance with the protocol, standard operating procedures, and good clinical practices and the applicable regulatory requirement(s)

Protocol -- A document that describes the objective(s), design, methodology, statistical consideration, and organization of a trial. The protocol usually gives the background and rationale for the trial, but these could be provided in other protocol-referenced documents

Site -- The location where trial-related activities are actually conducted. This may be in a hospital, private practice setting, or outpatient clinic dependent on the nature of the population

Source documents -- Original documents, data, and records (e.g., hospital records, clinical and office charts, laboratory notes, memoranda, subject diaries or evaluation checklists, pharmacy dispensing records, recorded data from automated instruments, copies or transcriptions certified after verification as being accurate copies, microfiche, 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)

Sponsor -- An individual, company, government agency, institution, or organization that takes responsibility for the initiation, management, and financing of a clinical trial but does not actually conduct the investigation Standard operating procedures (SOPs) -- Detailed, written instructions to achieve uniformity of the performance of a specific function

Subject -- An individual who participates in a clinical trial, either as recipient of the investigational product(s) or as a control. A subject may be a healthy human or a patient with a disease

Off the Beaten Path

Christine Collins *is manager of medical documentation at Fujisawa Healthcare, Inc., Deerfield, IL.*

Article citation:
Collins, Christine M. "Ensuring Data Quality in the Pharmaceutical Industry." *Journal of AHIMA* 69, no.6 (1998): 34-39.