

# DATA MANAGEMENT OF GLOBAL CLINICAL TRIALS

Jeanne Ashton, Executive Director, Data Management

## Introduction

Clinical data management and data quality are key strategic assets of large pharmaceutical companies and clinical research organizations that conduct human clinical trials. Data must be of the highest possible quality to meet both the requirements of accuracy in human trials and historically strict government regulatory conditions.

The complexity of data management continues to grow as both technological and business models for conducting clinical trials evolve to meet the demands of large, multicenter trials. In recent years, companies have begun to shift data acquisition from the traditional (and once ubiquitous) paper to electronic means. Electronic technologies now enable pharmaceutical companies and CROs to expand clinical trial data management and integrate far-flung data centers around the world. With this shift, the role of data management is more important than ever, even as the fundamental nature of clinical data capture and database management has begun to change.

## The Role of Data Management in Clinical Trials

It's no secret that pharmaceutical companies commit a large percentage of their R&D budgets to the clinical development process, and clinical data management is one of the cornerstones for successful drug development and regulatory approval. According to the Tufts University Center for the Study of Drug Development, each new drug requires, on average, 37 separate Phase I–III clinical studies before FDA approval.<sup>1</sup>

Even after FDA approval, Phase IV trials may continue indefinitely while companies further test efficacy and identify new indications for clinical therapies. In fact, the role of data management hardly diminishes after a new drug is approved. Pharmacovigilance data managers continue to amass clinical results

on ongoing trials, drug safety, and any serious adverse effects in study populations around the world. Indeed, in Phase IV trials with large patient populations, data management and pharmacovigilance become increasingly critical as more patients elect to begin taking for the first time a newly approved drug. Only after large numbers of patients take the drug over an extended period can subtle benefits, indications, and adverse events (AEs) become apparent.

**Vioxx® and Celebrex®** — Popular arthritis and pain medicines that were discovered to increase risk of heart attack and stroke.

**Baycol®** — Cholesterol-lowering statin that can cause severe muscle breakdown known as rhabdomyolysis, a rare but well-recognized life-threatening side effect of many cholesterol-lowering drugs

**Fen-Phen** — Weight loss therapy that has potentially life-threatening cardiovascular side effects.

**Tysabri®** — Biogen and Elan's therapy for multiple sclerosis was pulled from the clinic just 3 months after FDA approval because of 2 patient deaths.

Even longstanding therapies such as hormone replacement therapy in postmenopausal women continue to be studied for safety and efficacy in large, ongoing international clinical trials.

Techniques in data management also must keep pace with new technologies in molecular medicine, such as the emerging field of pharmacogenomics, in which responses and adverse effects to therapies are correlated with genomic markers to predict efficacy. An emergent example of genomics co-mingling with therapies is Gleevec®, an anti-cancer drug that has quickly become a front-line chemotherapy agent for patients with chronic myelogenous leukemia – except for patients with specific mutations in the protein targeted by the drug. For those patients, genetic

testing can confirm their resistance to Gleevec, and newer second-generation drugs designed specifically to overcome resistance can be prescribed.

Clinical trials today can be expensive, time-consuming, and ongoing. Clinical data management is one of the fundamental, and often underappreciated, processes that control both data accuracy and timelines of every trial. Clinical data management links clinical research coordinators—who monitor all patient sites and collect all data—with biostatisticians—who analyze, interpret, and report data in clinically meaningful findings. In this sense, data management departments are much like commodity producers. First they harvest and refine raw clinical data. Then they package it in the form of databases for analysis by biostatisticians. All the while, they ensure its quality and accuracy for regulatory agencies, such as FDA.

Global data management groups comprise experienced, dedicated professionals who help oversee both traditional paper trials and e-clinical trials. Data management teams have the requisite experience to manage programs in any phase and of any complexity, including large-scale Phase IV trials and expanded-access programs generating data from thousands of patients. Because each clinical trial is unique, data management groups typically first consult on strategies, systems, databases, processes, procedures, and metrics, including:

- Database design, implementation, and validation
- Data capture, entry, validation, cleaning, and review
- Medical coding using standard dictionaries and thesauruses, including COSART, WHO, MedDRA, and proprietary systems
- Safety procedures, including SAE reconciliation and management of protocol violations, deviations, and endpoints
- Quality assurance to client-specified levels
- Data consolidation, migration, and conversion

## ■ **Data Trails: Paper or PC?**

Case Report Forms (CRFs) made of paper have been the traditional means of acquiring data from clinical trials. However, ICH guidelines on Good Clinical Practice (GCP) refer to both paper and electronic records.<sup>2</sup> Clinical data on CRFs flows to data management professionals from many sources:

- Physicians' offices
- Centralized testing laboratories
- Nurses
- Research coordinators

Data can even come directly from patients in the form of surveys, diaries, self-reports, and, more recently, electronic patient-reported outcomes, or ePROs.

Local EDC systems, file transfers, and central web-based systems all have facilitated the greater use of global clinical trials and data management. Well-managed EDC systems can be common denominators of the global clinical trials information management process. As EDC evolves, data managers can provide significant integration of databases with ongoing EDC information and input processes to efficiently make the transition from paper CRFs to EDC. During the transition, continued implementation of surveillance strategies will maintain data integrity and accuracy.

Developing databases to collect data, cross-check it for accuracy, and analyze it efficiently has never been simple. Data management encompasses many roles and duties, from programmers who customize databases specifically for individual clinical trials to data coordinators who play a lead on-site role in data management throughout the study.

Data management also works proactively with biostatisticians to determine guidelines and specifications about how data will be handled. The two also work

with off-site vendors, such as central testing laboratories, that supply data as electronic files that then are uploaded into a central database as part of a clinical study.

Other responsibilities include organizing non-electronic and non-numerical data from CRFs, such as side effects or AEs caused by clinical treatment. Data management groups categorize these through “dictionary coding,” resulting in a table of AEs and their frequency of occurrence. This information is ultimately included in a product package insert. Although numeric data collection and database programming are not barriers to multinational trials, dictionaries used in the processing, analyzing, and reporting of data collected in clinical trials must be tailored with language and cultural barriers in mind. With global data management of reported AEs, dictionaries may range in size and complexity – from simple code lists with a limited number of categories to complex dictionary systems such as MedDRA (Medical Dictionary for Regulatory Activities) and WHODRUG (World Health Organization Drug Dictionary) with thousands of entries and related tables.

One of the challenges of establishing global clinical trials and data management is to establish and maintain coding dictionaries. Translating and/or re-writing complex coding dictionary detailing symptoms and AEs presents multiple challenges to global data management.

Despite the challenges posed by rapidly changing technologies in data acquisition and storage, the relative ease of data manipulation also enables opportunities for data linkage and mining strategies. However, in light of recent security breaches at consumer database companies, the global policies and procedures of pharmaceutical companies regarding data privacy must be considered.

Data privacy means the norms and standards applicable to the protection of trial subject’s personal data. These norms and standards were codified by the

US Congress in 1996 as the Health Insurance Portability and Accountability Act (HIPAA). The European equivalent is the European Union Directive on Data Privacy, commonly referred to as the EU Directive. Personal data are defined as any information relating to a research subject—patient names, initials, addresses, genetic code, etc.—that can facilitate the identification, either directly or inferentially, of that subject.<sup>3</sup>

## ■ **Pharmaceutical Companies Expand Their Global Data Management Capacity**

As in other industries, clinical data management has been affected by both Internet telecommunications and increasing globalization. Data management is critical for any clinical trial, large or small. However, pharmaceutical companies and CROs have responded to the increase of large, multi-center international trials by building global data management strategies and protocols into their operations. Changes in the landscape of clinical trials have made global data management at once both more practical and feasible as well as increasingly complex. These changes include:

- Globalization of clinical trials
- Revolutionary changes in IT such as direct EDC
- Tighter regulatory oversight requiring thorough documentation and clear audit trails of data changes

Larger pharmaceutical companies, such as Pfizer and Eli Lilly in the USA and GlaxoSmithKline and Novartis in Europe, are increasing their interest and commitment to outsourcing clinical research in the form of global clinical trials. Development of consistent, workable strategies and a standardized set of common protocols and practices for data management of global clinical trials has never been more complex. The task requires vision and discipline and is a

high priority because clinical data management remains the linchpin of the regulatory approval process.

## ■ **Global Integration of Data Management**

For many US companies, the first strategic move toward globalization of clinical data management was to integrate more efficiently with clinical trials professional staff strategically located throughout Europe, giving companies an immediate presence there. Data centers are staffed with data management professionals with experience in protocol design, IT, and biostatistics. These centers also typically have expertise in regulatory requirements of the European Agency for the Evaluation of Medical Products as well as European rules and regulations.

Even more recently, companies have expanded their data management presence in Asia and Australia by opening offices in India, an outsourcing strategy that is hardly unique to the pharmaceutical and biotechnology industries. American companies have been relocating manufacturing capacity to that region for decades. However, in recent years the relocation of white-collar professional jobs has been surging. Pharmaceutical companies and clinical research organizations have moved entire data management departments offshore. Large financial corporations, such as Citibank, Visa, MasterCard, and Bank of America, have set up operations in India, as have technology companies such as Verizon, Hewlett-Packard, and Electronic Data System. Nor is the relocation of jobs limited to back-office functions such as call centers and billing departments. Not surprisingly, the convergence of Internet technologies, an increased reliance on global multi-center clinical trials, and the same economic realities that have led companies of every type to embrace globalization and outsourcing, have led pharmaceutical companies to implement a truly global clinical data management strategy.

“We could ship CRFs from anywhere in the world to our data management offices in the UK or the US,” says Jeanne Ashton, Executive Director, Data Management, at PharmaNet. “But the cost savings in coordinating data management at our Bangalore office are so great that we, like many CROs, are expanding data management globally. Many pharmaceutical companies have been entering the Indian market for years, chiefly because of the lower labor costs. Also, India has great infrastructure for Internet and IT wide area network capacity, which is ideal for data management between offices around the world.”

As companies expand into global clinical trials, harmonizing data collected from several clinical sites is only reinforcing the trend away from paper to EDC. With EDC, a data management center provides each clinical research site with a system to log into a secure server for direct EDC. This ‘real-time’ EDC has been evolving since the 1980s, when the standard was remote data entry by capturing data on a remote PC at the clinical research site and then performing a remote data transfer. Now, with the Internet, secure data encryption, and web-based applications, it’s routine to enter data from any computer for EDC.

Among its many advantages, EDC eliminates data transcription errors from CRFs, as well as illegible data. However, quality control with EDC requires data review via source documentation. This is particularly true since passage of the FDA regulation 21 CFR Part 11, which places tight controls over data changes and the manner in which they are carried out.<sup>4</sup>

For example, in a clinical study that calls for a patient age range of 40 to 60 years, if the database shows that one patient’s age was entered as 73 years, then data managers are required to single out that data point and ask the site “was that patient actually 73, or was that an input error?”

Data programmers are increasingly creating programs that they run against the data to verify it, and send out questions—known as ‘queries’ or ‘data

clarifications'—to the sites so that data being analyzed by biostatisticians and prepared for submission is as clean and error-free as possible. FDA's Part 11 requires an 'audit trail' to show the source of errors, which may include:

- Input technicians
- Site coordinators
- Healthcare workers
- Data management professionals

The failure to document such changes can result in fines or suspension of clinical trials. So, although EDC makes the logistics of global data management more streamlined, it carries with it its own regulatory burden of proof. To manage the documentation of data queries and database changes, important details about each record that is created are saved in the database, including:

- Author
- Creation date
- Ownership
- Searchable keywords that can be used to classify the document
- Details as to the type of data found in the document
- Relationships between data components

Before providing any data to regulatory agencies, such as the FDA, for review, data management groups are confident that there are accurate checks and balances for evaluating databases, and with any database changes there is a fully transparent audit trail. With global data management, this may now require viewing source documentation from halfway around the world, regardless of whether it is an original EDC trail or a paper CRF. This task can be more quickly accomplished with faxes and e-mail attachments than any worldwide courier could deliver.

## Summary

Both technology and market forces are reinforcing reliance on integrated strategies for global data management. Many traditional strategies and Best Practices of the industry are being modified and updated to incorporate new Internet and computer-based data acquisition, and such systems are evolving to respond to governmental regulatory requirements both in the US and Europe. The latest industry response to global trials has been to expand data management operations into Asia, which has both the infrastructure and white-collar professionals necessary to oversee data collection, integrate with other centers around the world, and build a unified system of global data management.

## References

1. Dimasi JA, Paquette C. The economics of follow-on drug research and development: trends in entry rates and the timing of development. *Pharmacoeconomics*. 2004;22:1-14.
2. Management, S.f.C.D. 2003. Good Clinical Data Management Practices, Version 3.
3. Hansen E. HIPAA (Health Insurance Portability and Accountability Act) rules: federal and state enforcement. *Med Interface*. 1997;10:96-98, 101-102.
4. Keatley KL. A review of US EPA and FDA requirements for electronic records, electronic signatures, and electronic submissions. *Qual Assur*. 1999;7:77-89.