ClinAccess™: An Integrated Client/Server Approach to Clinical Data Management and Regulatory Approval

Martin J. Rosenberg, Ph.D., MAJARO INFOSYSTEMS, INC.

ABSTRACT
Motivated by the desire to better manage an ever increasing volume of information and to reduce the length of time required to introduce new drugs around the world, the pharmaceutical industry and regulatory agencies such as the U.S. Food and Drug Administration have sought to facilitate the drug development and approval processes through innovative uses of computer technology. One such effort that has received great attention is the computer assisted NDA or CANDA and its cousin for biologicals the computer assisted PLA or CAPLA.

The first CANDA's were created with much effort, and at great expense, after the paper NDA's had been filed. From these early experiments, it became evident that if CANDA's were to ever become routinely used throughout the industry by both large and small companies, their development would need to be simplified and to begin earlier in the drug development process. Medical reviewers at regulatory agencies have frequently commented that for a system to be useful as a CANDA, it should also be beneficial for in-house use. While recent advances in computer technology now permit the construction of sophisticated systems which remain relatively simple to use, in order to maintain data integrity, it is essential that data tables do not require transformation or restructuring by the end user. Hence ideally, the development of a CANDA should begin at the time of study definition and data entry.

With these goals in mind, MAJARO has explored ways of developing systems which seamlessly integrate data entry and in-house medical review, so as to facilitate CANDA development. This paper discusses CLINACCESS™ an integrated clinical trials system developed with SAS® software, that combines the data entry and data management capabilities of traditional clinical information systems, with clinical data review features that permit monitors, CRA's, managers, and other members of the clinical staff to monitor the progress and quality of ongoing clinical trials. It achieves its ease of use through a graphical user interface that incorporates pull-down menus, scroll bars, dialog boxes, radio buttons, and a mouse. Initially developed for use on PC compatibles, the code has been written to permit easy porting to other computer platforms and exploitation of client-server technology.

THE DRUG APPROVAL PROCESS
Unlike most other industries, companies in the pharmaceutical and biotechnology industry must first obtain government approval before they can bring a new product to the market. The process of obtaining government approval can often be long and complicated. In the United States, the company must first use the proposed drug in animals to show that there are no gross toxic effects. With this knowledge, the sponsoring company obtains an IND (Investigational New Drug) from the Food and Drug Administration (FDA). An IND permits the sponsor to ship the drug across interstate lines for tests in humans.

Once an IND is obtained, the three pre-marketing phases of clinical research begin. Phase I typically starts by using small single doses in healthy human volunteers in order to look for toxic effects in man. As evidence of safety is accumulated, dosages are increased to therapeutic levels. In Phase II, the drug is tried in human volunteers with the targeted disease. Sample sizes are small, and information is gathered concerning the appropriate dose and regimen to use to obtain a therapeutic effect. Once sufficient information is gathered and hypotheses about the drug's efficacy in various indications have been formulated, large scale Phase III testing in human volunteers begins.

The research process can take five or more years. Throughout the process additional information is gathered concerning the safety of the drug in both humans and animals; the stability of the drug (how long it can remain on the shelf without degrading); the pharmacokinetics of the drug, i.e., how it is metabolized in humans; and the ability of the company to manufacture the drug in production quantities.
When all the required research is completed, the drug company submits its information to the FDA and makes a formal request for permission to market the drug in specified diseases. This registration process is known as a New Drug Application or NDA in the United States. An NDA is frequently between 200 and 400 volumes long or more than one hundred thousand pages. Although information from many scientific disciplines is presented, frequently the largest single portion of the NDA is the clinical section. The FDA reviews the NDA and can take three actions: approve the drug for all or some of the indications; not approve the drug; or request more information be gathered. The review process frequently takes another two or more years.

Although the details differ, the drug approval process is conceptually the same in other countries. Experimental information is gathered according to the requirements of the particular country until sufficient information is gathered to submit the drug for approval to a governmental regulatory agency.

**CLINICAL INFORMATION SYSTEMS**

As can be imagined from the magnitude of information that must be collected, computer systems have long been a part of the drug development process. We call a computer system for the collection, retrieval, and analysis of clinical trial information, a Clinical Information System or CIS. The three functions comprising the current concept of a Clinical Information System are shown in Figure 1.

<table>
<thead>
<tr>
<th>In-House Data Entry and Storage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistical Analysis</td>
</tr>
<tr>
<td>Table and Graph Generation</td>
</tr>
</tbody>
</table>

**Figure 1: Current CIS**

While such systems have long existed within the pharmaceutical industry, their use has generally been limited to the data processing, biostatistics, and programming staffs. Other users who could benefit from the information stored in the CIS, such as medical monitors and CRA's have not had direct access. Such access would be useful to reduce paperwork, detect and prevent protocol violations, and provide valuable information for planning future studies. We call such a system for use by medical staff, a Clinical Data Review System.

Additionally, there has been much interest in computer systems that permit the investigator to enter data and the sponsor to remotely monitor the trial. The intent of such systems is to collect more timely and accurate information. We call such systems Remote Study Monitoring systems.

**CANDAS**

Spurred on by the desire to better manage an ever increasing volume of information and to reduce the length of time required to introduce new drugs around the world, the pharmaceutical industry and regulatory agencies such as the U.S. Food and Drug Administration have sought to facilitate the drug development and approval processes through innovative uses of computer technology. One such effort that has received great attention is the computer assisted NDA or CANDA and its cousin for biologicals the computer assisted PLA or CAPLA. The features and technology utilized by CANDA's have varied greatly. Nonetheless, three general areas have emerged: a data portion, with the ability to query and analyze an on-line database; an image portion, that presents electronic images of case report forms; and a text portion, that displays and frequently permits the manipulation of the text from the paper NDA submission.

Experimentation with CANDA's began in 1986. The earliest CANDA's were developed after the paper NDA's had been submitted. Because of this, they involved tremendous effort and frequently great expense. As the use of CANDA's increased, it became clear that similar tools would be useful for review by in-house medical and regulatory personnel.
Figure 2: Future CIS

With the introduction of these three new classes of users (investigators, in-house medical staff, and FDA reviewers) future clinical information systems will resemble Figure 2. Not every study will require all the facilities of the future CIS's. In particular, Remote Study Monitoring is likely to be used in selected studies only. However, pharmaceutical firms will increasingly expect to have these capabilities at their disposal.

INTEGRATING CANDA DEVELOPMENT INTO THE CLINICAL RESEARCH PROCESS

CANDAs have evolved into three components: data, image, and text. The data component is comprised of the database, analysis programs, and tools to let medical and biometric reviewers access the database. The image component consists of images of the CRFs and possibly the entire registration package. The text component often means a copy of the registration package stored in word processing files so as to permit cutting and pasting from the NDA into the documents that need to be prepared by the reviewer. With careful planning, it is possible to integrate the data and image components into the clinical data management operations already being performed.

CRFs Logged and Scanned

As CRFs are collected, it is common to stamp them with a unique document ID number (also known as an accession number) and log them into the database as received. Increasingly, companies are also scanning the CRFs into an image database (Figure 3). Image databases have several advantages: the original CRF can be safely stored; many people can access the CRF at the same time, e.g. data entry and monitors; and it provides a single source for the most up to date copy of the CRF.

The log-in process, used in clinical data management to track CRFs, can also serve as the index required by the image database. With modern database technology, the image and data entry screen can be displayed side by side to facilitate data entry (Figure 4). As data entry is performed, the links between the records in the database and the CRF images are created. Hence, the database and CRF image components of the CANDA can be
created by substantially leveraging the effort already being expended in clinical data management.

CLIENT/SERVER COMPUTING AND THE SAS SYSTEM

With the introduction and enhancement of SAS/CONNECT and SAS/SHARE software in SAS 6.08 and higher, the SAS system has become an excellent software platform for client/server computing. The SAS system follows three models for distributed and client/server computing: remote transfer services, remote submission services, and remote library services. Remote transfer services permit all or part of a database table to be up or down loaded between a client and a host server. Remote submission services permit a client to submit a program which executes on the host server using data stored on the host server, with the results sent back automatically to the client. This can be used to harness the ease of use of PC computers to the speed of mainframes. Remote library services permits a client to interactively enter, modify, or query data stored on a host server. In each case, the client can be either the same or a different hardware platform than the server. For example, PC or Macintosh clients can connect to a UNIX or VAX VMS host acting as a server.

Among PC compatible installations, the most commonly used network is Novell NetWare®. To perform client/server computing, one connects a PC, which will act as the SAS/SHARE database server, to the network as an ordinary client as shown in Figure 5. To the user however, it appears that each client workstation is connected to the database server which in turn can access data stored on the file server (Figure 6).

In operating systems (such as UNIX and Windows NT) that can act as both a file server and an applications server, it is possible for the same machine to provide both network file server and SAS/SHARE database server services.
While recent advances in computer technology have made it possible to develop systems capable of performing complex tasks while requiring minimal knowledge on the part of the user, transformation and restructuring of data tables are best left to computer professionals due to the potential for compromising the integrity of the database. Thus, for in-house review systems to become practical, the data must be available in a form that does not require transformation. If such systems are to be used while studies are in progress, the data storage and retrieval structures used for data entry must already be amenable to clinical data review. With these objectives in mind, MAJARO InfoSystems has explored ways of developing systems which seamlessly integrate data entry and in-house medical review into CANDA development. Our first product is the CLINACCESS™ clinical trials system.

CLINACCESS™ is an integrated clinical trials system that combines the data entry and data management capabilities of traditional clinical information systems, with clinical data review features that permit monitors, CRA's, managers, and other members of the clinical staff to monitor the progress and quality of ongoing clinical trials. Originally written in Version 5.18 SAS/AF® software for use on IBM mainframes (Rosenberg 1989a and 1989b), CLINACCESS has been ported to Version 6 SAS software and is available on desktop platforms.

CLINACCESS 2.7 is a full featured data entry system containing all the capabilities required in a pharmaceutical environment: an integrated data dictionary to facilitate new study definition and pooling of data, single and double-key data entry, customization of data entry screens to resemble Case Report Forms, interactive and batch edit checks, support for a thesaurus such as COSTART, and a full audit trail facility which records all changes and the reason for the changes. However, unlike similar systems, CLINACCESS not only captures data, but delivers information to users on the clinical staff such as medical monitors, clinical research associates, and managers. This information permits the clinical staff to track the progress of ongoing clinical trials and use this information to plan future studies. Designed for the 90's, CLINACCESS has a graphical user interface to facilitate training and ease of use.

An Example

These concepts will be illustrated by an example of data entry and clinical data review.

Data Entry Example

CLINACCESS is designed as an integrated clinical information system, which is distinguished from other such systems by its graphical user interface (GUI), ease of use, and strong clinical data review component.

The data entry component of CLINACCESS 2.7 consists of study definition, data entry with on-line edit checking, data verification, post-entry data validation, and audit trails.

One or more users are identified to CLINACCESS as Database Administrators (DBA). To protect the integrity of the data, only the DBA's have access to a subsystem of CLINACCESS which is used to: create study libraries, define new studies to the system, manage existing studies, verify and validate data, update the Clinical Questions Catalog, and add new users to the system. Similarly, only those users specifically granted data entry and verification privileges by the DBA are permitted to enter or modify data. All other users may view but not modify the database.

To define a new study to the system, the DBA first creates a study library. Then one or more tables are created to hold the data. CLINACCESS was specifically designed to facilitate the pharmaceutical industry's need to create data entry applications rapidly. To meet this need, tables can be defined in two ways. First, the table can be created from the Clinical Questions Catalog by merely selecting the names of variables to be included from a list. All defining information such as variable type, length, label, format, and informat are automatically included. To provide flexibility while enforcing standardization, the label, informat, and format may be customized to the study, while the name, type, and length of the variable are fixed.

The second method recognizes the fact, that pharmaceutical firms frequently run similar trials on a compound. To accommodate this, a table may be created from a previous study which is similar to the new study. The two studies need not be identical as variables may be added or deleted from the definition. To maintain consistency, CLINACCESS automatically checks to make sure that any variable which is added is defined in the Clinical Questions Catalog. The final step in defining a new table is to
specify the primary key. The primary key is the variable or variables which uniquely identify each record in the table. This primary key is stored in a meta-data file and simplifies use of the data review components of the system.

Once the tables are defined, screens can be customized to resemble case report forms, and powerful cross field edit checks and computations can be performed during data entry. For example, to increase accuracy, clinical trial protocols often require blood pressure to be measured three times at each reading and the average used as the response. As shown in Figure 7, the data entry operator can enter the three sets of blood pressures and the mean will be accurately computed and stored in the dataset, available for immediate analysis.

Optionally, data can be entered a second time (double-key entry) to detect and prevent key stroke errors. Changes made to the data subsequently are captured in an audit trail.

Finally, the DBA can use the complete power of the SAS system's DATA and PROC steps to create batch data validation programs which can be customized for each study and run from a menu option.

Clinical Data Review Example

CLINACCESS 2.7's unique strength continues to be the ease with which data can be accessed and manipulated. To illustrate this, let's take an example of how CLINACCESS can assist with reviewing laboratory data. The user will view the laboratory data, create a report displaying the data to take along on a site visit, generate a graph to detect a trend in a specific lab test over the course of the study, and then generate another report to examine the trend in detail.

Throughout this process, data integrity is constantly maintained. As mentioned previously, only those users specifically granted data entry privileges by the DBA may modify the database tables. The data review component of CLINACCESS cannot change data in the underlying database. Instead, all data manipulation, such as selecting subsets of patients, is performed on copies of the data and stored in the user's personal library.

To begin, the user selects the View option from the Main Menu. The user is given a choice of viewing data in Case Report Form or Table formats and decides to view the data in CRF format. The user then selects a study from a list of studies. For clarity, the list includes the study name or number and a brief description of the study (Figure 8). Thanks to CLINACCESS's graphical user interface, the operation is completely point and shoot. The user merely positions the cursor on the study name by using a mouse or the tab key and clicks the mouse button or presses enter to make the selection. The user is presented with a list of data available to be viewed in the study and similarly makes a choice.

Figure 7: Cross-field edits and computations can be performed during data entry

Figure 8: The user selects a study from a list which includes the study number and a brief description.
Finally there is some concern that similar compounds have caused a decline in White Blood Cell counts. To detect a trend, the user selects the Graphics option and chooses a vertical bar chart. After selecting the variables to be graphed a chart is generated which appears to detect a decrease in WBC counts over time (Figure 11). To get more detail, the user selects the Descriptive Statistics report option from the Reports menu, selects the variable WBC for analysis, chooses statistics to compute from a list, and requests that the statistics be computed for each treatment group at each visit. The report indicates that there may indeed be some cause for concern about decreased white blood cell counts. The reviewer can then discuss performing some definitive tests with a biostatistician.

FUTURE DIRECTIONS

In 1990, SAS Institute introduced Version 6, a new generation of the SAS System. This release adds many capabilities, such as indexing and SQL, usually associated with relational databases. Additionally, the introduction of multiple engine architecture permits applications written with SAS software to directly access data stored in such popular databases as Oracle® and IBM's DB2. Further details of these new capabilities are described in Rosenberg 1990.

In 1994, SAS Institute announced three new important features: SAS/GRAPH software with image extensions; remote library services for SAS/CONNECT® and SAS/SHARE® software; and the development of a version of SAS for the Apple® Macintosh™ computer. Work is underway to exploit...
these new capabilities in future CLINACCESS releases so as to provide a total integrated solution from data capture to regulatory reporting.

SUMMARY

The CANDA Initiative begun in 1986 has advanced to the stage where nearly 1 out of 3 new submissions have a CANDA component. In order for CANDAs to continue to become a routine requirement, the creation of CANDA's must be simplified and begun early in the clinical development process. The CLINACCESS clinical trials system is an important step in this direction. Developed entirely with SAS software, CLINACCESS is designed to provide monitors, CRA's, and other non-traditional users with access to the information stored in clinical databases. CLINACCESS provides: single or double-key data entry; viewing and querying of data; graphics; descriptive statistics; and report generation and is available now on PC compatibles running Microsoft Windows or IBM OS/2®. Versions for UNIX workstations, DEC Alpha AXP, Microsoft Windows NT, Microsoft Windows 95, and Apple Macintosh are under consideration or development.

For a number of years now, we've witnessed the evolution of SAS software into a product with greater interactivity and data management capabilities. Version 6 has been a major step in that evolution. Companies that start now to exploit these new capabilities through systems such as CLINACCESS, have the potential of realizing substantial advantages over their competitors in terms of reducing the cost and time needed to bring a new drug to the market.

REFERENCES


ACKNOWLEDGMENTS

CLINACCESS is a trademark of MAJARO INFOSYSTEMS, INC., in the USA and other countries.

All CLINACCESS screens shown are Copyrighted © 1988-1992 by MAJARO INFOSYSTEMS, INC. and are used by permission. All rights reserved.

SAS, SAS/AF, SAS/FSP, SAS/GRAPH, SAS/CONNECT, SAS/SHARE and MultiVendor Architecture are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

IBM and OS/2 are registered trademarks of International Business Machines Corporation.

Oracle is a registered trademark of Oracle Corporation.

Novell and NetWare are registered trademarks of Novell, Inc.

Other brand and product names are the registered trademarks or trademarks of their respective companies.

MAJARO reserves the right to modify CLINACCESS specifications and screen designs without notice.

MAJARO INFOSYSTEMS, INC. provides statistical and information management services to the pharmaceutical, biotechnology, and food products industries, and specializes in extending computer technology to non-traditional users.

For further information regarding this paper, please contact:

Martin J. Rosenberg, Ph.D.
MAJARO INFOSYSTEMS, INC.
2700 Augustine Drive
Suite 230
Santa Clara, CA 95054
tel. (408) 562-1890
fax. (408) 562-1899