Drug Information Journal, Vol. 33, pp. 1061–1065, 1999
Printed in the USA. All rights reserved.
0092-8615/99
Copyright © 1999 Drug Information Association Inc.

CLINICAL TRIAL MANAGEMENT AND REMOTE DATA ENTRY ON THE INTERNET

ROBERTO SCOGNAMILLO

Senior Data Manager and Clint&rnet Project Leader, Data Management and Statistics Unit, Medical Department, Glaxo Wellcome S.p.A.,Verona, Italy

CARLO STROZZI

Software Engineer, IBM Network Services, IBM SEMEA S.p.A., Milano, Italy

BEATRICE VINCENZI

Head of Unit, Data Management and Statistics Unit, Medical Department, Glaxo Wellcome S.p.A., Verona, Italy

GIUSEPPE RECCHIA

Medical Director, Medical Department, Glaxo Wellcome S.p.A., Verona, Italy

Despite the widespread use of information technology and communication advances in the pharmaceutical industry, there are still some steps of the clinical process, such as clinical data collection, that do not take advantage of this technological revolution. In order to verify if a new technology such as the Internet could both speed data capture and improve the overall quality of data, the Italian Glaxo Wellcome company and IBM have cosponsored a pilot project called CLINical Trial & Research Management via InterNET (CLINT&RNET). Within this joint project, IBM has developed the data capture Internet application and Glaxo Wellcome has handled the clinical and organizational aspects, always in compliance with regulatory laws. This paper summarizes the technical aspects of the project.

Key Words: Remote data entry; RDE; Electronic data capture; EDC; Remote data management; RDM; Clinical trials; Internet

INTRODUCTION

THE CLINICAL DEVELOPMENT of a drug is a complex and costly process which must be conducted under numerous bonds and regulations. Monitoring clinical studies, in particular, data management, is quite demanding, both in terms of the time required and the human resources dedicated to this process.

Pharmaceutical companies have been trying to improve the overall quality, timeliness, and efficiency of their clinical development processes. Information technology offers them an opportunity to do so. In fact, the pervasive presence of innovation due to information technology is changing the way we work, frequently faster than we would like. Over the last few years, technology companies have proposed many exciting

Reprint address: Roberto Scognamillo, Senior Data Manager and Clint&rnet Project Leader, Data Management and Statistical Unit, Medical Department, Glaxo Wellcome S.p.A., via A. Fleming 4, 37100 Verona, Italy.

tools to clinical trial management and data management to improve clinical data capture: interactive voice response, faxed and scanned optical image data capture, smart phones, PC-based remote data entry, and the Internet.

Glaxo Wellcome S.p.A. (GW) in Italy and IBM have launched CLINical Trial & Research Management via InterNET (CLINT&RNET), an application to remotely manage clinical trials and capture all patients' data via the Internet. This project must answer the following questions:

- Is it possible to concentrate the case report form (CRF) design, database set-up, and data entry screen programming for the three phases into one faster step?
- Is manual input to the Internet by the investigator the most appropriate method both for data collection and to increase the investigator's responsibility for delivering accurate information? and
- Does this web application allow the sponsor to follow the qualitative and quantitative progression of the clinical trial online?

A randomized, double-blind, Phase III, multicenter trial, FLIC14, was used to test the hypothesis of system development. The study objective was: "To evaluate if adding salmeterol 50 μg bid to low-dose fluticasone propionate (100 μg bid) or using a higher dose of inhaled fluticasone propionate (250mcg bid) is more effective than low-dose inhaled fluticasone propionate alone (100 μg bid) on pulmonary function and inflammatory markers of induced sputum in naive patients with mild to moderate asthma poorly controlled with one-month of therapy with low-dose fluticasone propionate (100 μg bid)." Seven investigators were charged with recruiting 100 patients to be treated for six months. All investigators used the hardware and the Internet connection available at their clinical center.

This study involved managing a large quantity of data. FLIC14 consists of 10 visits, one follow-up visit, and a withdrawal visit, along with three special sections: to collect laboratory data, for concomitant medications, and for adverse events reporting. Thus, there are 15 visits/sections, which translates into 320 fields plus 13 for each concomitant medication and 15 for each adverse event.

IBM developed the application based on the ad-hoc paper system designed by the GW Data Management team. It was tested and validated by both IBM and GW and was loaded on the IBM server based in Milan. All parties (investigators, and GW clinical monitors and clinical data managers) received adequate training and access to the server, according to their level of authority, by specific passwords.

Since the trial is still ongoing, it is too early to discuss results. This paper focuses on the technical solution adopted, which is working very well and appears to be compliant with regulatory requirements.

METHODS

The general architecture (Figure 1) of the application developed by IBM is based upon the classical World Wide Web (1) two-tier model, where web clients access the application, located on a web server, with an ordinary Mozilla-compatible web browser program.

The web server runs on a set of common gateway interface (CGI) (2) programs, written primarily in the Perl (3) programming language and backed by a simple relational database management system (RDBMS) engine. Additional JavaScript™ (4) code embedded in the HTML forms is used to carry out data validation checks on the client side.

We decided to use this time-honored architecture after having verified the clumsiness, inefficiency, poor scalability, and platform-specificity of a pure Java® approach, where it was virtually impossible to ensure consistent behavior across different Java® implementations on client browsers. On the World Wide Web server side, the leading web server, Apache (ssl-Apache) (5), which is widely available, was used.

The key criteria to achieve the final product were:

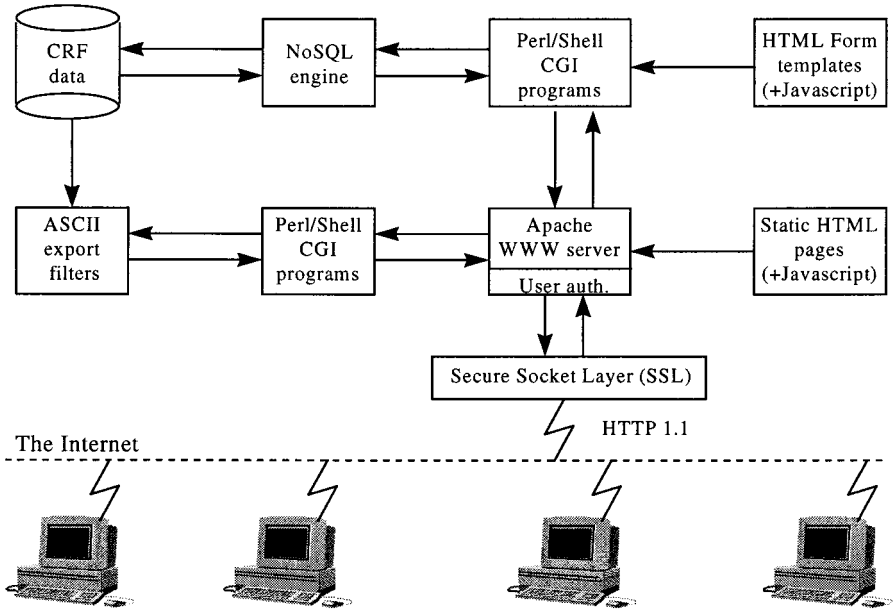

FIGURE 1. The CLINT&RNET flow chart by Carlo Strozzi of IBM Network Services.

- *Compliance with regulatory guidelines*: The application has a security system that prevents unauthorized access to the database, an audit trail log which documents data changes, and a list of the individuals who are authorized to make data changes. All data are adequately backed-up. Due to the nature of the data being sent back and forth between the web server and the investigators, security of the application was not an option. Data flow confidentiality has been ensured by the use of the Secure Socket Layer (SSL) communication protocol; we decided to use the widely available, cryptographically strong open source (6) version of SSL, SSLeay (7). This was chosen because the development team is based in Australia and SSLeay is not subject to export restrictions,
- *Ability to move between the GW Clinical Trial Data Management Database and the Internet*: All CRF data on the web server are stored in ASCII files which can be easily extracted and transferred from the IBM server to the GW Clinical Trial Data Management Database (CTDMD). This strategy of implementing only the data acquisition and validation front-end on the World Wide Web platform eliminates the need to replicate online the complex data processing process already performed by the CTDMD,
- *Openness*: The entire application is based on the UNIX Operating System (8), which offers the required degree of flexibility. Many handy open source programs and tools are available for free downloading from the Internet. Specifically, the current version of the system runs on the Linux operating system, a free UNIX clone that is rapidly gaining in popularity over both traditional UNIX platforms and other operating systems (9),
- *Database back-end*: The database engine is NoSQL (10), a simple relational database management system for UNIX (open source, and free of charge, software developed by Carlo Strozzi), with no arbitrary limits other than memory and processor speed, and
- *Flexibility*: The e-CRF system is completely customizable through external ASCII tables (in NoSQL format), with no need to change the central programs on the web server (CGI programs) to add more

functionalities, variables, checks, or other features.

The main steps in setting up the electronic CRF (e-CRF) with this application will be outlined in general terms. The table which contains CRF-related variables is the most important table. Its role is to define all necessary variables, and to describe them and other associated attributes, such as the list of permitted values. The HTML forms which constitute the actual e-CRF can be designed just as if they were ordinary paper CRF sections, with readily available tools, such as word processors, preferably with the ability to export documents in HTML format.

Variables, alongside the queries, can be added by following the simple syntax required by CGI programs (ie, few ASCII characters) to the resulting HTML pages. This can also be done in a subsequent step, by means of an ordinary text editor (such as Notepad) and, if necessary, by someone other than the form designer. An HTML page may also refer to variables that belong to a different form, to allow the CGI programs on the web server to carry out consistency checks across different pages/visits. JavaScript programmers can then add the relevant code to the web pages, to perform the necessary validation checks on the client side when the forms are filled out by the investigators.

Finally, another NoSQL table defines web user IDs and passwords for the subjects involved, such as investigators, clinical monitors, and clinical data managers. These IDs and passwords are the credentials used by the web server to decide who does what, depending on how the client has authenticated himself upon connecting to the server.

CONCLUSION

The deployment of stock Internet tools has resulted in a final application which is both effective and low-cost. It can be successfully adopted by small research groups where funding may be scarce, and by large organizations where the availability of the source code allows for extensive customization and the ability to easily fix problems. The adoption of open source software makes operating and modifying the system substantially less expensive than using proprietary systems, as the program source code is widely available and program licensing allows use at no charge.

The proposed two-tier model is very functional for the relatively small amount of processing needed to scan an electronic CRF page on the web server. Should the need for heavier processing arise, however, the system can easily be extended into a three-tier architecture, that is, into an application server, by backing the CGI programs with a robust batch processing system. This is the approach taken by some well-known CPU- and memory-intensive web applications, such as those used for DNA sequencing analysis.

These technological developments make new things possible. The Internet represents the natural evolution of the data capture. As with all new technologies, however, the Internet could be subjected to excessive hype but even if its use in clinical data collection proves not to be advantageous, these opportunities need to be fully explored. At present, the Clint&rnet site is hosted on an IBM server at the following address:

http://Clinternet.glaxowellcome.it

REFERENCES

1. Exploring the World Wide Web—University of Georgia, Center for Continuing Education. http://www.gactr.uga.edu/exploring/index.html._.
2. CGI Programming OpenFAQ—Boutell.Com, Inc. Seattle WA98102, USA. Http://www.boutell.com/faq/.
3. Wall L, Schwartz R. “Programming Perl,” O’Reilly & Associates,1991. Http://www.ora.com and http://www.perl.com.
4. A Beginner’s Guide to JavaScript—SimpleNet, a subsidiary of Broadcast.com. http://jsguide.simplenet.com/.
5. Welcome to the Apache Server Project—Apache Group. http://www.apache.org and http://www.apache-ssl.org
6. Open Source: The Future Is Here—Opensource.org. Http://www.opensource.org/.
7. SSLeay and SSLapps FAQ—The University of

Queensland. Http://www.psy.uq.edu.au:8080/~ftp/Crypto/.
8. The Kirch Paper—The UNIX versus NT Organization. Http://www.unix-vs-nt.org.
9. Linux Online! and Linux NOW—by Linux Programmers Worldwide. Http://www.linux.org and http://linuxnow.com.
10. Strozzi C, "NoSQL: A Relational Database Management System." Italian Linux Society. Http://mi.linux.it/nosql.