A Web-Based Clinical Trial Expert System (WebCTES)

Luigi G. Franciosi, Bernard A. MacLeod, Mihai I. Huzmezan, Ciprian Ticea and Adrian Cosma
Division of Control Systems, Departments of Pharmacology & Therapeutics and Electrical & Computer Engineering, Faculties of Medicine and Applied Science, University of British Columbia
2176 Health Sciences Mall, Vancouver, British Columbia, V6T 1Z3, Canada
http://www.pharmacology.ubc.ca/cspt

ABSTRACT

The present problem of poorly conducted clinical research may be reduced by the use of a web-based clinical trial expert system (WebCTES) for the design and implementation of clinical trials. This comprehensive online system guides health care professionals and students through the reasoning process of conducting clinical trials. The system contains a domain-specific knowledge base and a list of typical tasks or operations common to all clinical trials. In particular, WebCTES is made up of five modules, each representing a principle of scientific conduct: *question*, *design*, *statistics*, *ethics*, and *standard operating procedures*. The value of system's expertise, which is a first level knowledge-based expert system, is to effectively guide students or researchers through each module or "Good Clinical Practice" and enable them to clarify their reasoning. These modules contain a number of question prompts that force users to consider the details of planning, initiating, monitoring and completing their trials. Typical tasks in the clinical trial workflow such as drafting protocols or preparing applications for institutional review boards are immediately made possible through the system's information retrieval and printing properties. To assess the system's performance, a plan is outlined to study its usefulness, capability, ease-of-use, intelligibility, efficiency, and reliability in certain user populations. Randomized controlled testing is proposed to establish its utility compared to conventional methods.

Keywords: Clinical Trials, Expert Systems, Internet, Computer-Aided Design and Implementation, Knowledge-Based Systems, Applications: Linux, Oracle, Java.

1. INTRODUCTION

Over the past two decades, many articles and books have been published that call for improvement in the quality of clinical trials conducted [1-12]. The inability of clinical investigators to consider factors that may confound their results, the improper use of design and statistical methods, and the insufficient training, time and monetary support to implement standard operating procedures are some of the major reasons cited for poorly conducted clinical trials [1,2,6,7,8]. Methods used to improve clinical trial process usually involve the education of clinical investigators about the randomized controlled trial (RCT) design through academic publications, formal seminars, certification programs, and the use of statistical software. Evidence-based organizations such as the Cochrane Collaboration have been established to analyze and disseminate published clinical trials for use by not only practicing physicians but also by clinical investigators to assist their decision making during the planning process [9].

The World Wide Web is becoming increasingly a popular medium for improving the conduct of clinical trials [4]. New training courses and methods of electronic data capture are being offered online to clinical investigators [4]. Companies such as Pharsight Corporation of Mountain View, California, USA (http://www.pharsight.com) and Phase Forward Incorporated of Waltham, Massachusetts, USA (http://www.phaseforward.com) are providing online statistical software and informatic tools to 'optimize' the clinical trial development process for clinical investigators and drug companies. However, the worthiness of these methods and tools for improving clinical trial quality, and in turn, medical practice is not entirely known. Evidence for which methods work is often rare and usually not based on the same rigorous randomized controlled trial testing that the very methods promote or rely upon [25].

The purpose of this paper is to present a web-based expert system that has been developed by the University of British Columbia as a tool for academic clinical investigators to design and implement their clinical trials. The principal research objectives are to better understand the properties of such online expert systems and to determine whether they can actually improve the quality of planned clinical trials as compared to traditional methods by randomized controlled trial testing.

2. THE USE OF ARTIFICIAL INTELLIGENCE IN MEDICINE AND CLINICAL RESEARCH

With the advent of artificial intelligence (AI) in 1960s, many industries have moved toward automating tasks that normally would be too monotonous and too time consuming for a human to conduct [20]. This trend is most notably in the production of automobiles, in which the field of robotics came about. Artificial intelligence has also made its way into medicine, in particular, in many medical devices found in the operating rooms and intensive care units of hospitals. It has also been used in the development of algorithms for diagnosing disease and providing optimal treatment strategies. An example of such an AI or 'expert' system is MYCIN for the treatment of blood infections [20].

What are the advantages and disadvantages of AI? A major advantage of using such a system in medicine is that physicians would be better able to identify rare diseases that would otherwise go unnoticed [20]. Another advantage is that it opens up the possibility of patient self-diagnosis for certain non-fatal diseases like colds. However, the disadvantages are that not all

diseases can be placed into a ruled-based format of an expert system. As well, it takes time and effort for such technologies to be adopted by most physicians as clinical protocols.

In clinical research, low level expert systems such as wizards can be found in statistical software packages like Prism, NCSS and SAS where the clinical investigators are assisted with clinical trial design and statistical analysis. As well, there are systems available that assist in the development and implementation of standard operating procedures for large-scale, industrial clinical trials. However, there are few expert systems that try to bring together the clinical trial planning stages of design and implementation for the individual academic investigator with limited resources [21-24]. Hence, an attempt has been made by the authors to further close this gap by developing a web-based expert system for the conduct of small-scale, academic clinical trials.

Why use a web-based application? The Internet is becoming a popular tool to search for information as well as to provide cheap and reliable training to the 'cyber' generation. It has features that enable users to access information anytime and anywhere around the world. Current developments include the immediate access and portability of the Internet through wireless technologies like the hand-held Palms. Such features of "online" and "anywhere" make the Internet attractive, in particular, for systems developers when the ease of access and the necessity of on-demand information are both important.

In academic research, the Internet has already become a commonly used tool by granting agencies to implement their activities. For example, the Natural Sciences and Engineering Research Council of Canada (http://www.nserc.ca) has a web-based application form that is filled out by registered users and then submitted online. The time necessary to submit subsequent applications is significantly reduced since only small changes or updates are necessary to the research experiment or the user's curriculum vitae.

In answering the question "What does an expert system offer to clinical researchers?", an expert system plays an important role as a software program to provide clinical researchers with immediate and sufficient human expert knowledge to aid in solving of problems that occur during the conduct of their clinical research [20]. Such a system is one step up from online applications where, for instance, there is no assistance provided by the granting agency on how the experiment should be conducted. In the proposed system, there is direct knowledge apart from the system help files to assist in the implementation of the research.

In the past, a number of expert systems have been developed for clinical researchers. In 1994, J.C. Wyatt published 'Design-a-Trial' (DaT), which was a knowledge-based, critiquing system that helped physicians in selecting proper trial design and statistics, and later produce a six-page draft protocol [21,22]. In 1997, U. Haaug and group also published a similar system known as the 'PATriCIa Project' but it mainly concentrated on directing investigators in clinical trial statistics [23,24]. Other variants of Wyatt's and Haaug's software appear as 'complex forms' offered by statisticians to clinical investigators in which a design and statistics are chosen from a list of possible research problems or questions of similar interest. When the form is completed, the clinical investigators are usually required to meet with the statistician to discuss the details of the statistics and later implement an action plan for analysis.

However, it becomes apparent that these expert systems and complex forms provide little or no direction to clinical investigators on how to actually implement their research protocols based on the time and resources that they have available. In other words, the necessary training or assistance for developing and implementing standard operating procedures (SOPs) for the ongoing conduct of their clinical trials is lacking.

Academic clinical investigators usually rely upon the SOPs of drug companies and contract research organizations that are provided to them for the conduct of industrial trials according to regulatory guidelines. Hence, clinical investigators that are either inexperienced or have limited resources could benefit from an online expert system that is capable of guiding them through the design, and most importantly, the implementation of a clinical trial when they need it. Laborious tasks such as preparing ethical reviews or granting agency applications could be automated based on past and current information provided to an expert system. At this time, there appears to be little available in terms of a Web-based Clinical Trial Expert System, or WebCTES, for academic physicians or health care professionals.

3. THE DEVELOPMENT OF A WEB-BASED CLINICAL TRIAL EXPERT SYSTEM

In 1991 at the University of British Columbia, the Clinical Pharmacology Research Organization (http://www.cpro.ubc.ca) was created to assist clinical investigators by providing the logistics, personnel and financing to conduct their clinical research. Both undergraduate and graduate students in the health sciences were the primary source of assistance that investigators could rely upon to see their research realized. However, during the years it became apparent that a clinical trial manager was necessary to coordinate both the investigator and the students for each clinical research project. (In industry, the need for management services was also apparent with the hiring of site management organizations, or SMOs, to aid in conduct of clinical trials, in particular, at academic settings). For the clinical investigator with the small budgets, this would mean an expense that would be difficult or impossible to absorb. If there was a system that provided sufficient clinical trial expertise to both the student and the investigator, it was thought that a student could 'manage' the trial using standard operating procedures and the investigator would only have to provide the required time to recruit and screen patients and the necessary clinical expertise to implement the proposed research. Hence, this idea formed the basis of developing a web-based clinical trial expert system at The University of British Columbia.

In the first phase of development, the aim was to create the expert system and evaluate its effectiveness as a tool for generating clinical trial protocols. The expert system covers five main modules: *question, design, statistics, ethics,* and *standard operating procedures.* Each module consists of prompts for questions based on Good Clinical Practices (GCPs). The *question* module helps the user consider the

nature and feasibility of the research problem at hand and then formulates a research question for them. The *design* and *statistics* modules help the user consider the possible experimental design and associated statistical tests in relation to research question being asked. A summary of the reasoning process is then provided. The *ethics* module assists the user in considering the appropriateness of the question, design and statistics and then reminds the user about historical documents related to research ethics, as well as provides ethics committee applications and submission deadlines. The last module, *standard operating procedures*, gives the user a step-by-step approach to implementing their trial from start to finish as well as offers them an audit checklist to assess the quality of the clinical trial. The question-prompts of each module are linked to the system's help topics containing many resource documents and related web links. Once all responses have been gathered through the system's prompts, a draft protocol cab be printed.

In the first version of WebCTES, a Microsoft Word Document Template from MS Office 97 was created, which contained the five modules that were laid out in the format of a clinical trial protocol. The advantages of this framework for the expert system were that it was readily accessible in our laboratory and it appeared to be easy and fast to develop so that pilot work could be undertaken. It also had the unique feature of linking different parts of the document to HELP files, which gave the feel of being on the Internet. Unfortunately, some of the first subjects, who evaluated the system, found it difficult to navigate in the document when wanting to work on different parts. As well, they also found it to be not esthetically pleasing when the template was converted to Word's hypertext mark-up language (HTML) format in order for them to access other resources on the World Wide Web.

Thus, in the second version of WebCTES, a main data acquisition form (MDAF) was developed using Netscape HTML composer with buttons, links, and more windows. With this approach, it was felt that data collected in this form could be imported into a protocol form and other clinical trial-related forms using the Adobe PDF Reader. However, for importing to occur, it was realized that expensive third party software was necessary. Thus, a cheaper alternative was to develop an MDAF using Adobe Acrobat. This approach allowed information that was collected from the user to be saved and then exported into another PDF form that served as the research protocol. Hence, a combination of Adobe Acrobat forms and HTML HELP files provided the framework for third version of the expert system, but it was also too cumbersome to use when evaluated by subjects. Hence, by integrating the features of most commonly available software into one platform over the Internet, the current structure of WebCTES has been optimized (Figure 1).

4. THE SOFTWARE STRUCTURE

Operational Specifications of the WebCTES Architecture

The Main Data Acquisition Form (MDAF) of WebCTES: The system contains a main data acquisition form with 191 fields for the user to enter data. The following rules were implemented: the average number of words allowed for each field must be 200; the standard font size for each answer must be 10 point; the font style must be Times New Roman; the noun clause of each answer must be pre-entered into each field (in other words, the system must start the opening sentence or 'noun clause' that the user needed to finish; the opening sentence would usually be based on structure of the question); and for any field considered 'required' information, the system must alert the user that the field would have to be filled for the system to proceed further.

The Draft Clinical Trial Protocol from WebCTES: Once the data is entered into main data acquisition form, a 20-page protocol is printed. The protocol is based on the concepts presented in the MDAF. The cover page contains the clinical trial title as well as the personal and institutional information of the investigators. This page is followed by a table of contents that is laid out according to the five basic good clinical trial practices (GCPs): question, design, statistics, ethics, and standard operating procedures. In the draft clinical trial protocol, the questions of each GCP presented in the MDAF are turned into headings and subheadings of the protocol. The following rules were implemented for the protocol format: information collected from fields in the MDAF must be entered into corresponding fields of the protocol; fields in the MDAF that are not filled in by the user must not be presented in the protocol; information placed into the protocol from the MDAF must have a 10-point font size; the headings and subheadings must be in numbered-outline format with a 12-point font size; the font style of all text must be Arial, in keeping with most clinical trial industry and granting agency form standards; the header of each page must contain the title for the clinical trial and must be a font size of 8-point; the footer of each page following the cover page must contain the date of when the protocol was created on the left hand side and the page numbers on the right hand side; the font size must be 10 point; and the protocol's table of contents must be updated automatically based on the length of responses entered in the MDAF.

Technical Specifications of the WebCTES Architecture

The minimum hardware requirements for building a server for an online clinical trial expert system were: a Pentium III 400 MHz, 250 MB of RAM, and a SCSI interface three distinct HDDs (containing the OS/WEBCTES, DBase, Backup DBase, respectively - each on a different HDD). The following software components were used to build the online Clinical Trial Expert System:

-Operating System: Linux Redhat 6.0 (kernel de tip 2.2.5 and glib 2.1.1); http://www.linux.org/index.html; This included the Apache server [26,28,29].

-Database: Oracle Database Server 8i release 8.1.5. for Linux [27]; http://platforms.oracle.com/linux/

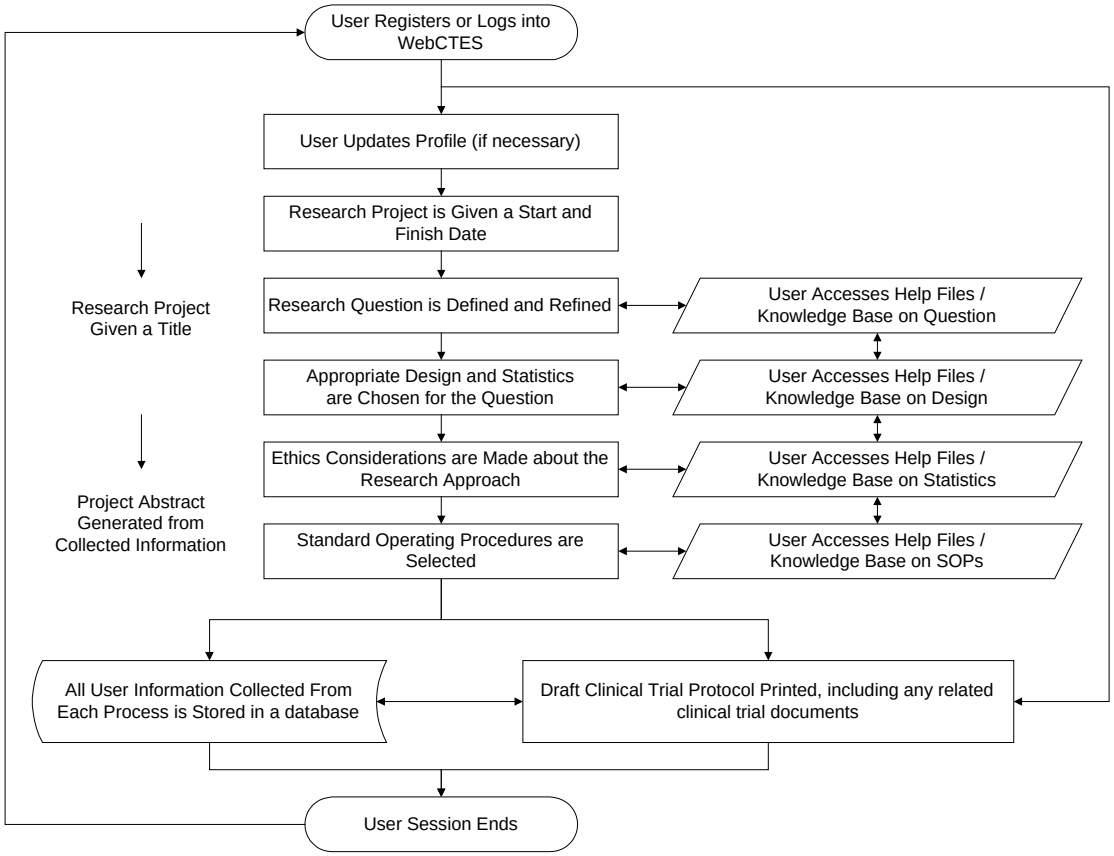
-HMI/Environment*:* Symantec Visual Cafe, Expert edition, http://www.symantec.com/domain/cafe/vcafexe40/

-Latex Document Formatting Compiler

-Post Script to PDF converter

The development of WebCTES was undertaken in one phase and it was regarded as a 'proof of concept', in which the system would be built in such way as to keep the cost, time and material resources to a minimum but support up to 100 users.

Figure 1. WebCTES application



Features of the WebCTES Architecture

Original discussions about the general features of the online clinical trial expert system concentrated at the level of the problem statement, the expert system's help file structure, how it would generate statistics for a given design, and important security features.

With respect to the help files, over 2500 concepts and words were incorporated into the knowledge base of the clinical trial expert system. Three levels of help were instituted:

Beginner Help: Commonly used clinical research terms or word phrases are defined for the user [13].

Intermediate Help: Scientific or clinical research concepts or tasks are defined and examples would be provided. In particular, all Good Clinical Practices are presented [12-19].

Advanced Help: A virtual help center is provided that includes: the implementation of frequently asked questions; database references to books, journals, Good Clinical Practices web page URLs and e-mails as links; an online manual of standard operating procedures; and online consultants that are available as web page URLs and/or e-mail links.

Another feature implemented in the system's help files was the determination of the appropriate statistics and sample size for the specific experimental design in question [16,17]. These statistics were: the Analysis of Variance, Chi square, and Student t-Test. Since these statistics are also available from many academic institutions on the web, their hot links were incorporated into the system in order for users to consider alternative algorithms when selecting test types or choosing a sample size.

User cases or scenarios were also considered, in particular, in the case of user password authentication. The first page of the expert system contains the 'USER ID' and 'Password' fields. It was decided that for this phase of development, the first time user must e-mail the system administrator to request for a password for future verification and entry.

5. THE IMPLEMENTATION

WebCTES was implemented following a 'Rapid Development' process that was mainly consisting of the stages shown in Figure 2.

Joint Application Development (JAD) Sessions

Software engineers met weekly with representatives of the Clinical Pharmacology Research Organization to clearly understand and define scope of the expert system. During this period, the features of the system have been specified and further, initial technical specifications defined. This was a necessary exercise since some of the divergent views had to be focused into a final product, i.e., WebCTES.

Project Scope definition
The outcome of the JAD meetings was a clear list of goals and constrains, further, reviewed and prioritized in the light of an academic type budget and schedule. Hence, a working development protocol was instituted.

Requirements Refinement
A third iteration in defining requirements was perused such as to ensure completeness. As a result, various user-case scenarios were drafted and appended. Other additions generated by this phase were: user registration and certification, user update profile function, tools to search for other certified WebCTES users, and a consulting system used to activate/certify an application.

High Level Design
Once the requirements were finalized, a system high level design was drafted. In this phase, the aim was to develop an adequate design that would allow for fast prototyping. As an example of this procedure, we have created the backbone of WebCTES. This prototyping procedure was intended to become a starting point for the development of other similar applications. The challenge here was to find a cost-effective solution that would be reliable, expandable, and portable.

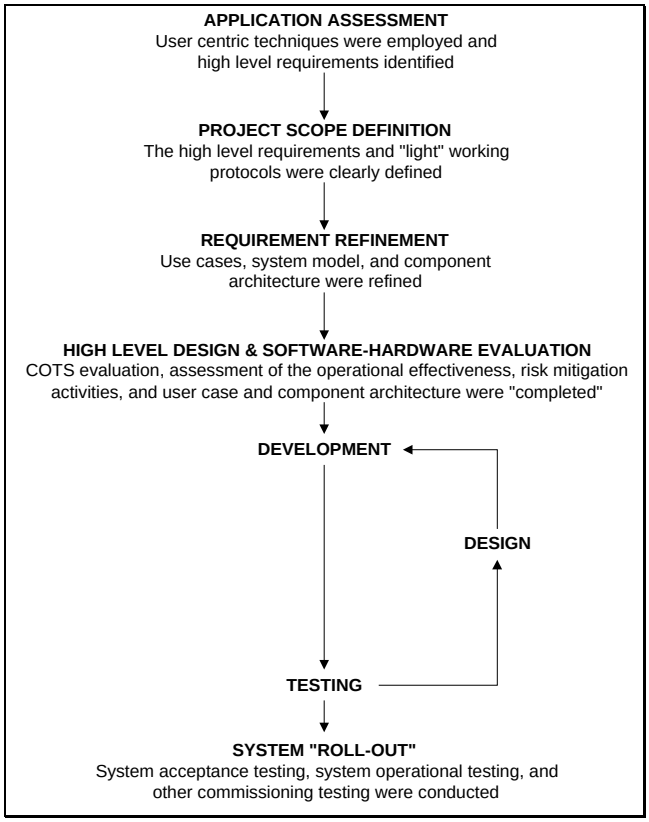
Software and Hardware Selected
In the beginning, Linux was the first choice as an operating system since it offered stability. Other software like the HTTP Apache Server, Latex, and Oracle, were added to the core of the WebCTES system. The first lesson learned at this level was that a significant amount of money was saved during the development procedure through the use of licenses that need to be paid only at deployment. The second lesson taught us that the lack of documentation for Oracle on Linux installation required about two weeks for a complete installation and configuration.

Iterative Development and Testing Phases
The development process was an iterative, incremental and evolutionary approach. The iterative process of "building a little, testing a little" established the framework of the system, based on the priority list defined in the previous phases. This process was beneficial as it allowed users to evaluate the system as it was being built and even more to provide corrective feedback. Through an incremental approach, components were added to extend the functionality of the earlier version. Re-prioritizing requirements such as to cope with limited resources, the list of "must-have" features were achieved during this phase of development. The management of new versus existing requirements was achieved through a “stack type” approach.

Assessing the System’s Performance
A series of experiments were planned to evaluate the system’s functionality and user-friendliness. A randomized controlled clinical trial will be implemented using a grant provided by the University of British Columbia's Teach and Learning Enhancement Fund. The principal endpoints for this evaluation will be system usefulness, capability, ease-of-use, intelligibility, efficiency and reliability in certain user populations. One hundred subjects from various disciplines and work backgrounds will be asked to participate in this testing phase.

System Commissioning and Operational Use
WebCTES was commissioned using a single computer and one Internet Protocol address. A proposed Internet domain for the web site is http://www.webctes.org.

Figure 2. WebCTES Development Process



6. CONCLUSIONS

Based on past and current approaches of applying artificial intelligence technologies to various industries, in particular, the clinical trial development process, a web-based clinical trial expert system, or WebCTES, was developed for clinical investigators at the University of British Columbia. The system directs users in the design, and most importantly, in the implementation of clinical trials by having a comprehensive online knowledge-base of Good Clinical Practices and manual of standard operating procedures, and automation tools for the development of important clinical trial documents, in particular, the clinical trial protocol. Randomized controlled testing will be undertaken to evaluate the functionality and user-friendliness of this expert system in clinical research. Based on the results found, the expert system will be improved through further development and testing.

7. REFERENCES

[1] D.G. Altman, "Statistics and ethics in medical research: Misuse of statistics is unethical", British Medical Journal, Vol. 281, 1980, pp. 1182-1184.

[2] W. Lawrence, Jr., Some Problems with Clinical Trials: James Ewing Lecture, Archives of Surgery, Vol. 126, 1991, pp. 370-378.

[3] R. Smith, "Where is the wisdom...? The poverty of medical evidence", British Medical Journal, Vol. 303, 1991, pp. 798-799.

[4] C.T. Vogelson, "Tide of Trials", Modern Drug Discovery, March Issue, 2001, pp. 29-34.

[5] R.A.K. Nadkarni, S. Antel, and K. Sargent, "The Value of Site-Based Quality Assurance Systems for Clinical Testing Sites", Monitor, Winter Issue, 2000, pp. 29-32.

[6] W. Bohaychuk, G. Ball, G. Lawrence, and Katy Sotirov, "Protocol Content and Management", Applied Clinical Trials, March Issue, 1999, pp. 67-77.

[7] B. Farrell, "Efficient management of randomised controlled trials: nature and nuture", British Medical Journal, Vol. 317, 1998, pp. 1236-1239.

[8] F.O. Winther and O.P. Hole, "Scientific quality of clinical research: an analysis of 40 research projects in pharmacology / pharmacotherapy", European Journal of Clinical Pharmacology, Vol. 51, 1997, pp. 351-354.

[9] D.L. Sackett, R.B. Haynes, G.H. Guyatt, P. Tugwell, Clinical Epidemiology: A Basic Science for Clinical Medicine, Second Edition, Boston: Little, Brown, 1991.

[10] D.L. Sackett and W.M.C. Rosenberg, "The need for evidence-based medicine", Journal of the Royal Society of Medicine, Vol. 88, 1995, pp. 620-624.

[11] K.F. Schulz, I. Chalmers, R.J. Hayes, and D.G. Altman, "Empirical Evidence of Bias: Dimensions of Methodological Quality Associated with Estimates of Treatment Effects in Controlled Trials", Journal of the American Medical Association, Vol. 273, 1995, pp. 408-412.

[12] C. Hutchinson, "History, Development & Application of GCP Guidelines", Monitor, Fall Issue, 2000, pp. 33-36.

[13] L.M. Friedman, C.D. Furberg, and D.L. DeMets, Fundamentals of Clinical Trials, Third Edition, New York: Springer, 1998.

[14] T.A. Lang and M. Secic, How to Report Statistics in Medicine: Annotated Guidelines for Authors, Editors, and Reviewers, Philadelphia: American College of Physicians, 1997.

[15] A. Spriet and T. Dupin-Spriet, Good Practice of Clinical Trials, Paris: Karger, 1992.

[16] D.G. Altman (ed.), Practical Statistics for Medical Research, London: Chapman and Hall, 1991.

[17] P. Armitage and G. Berry, Statistical Methods in Medical Research, Third Edition, Oxford: Blackwell Scientific Publications, 1994.

[18] J.E. Sayre, GCP Quality Audit Manual, Second Edition, Buffalo Grove: Interpharm Press, Inc., 1994.

[19] B. Spilker, Guide to Clinical Trials, New York: Raven Press, 1991.

[20] P. Jackson, Introduction to Expert Systems, Third Edition, Essex: Addison Wesley Longman, 1999.

[21] J.C. Wyatt, D.G. Altman, H.A. Heathfield, and C.F.A. Pantin, "Development of Design-a-Trial, a knowledge-based critiquing system for authors of clinical trial protocols," Computer Methods and Programs in Biomedicine, Vol. 43, 1994, pp. 283-291.

[22] S. Modgil, P. Hammond, J.C. Wyatt, and H. Potts, "The Design-A-Trial Project: Developing a Knowledge-Based Tool for Authoring Clinical Trials Protocols", Paper Presentation, EWGLP 2000, First European Workshop on Computer-Based Support Guidelines and Protocols, Leipzig, Germany, 13-14 November, 2000.

[23] U. Haag, "Knowledge Representation for Computer-Aided Planning of Controlled Clinical Trials: the PATriCIa Project", Methods of Information in Medicine, Vol. 36, 1997, pp. 172-178.

[24] U. Haag, G. Antes, M. Bismarck, R. Haux, R. Holle, A. NeiB and G. Nurnberg, "Knowledge-Based Planning of Clinical Trials", In SoftStat '91, Advances in Statistical Software, pp. 13-18, New York: Gustav Fischer, 1992.

[25] C. Friedman, J. Wyatt, Evaluation Methods in Clinical Informatics, New York: Springer Verlag, 1997.

[26] Peterson, Linux: The Complete Reference Third Edition, New York: Osborne McGraw Hill, 1999.

[27] C. Pribyl and S. Feuerstein, Oracle PL/SQL Programming Second Edition, Cambridge: O'Reilly and Associates, 1998.

[28] D. Flanagan, W. Crawford, J. Farley, and K. Magnusson, Java Enterprise in a Nutshell, Cambridge: O'Reilly and Associates, 1999.

[29] G.K. Koch and K. Loney, Oracle 8: The Complete Reference, New York: McGraw Hill, 1997.

[30] J. Hunter and W. Crawford, Java Servlet Programming, Cambridge: O'Reilly and Associates, 1998.