1995 IEEE-EMBC and CMBEC
Theme 3: Clinical Engineering/Medical Informatics

Knowledge-based support in study design of scientific urological research

Wendy d'Hollosy, Wim P.J. Witjes, Nicolaas J.I. Mars**,
Pieter F. de Vries Robbé*, Frans M.J. Debruyne, Hessel Wijkstra

University Hospital Nijmegen, Dept. of Urology, Nijmegen, The Netherlands.
* University of Nijmegen, Dept. of Medical Informatics, Epidemiology and Statistics, Nijmegen, The Netherlands.
** University of Twente, Dept. of Computer Science, Enschede, The Netherlands.

Abstract - **The process of scientific research in medicine is a complex matter. This process consists, among others, of the design of a study and data management during the study. Computers turned out to be useful tools to support parts of this research process. URIS® is an example of a tool for computerized data management [1,2]. This paper describes the design of a knowledge-based system to support the development process of studies in scientific urological research.**

I. Introduction

In university hospitals research is performed by means of studies to achieve improvements in medical science. An example of a study often used for scientific clinical research is a *clinical trial.* Clinical trials are used to investigate the efficacy and safety of new treatments. Several types of trial design exist, e.g. randomized, blinded (single or double), cross-over, multi-centre. Furthermore clinical trials can be classified into phases [3]. A clinical trial should be designed with utmost care and in detail before its implementation.
The design of a trial is described in the protocol, which provides the rules to be followed during the trial. This includes not only drug schedule, evaluation of response and statistical considerations, but also sufficient information to avoid unwritten decisions and on-the-spot judgements and a scientific and ethical foundation of the study. The protocol has to be judged by several committees on its ethical and scientific contents before the trial can be started.
Designing a trial and writing the corresponding protocol appears to be a very complex matter [4]. Most problems are encountered in:

- coherence between the protocol contents
- unambiguous and complete protocol contents
- statistical design of the trial

These problems did not remain unnoticed. Different groups have designed or are designing tools to improve the design and execution of clinical trial protocols. One example is the ONCOCIN/OPAL system, developed at Stanford. The ONCOCIN/OPAL system is used to review and manage existing cancer trial protocols. The ONCOCIN knowledge base is maintained by the physicians using OPAL [5,6].
Another system is DaT (Design-a-Trial) [4]. This system interviews writers of clinical trial protocols by prompting for the information needed and performs calculations and provides for comments on important statistical aspects of the proposed design. At the end, the system generates a 6-page draft protocol document. The knowledge base of this last system is maintained by knowledge engineers.
We are interested in a knowledge based system that supports all aspects of the design process of new clinical trials and their corresponding protocols. The system should also be able to support parts of this process. Therefore we started the development of a system called PROSYS (= PROtocol design SYStem). This abstract describes the general design approach of PROSYS.

II. Materials and methods

The design process of a clinical trial consists of several parts that are related to each other. The different parts of this process lead to information that should be described in the corresponding protocol. The contents of five approved clinical trial protocols have been analyzed, just as the guidelines of the European Organisation for research on Treatment in Cancer (EORTC) [7] and Good Clinical Practice (GCP) recommendations [3] to determine which information is most important and most frequently found in a clinical trial protocol.

III. Results

10 different categories of information were found that are most important and most frequently found in a clinical trial protocol. Figure 1 shows these different *information categories.* Each category is divided into one or more detailed *information blocks.* For example, the information category *evaluation* is divided into the information blocks *study variables and measurement methods* and *time schedule of recording (pre-treatment, follow-up).* In total 29 information blocks have been defined at this moment.

0-7803-2475-7/97 $10.00 ©1997 IEEE

The determination of the contents of one information block is represented as follows:

input → reasoning process → desired information

in which the *input* is information retrieved from other information blocks, from communication with the end-user or from the PROSYS knowledge base. The *reasoning process* uses the incoming information to determine the *desired information*. This representation is called a *PROSYS part*.

We started to describe the input and the reasoning process for the PROSYS part to determine the desired information for the information block *study variables and measurement methods*. This information block provides for information on data that should be collected during the trial. This information is given in the form of study variables (e.g. weight, WBC) and measurement methods (e.g. bone scan, ECG).

The two main reasons to collect data are to answer the question of the study and to monitor the safety of the studied subjects with respect to the treatments [8]. Therefore the reasoning process of this PROSYS part minimally needs information from information blocks *study question* and *toxicity effects* from the categories objectives and treatments respectively. This information should be given in the form of variables. These are the starting variables of the reasoning process. If, for instance, the study question is: "Is there a relationship between prostatic carcinoma and the prostate specific antigen (PSA) level?" then the starting variables are *PSA* and *prostate cancer*.

The reasoning process uses constraints, given by the end-user, and knowledge from the knowledge base to determine the most relevant study variables and measurement methods for that particular study. The knowledge is a complex network of relations between variables and measurement methods. For example, the PSA level can be determined by different blood tests and prostate cancer can be confirmed by a histopathological diagnosis. The reasoning process reasons through this knowledge network according to the given constraints.

IV. Discussion

Study design is a very complex matter [3] and may introduce problems as described in the introduction. The most common problems will as much as possible be avoided by using PROSYS as a study design support tool in the near future.

Unambiguous and complete information in an information block will be achieved by a careful definition of the reasoning process and the input needed for this process to determine the contents of this information block. The coherence of the protocol will be accomplished by the needed input from other information blocks for most reasoning processes.

A well-considered statistical study design will be achieved by a good definition and implementation of the reasoning processes and their inputs to retrieve the contents of the information blocks of the *statistical considerations* category.

Data management in URIS® will benefit from PROSYS, because in the first instance PROSYS is focused on the determination of the most relevant study variables. Up to now, researchers tend to define superfluous study variables in URIS®. On the other hand, they may discover an overlooked variable during the course of a study. Not well considered study variable specification is another problem that occurs. Support from PROSYS will reduce the occurrence of these problems in the near future. Further, if PROSYS is used to determine the most relevant study variables, the specifications of these variables should be transferred from PROSYS into URIS® to create the study specific database in the corresponding project.

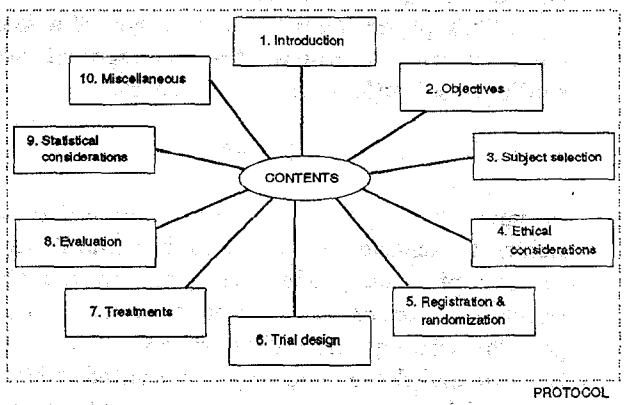

Figure 1. The information categories in a clinical trial protocol.

References

[1] W.d'Hollosy, B.Th.Hendriks, F.M.J.Debruyne, H.Wijkstra, "URIS® - An Universal Information System.", Proc. IEEE-EMBS, vol. 14, pp. 858-859, 1992.

[2] W.d'Hollosy, B.Th.Hendriks, W.P.Witjes, F.M.J.Debruyne, H.Wijkstra, "Improvement of data management in scientific urological research.", Proc. IEEE-EMBS, vol. 16, pp. 1394-1395, 1994.

[3] EEC Note for Guidance: Good Clinical practice for Trials on Medicinal Products in the European Community. Farmacology & Toxicology, vol. 67, pp. 631-372, 1990.

[4] J.Wyatt, D.Altman , C.Pantin , H.Heathfield , R.Cuff, N.Foster , M.Peters , V. Singleton, "Design-a-Trial, a knowledge-based aid for writers of clinical protocols.", Proc. MIE '93, pp. 68-73, 1993.

[5] M.A.Musen, J.A.Rohn, L.M.Fagan, E.H.Shortliffe, "Knowledge engineering for a clinical trial advice system: uncovering errors in protocol specification", Bull. Cancer, 74, pp. 291-296, 1987.

[6] J.W.Egar, A.R.Puerta, M.A.Musen, "Diagrammatic Acquisition of Medical Information for Clinical Trials", Proc. Medinfo'92. Elseviers Science Publishers B.V. North Holland. pp. 1330-1336. 1992.

[7] M.Starquet, R.Sylvester, C.Jasmin, "Guidlines for the preparation of EORTC cancer clinical trial protocols", Eur.J.Cancer, vol. 16, pp. 871-875, 1980.

[8] G.Lawrence, "Case record form design", In: Clinical Data Management. Rondel RK., Varley SA., Webb C. (Eds.). New York. John Wiley & Sons Ltd. 1993. Chapter 9. pp. 133-152. ISBN 0-471-94092-5.