

Management of data from clinical trials using the ArchiMed system

Georg Duftschmid, Walter Gall, Ernst Eigenbauer, Wolfgang Dorda

University of Vienna, Department of Medical Computer Sciences,
Spitalgasse 23, A-1090 Vienna, Austria

{georg.duftschmid, walter.gall, ernst.eigenbauer, wolfgang.dorda}@akh-wien.ac.at
www.akh-wien.ac.at/imc/

Abstract

Clinical trials constitute a key source of medical research and are therefore conducted on a regular basis at university hospitals. The professional execution of trials requires, among other things, a repertoire of tools that support efficient data management. Tasks that are essential for efficient data management in clinical trials include the following: the design of the trial database, the design of electronic case report forms, recruiting patients, collection of data, and statistical analysis. The present article reports the manner in which these tasks are supported by the ArchiMed system at the University of Vienna and Graz Medical Schools. ArchiMed is customised for clinical end users, allowing them to autonomously manage their clinical trials without having to consult computer experts. An evaluation of the ArchiMed system in 12 trials recently conducted at the University of Vienna Medical School shows that the individual system functions can be usefully applied for data management in clinical trials.

Keywords: Clinical trials, data management systems, clinical information systems

Address requests for reprints and correspondence to:

Georg Duftschmid

University of Vienna, Department of Medical Computer Sciences,
Spitalgasse 23, A-1090 Vienna, Austria

Tel.: +43-1-40400 / 6696

Fax.: +43-1-40400 / 6697

Email: georg.duftschmid@akh-wien.ac.at

1 Introduction

Clinical trials are conducted on an ongoing basis at the clinics of the Vienna University Medical School, with a 2000-bed capacity, located within the General Hospital of the city. The professional execution of such trials requires, among other things, a repertoire of tools that support efficient data management [1]. The process of trial data management involves several individual tasks, such as database design, design of electronic case report forms (CRFs), recruiting patients, collection of data, and statistical analysis. Standard applications occasionally used for this purpose, e.g. Access®, Excel® or FileMaker®, rapidly reach their limits when confronted with non-trivial requirements. Furthermore, the data management of complex trials will usually exceed the technical know-how of clinical researchers, if they cannot rely on profound tool support for each individual task involved.

Several commercially available products such as Clintrial® (<http://www.phaseforward.com/>), Oracle Clinical® [2] and MACRO® (<http://www.infermed.com>) allow for comprehensive data management in clinical trials, ranging from small-scale studies to complex multicenter trials. However, the systems are not entirely suitable for our purposes. Features that are essential for our applications, but are either available on a rudimentary basis or are not supported at all by commercial products, include the following:

- Support for clinical end users with little computer expertise: We are interested in systems that permit autonomous management of trials by clinical investigators without having to consult computer experts. System functions implemented in a manner that is too complex to be handled by a clinician include, for instance, the set-up process of a trial's database and the specification of CRFs with their underlying functionality (e.g. for data storage).
- Multi-platform support: Most available systems are customised for a single configuration of hardware and operating system. This contradicts our demand for extensive support of most, if not all, platforms used at our hospital. Multi-platform support is a crucial subject for us, as the computer infrastructure at the Vienna General Hospital is the outcome of historical evolution and therefore very heterogeneous; this feature is probably characteristic of many other health care centres as well.
- Integrated analysis component: While supporting the collection of clinical trial data in a sophisticated manner, most available products neglect built-in analysis functionality. The typical strategy is to provide mechanisms for data export to external statistical packages.

As the utilisation of the latter usually requires advanced statistical knowledge, it is largely reserved for experts in the field.

One of the main goals of the ArchiMed medical information and retrieval system [3] is to assist clinicians in clinical trial data management. In the design of the system, the above mentioned features were taken into account. ArchiMed has been in operation at the University of Vienna and Graz Medical Schools since 1997 and 1999, respectively.

The present paper reports the functionality of ArchiMed in supporting clinical trial data management. Section 2 provides a summary of prior work done in the field of clinical trial data management. Section 3 presents a detailed account of the ArchiMed system: In section 3.1, the key issues that governed the designing and implementation of ArchiMed will be addressed. The characteristics of ArchiMed's trial database and the underlying generic data model will be discussed in section 3.2. Section 3.3 will show, in detail, how ArchiMed supports data management in the following aspects of clinical trials: (1) the design of electronic CRFs, (2) recruiting patients, (3) collection of data, and (4) statistical analysis. Twelve clinical trials recently conducted at the University of Vienna's Department of Surgery and their utilisation of ArchiMed will be discussed in section 4. Section 5 will provide a comparison of ArchiMed with Yales' Trial/DB system [4], which supports clinical trial data management in a similar environment as ArchiMed. A concluding summary will be given in section 6.

2 Background

Computer-based support for data management in clinical trials has been a subject of long-standing research at the Department of Medical Computer Sciences. The WAMIS system employed in 1971 permitted collection of trial data through predefined forms [5]. The WAMAS system [6] allowed data to be retrieved within the WAMIS database, and their subsequent export to statistical analysis packages. The WAMASTAT system [7] was developed in 1982 and supports clinical trial data management in the areas of CRF design, data collection, and statistical analysis. The WAREL system [8] has been in operation at the University of Vienna Medical School since 1992. It receives data from the Vienna General Hospital's information system (HIS) and stores them in a retrieval-oriented manner. Data that are relevant for clinical trials may then be retrieved within WAREL, suitably prepared, and exported to statistical packages.

Several early trial management systems were analysed by Pollak in a review dating back to 1983 [9]. Another early system named MIDBIM (Medical Integrated Data Base Interactive Management) was described by Koval *et al.* [10]. This application integrates functionality such as database definition through the specification of forms, interactive data entry, report definition and report generation, and the execution of simple statistical procedures.

The URIS system [11] was developed to support the management of research data in an academic urology department. It allows forms to be defined by the end user; from these forms the system automatically derives databases. Documentation is supported including functions such as interactive data validation and computable form fields. As output facilities, a report generator and the computation of simple statistics are provided. Complex statistical analyses require that the data be exported to statistical software packages.

The ACT/DB system [12] is a client-server database system employed at the Yale Cancer Centre for clinical trial data management. Recently ACT/DB, which was originally oriented towards operation within a local area network, was reengineered to be operated via the internet. The upgraded version is now known as Trial/DB [4]. As we will show in section 5, the design objectives and implemented features of Trial/DB are very similar to those of the ArchiMed system.

3 The system

The ArchiMed system covers a wide range of study types, including experimental, observational, and diagnostic studies. Trials may be performed prospectively as well as retrospectively. Case-control studies cannot be handled within ArchiMed, as the system does not offer a mechanism for matching controls with each case. To provide full support for randomised trials, we are currently implementing a randomisation component. A web interface, which will allow the system to be applied for multicenter trials, is also being developed.

The design requirements of ArchiMed and how they were implemented will be discussed in section 3.1. Section 3.2 will describe ArchiMed's predefined trial database and the underlying generic data model. Section 3.3 will present four typical data management steps that are performed by a medical researcher in the course of a clinical trial, and how they are supported by the ArchiMed system.

3.1 Design considerations

One of the major goals we had in mind when developing the ArchiMed system was to provide clinical researchers with a tool that would allow them to autonomously manage their clinical trials without having to consult computer experts. We relieved end users of database design and management tasks by providing a predefined generic database that serves as a common data repository for all clinical trials. Further, we set up an intuitive user-friendly interface to all components of the system, which allows access to complex functionality, but conceals a large number of technical details that would be confusing for the user.

To achieve a high degree of portability, we chose the multi-platform development environments VisualWorks® [13] and SAS/AF® (Statistical Analysis System) for the implementation of the system. ArchiMed is based on a client-server architecture, where the graphic user interface and most applications are processed at the client's workplace. Data storage and management are handled by the server. The server is currently operating on an IBM SP2® machine under AIX®, whereas clients cover a variety of operating systems such as Windows®, PowerMac®, Sun Solaris® and most members of the UNIX family. We refrained from implementing the system as a web-based application, as web technology was not sufficiently advanced when the system was being developed (1993-1994).

A further key feature of the system is its ability to integrate data from different sources in order to make a combined analysis. The following capabilities were provided for: (1) To include data from routine patient care in clinical trials (for this purpose an interface to the HIS was created). (2) To integrate scientific data from preceding clinical trials; for this purpose, (a) a common data model was employed for all trials managed within ArchiMed; and (b) several import filters were installed for standard formats, such as Excel® or Tab-delimited data, to integrate data from externally managed trials.

The model was equipped with a security system that controls the access of data and the execution of system functions. Data access is controlled according to a hierarchy of user groups; each user is assigned to a specific group. Data are linked to the hierarchy through their owner; the owner may be a user group or an individual user. Users may access their own data, data of their user group, and data of all groups with a lower security level. For an efficient allocation of function privileges, each user may also be assigned a specific role. Thus, a trial administrator will have the authority to design electronic CRFs, whereas a regular user will only be permitted to enter data into a CRF.

The system provides sophisticated customisation functionality that is based on a hierarchy of profiles. Each clinic or research group may define its particular settings within its own profile. Examples of customisable properties are the system language, the size of CRFs when they are opened, the default owner of created data, and the output style of printouts. Frequently used CRFs and medical variables may be defined as personal component groups that are automatically loaded for fast access. Furthermore, individual system modules may be specified as preferred functions to be automatically opened at system start-up.

3.2 Design of the database

The ArchiMed system's database was implemented using Oracle®'s database server. The main design goal to be met was that the clinical end user should not have to set up a new database for each new trial. Instead, a predefined central database would be employed as the common data repository for all clinical trials. This required a generic data model that would have to be suitable for the storage of any trial data regardless of their origin (different forms within the system or even data imported from external systems).

Data model

When designing the data model, the central question was how to map the data for the different medical variables, registered on a form, into database tables. We decided against a model consisting of a single table that contains all variables as columns, since most mainstream relational database engines prescribe an upper limit for the number of columns per table (e.g. Access® permits 255 columns per table). Besides, this design would have required modification of the table structure whenever a new variable is created. Furthermore, a table of this kind would largely consist of null values, resulting in considerable wastage of space, as a mere fraction of the entire variable set is affected by a single insertion in most cases.

The standard approach of most report tools, wherein a new table is created for each new form, was also discounted: typical analyses involve data from more than one form. Using this approach it would not be feasible to predefine generic, form-independent queries. Every predefined query would have to be updated each time a new form is created that comprises a variable which is referred to within the query.

The data model we use allows dynamic expansion of medical variables and forms, and also permits predefinition of generic form-independent queries. It is based on the Entity-Attribute-Value (EAV) design [14-15], which allows a generic and uniform storage of the data and metadata originating from all clinical trials within an institution. EAV data models are used by several well-known systems, such as the HELP system [16], the Columbia-Presbyterian Medical Centre's clinical repository [14], Yale's Trial/DB system [4], and Oracle Clinical® [2]. The EAV design aims at representing a single entity's (e.g. patient's) data as multiple rows in a single table with few columns, rather than as a single row of data spread across multiple tables and columns. The term EAV refers to the data model's strategy to hold within

one row information on the entity (patient ID, document ID, etc.), the attribute (ID of variable being recorded), and the value of the variable.

[Insert figure 1 about here]

The core of our data model (see figure 1) is composed of the following elements:

- The variable catalogue represents the set of medical variables, for which patient data may be acquired. They may be semantically grouped within variable categories.
- A form allows data for a set of assigned variables to be collected.
- Several versions may exist for each form, representing its modification history. One version is designated as the form's current version.
- When placed on a form (version), a variable becomes a data field in which data can be entered. A data field has several attributes that concern the representation of its variable on the form (e.g. position on form, font, ...). It is further associated with a particular widget (e.g. input field, combo box, ...) that defines the data field's input behaviour.
- Numerical variables are associated with a unit category (e.g. length, weight), consisting of units that are convertible within their category. Within each category, one unit is designated as the category's standard unit to store the corresponding data uniformly. For each data field referring to a numerical variable, one unit is selected for displaying the corresponding values.
- A form (version) filled with patient data becomes a document when stored.
- A document consists of (text, number, code, date, time, timestamp) values, where each value refers to one data field and therefore also to one variable.
- Cases may be recorded for a patient, each associated with a set of documents. Documents may also be directly attached to patients, or may be left unassigned.

3.3 Trial data management functions

The following four data management tasks are usually performed in the course of a clinical trial: (1) design of electronic CRFs, (2) recruiting patients, (3) collection of data, and (4) statistical analysis. In the following sections we will discuss how each of these steps is supported by the ArchiMed system.

3.3.1 Design of electronic case report forms

The ArchiMed system is equipped with a Form Designer component to facilitate the laborious process of CRF designing. No programming know-how is needed to use this tool. With the Form Designer, clinical end users can create highly functional database-related forms by simple parameterisation of form widgets. Most technical details are automatically executed in the background, invisible to the user. Thus, clinicians may independently design the forms they need without having to consult computer experts. In the following, a few highlights of the Form Designer's functionality are described:

- **Component reuse:** The system is equipped with a central repository that contains components such as medical variables, 'choice lists', and classification schemes.
- **Manual and automatic form design:** Forms can be automatically generated using pre-selected variables if the user wants to minimise the time being spent on the design phase. Manual designing is also provided for (e.g. selection of the optimal widget, positioning of variables, choice of colours and fonts), if the user wishes to create a more sophisticated layout.
- **Conditional data entry:** Using a wizard, the user may define conditions that allow the dynamic activation or deactivation of individual variables or complete form pages that depend on the values of other variables during documentation. For instance, a variable pertaining to the stage of pregnancy might be automatically disabled for male patients.
- **Formula-based variables:** To allow automatic calculation of variables that may be derived from other variables (e.g. body mass index), the system is equipped with a formula editor that offers an extensive choice of logical, arithmetic, statistical, and temporal functions.

- **Query-based variables:** The ArchiMed system allows the user to pre-fill variables in a dynamic fashion, using data obtained from database queries. Pre-existing individual values in the database can be included in a new document, thus avoiding errors that would result from retyping the data. Furthermore, new values may be calculated by applying functions to a set of existing data, such as the mean value of all data recorded for a variable within a certain time interval for the current patient.

As a final element of the design, ArchiMed forms are stored in the central database server; this allows trial data to be entered immediately from any work station in the hospital. Forms are cached on the client's side to speed up instantiation.

3.3.2 Patient recruiting

The ArchiMed system offers two approaches to recruit patients for a clinical trial:

- **Use of formula-based variables and database triggers:** The trial eligibility condition is implemented by means of a form including a formula-based variable that generates the eligibility decision, depending on a set of input variables. If data for these input variables are expected to enter the system's database from other sources (e.g. other forms, integration of external data), the trial eligibility criteria can be installed as triggers in the system's database. Persons whose data satisfy a particular trial's eligibility condition are then added to a list of potential trial participants and are reported to the respective investigator.
- **Use of the retrieval component:** When all patient data required for a trial's eligibility decision are already present in the database (e.g. retrospective trials), suitable patient cohorts may be retrieved by specifying the appropriate queries within the system's retrieval component (see figure 2). These queries may include arithmetic, logical and temporal operators.

After a patient has been found to be eligible for a trial, he/she may be added to a list of enrolled patients. The system then ensures that, whenever a patient enrolled in a trial presents to a physician, the CRFs to be filled in are automatically displayed on the physician's screen. Hereby, clinicians are informed that they must review these forms and add or alter data as required.

[Insert figure 2 about here]

3.3.3 Collection of data

Trial data are recorded by filling in electronic CRFs and storing them as documents in the database (see figure 3). The process of data collection can be customised in many ways. For instance, certain forms may be automatically loaded and displayed for individual users. Thus, CRFs will be provided ready for documentation each time an investigator starts the system. The documentation process is supported by a set of tools and utilities that include the following:

- **Predefined choice lists:** Data may be selected from predefined choice lists for variables that have been structured as combo boxes or radio buttons; thus the user is saved the effort of typing. The use of predefined choice lists also enhances the quality of free text data, as typographical errors and different notations are avoided.
- **Interactive data validation:** A further mechanism to ensure good data quality is ArchiMed's interactive consistency checking function. Based on the properties of the variables (e.g. data type, value range), data are checked for consistency at the time of entry. Besides, data entry may be enforced for specific variables by defining them as 'required fields'.
- **Data coding:** The system is equipped with a coding wizard that allows identification of a specific code by entering its associated text as a search criterion. The code may then be directly imported from the wizard to the corresponding variable on the CRF. Variables for the collection of coded data may be linked to textual variables, so that the text associated with each code is automatically displayed. Currently, the ICD-10 classification scheme and a selection of local schemes are supported.
- **Formula-based variables:** Formula-based variables are automatically calculated by the system as soon as all required input data are present on the form.
- **Query-based variables:** Query-based variables may be automatically loaded by the system when a form is opened, or may be allocated upon request by pressing a button on the form. A special wizard may be used to copy data from existing documents to the current form. Data originating from query- as well as formula-based variables may be manually overridden by the user if desired.

[Insert figure 3 about here]

The scientific data collected within ArchiMed CRFs can be further supplemented by routine data that is imported from external systems (mainly the HIS). A further type of data import was specifically implemented for the Department of Transplantation; it allows donor data to be reused for the recipients of organ transplantations. After setting the patient ID of the donor in the recipient's form, the system automatically loads relevant donor data.

The system is equipped with a browser that allows rapid and efficient documentation of a specific form for a patient cohort (e.g. for post-documentation of partially filled CRFs). After selecting a patient cohort and a form, either new documents can be created or existing documents can be updated, based on this choice.

3.3.4 Statistical analysis

The analysis component allows clinicians who are untrained in statistics to correctly resolve their standard analysis problems; it provides a well considered selection of statistical functions in a predefined form (see figure 4). These functions can be executed via an intuitive interface, and cover commonly used methods such as descriptive statistics (means, extremes, medians, frequencies, etc.), statistical tests (pair-wise comparisons, the chi-square test, etc.), and graphic data visualisation methods (histograms, plots, etc.) to elucidate biometrical relations. The functions were selected on the basis of 15 years of experience in the statistical analysis of medical data with the system WAMASTAT [7].

[Insert figure 4 about here]

As the analysis component uses the SAS® format to store data, experienced users may further process their results in the regular SAS® environment. This is especially useful for handling non-standard complex analysis problems. The statistical modules SAS/INSIGHT® and SAS/ANALYZE® may be applied to the data as well as autonomously developed SAS®-procedures. For detailed information about the system's analysis component, refer to [17].

4 Results

ArchiMed is currently employed by about 270 clinical users from four clinics at the University of Vienna and Graz Medical Schools, where the system has been in operation since 1997 and 1999, respectively. Since this time, approximately 50 complex clinical trials have been managed within the system; some of the trials spanned a period of more than 12 months. From the 282 forms that were created, more than 100 are exclusively used for research purposes. The central repositories at Vienna and Graz each contain about 5000 medical variables. Including the data integrated from the HIS and preceding clinical trials, the databases of the Vienna and Graz installations together comprise more than 1.3 million documents containing 28 million values from over 320 000 patients. Most of the clinical trials managed within ArchiMed were conducted by the Department of Transplantation at the University of Vienna Medical School. This department was our closest medical partner when the system was being developed. Their typical strategy for managing data from clinical trials is as follows:

- **Design of electronic case report forms:** The forms designed within ArchiMed are general enough to provide basic data for several clinical trials addressing similar research issues. For instance, a form is designed for the documentation of every new kidney transplantation.
- **Patient recruiting:** To identify patients with relevant characteristics for a new clinical trial, ArchiMed's retrieval component is used.
- **Collection of data:** Data are collected by filling in the above mentioned forms. Occasionally complementary data may be imported from previous studies.
- **Statistical analysis:** Prior to statistical analysis, the collected data usually have to be pre-processed in order to customise them for the particular statistical procedure. This is either done within ArchiMed or by using external tools (usually Excel®) after the data have been exported. Excel® is chosen in cases where the pre-processing of data requires complex spread-sheet functionality. For statistical analysis, ArchiMed's analysis component is generally used in combination with external applications (usually SPSS®). SPSS® is given preference when graphic output is to be post-processed (e.g. for a presentation), as it offers advanced functionality for this purpose.

Twelve trials recently conducted at the University of Vienna Medical School have been analysed in table 1 with regard to their utilisation of ArchiMed in the tasks of CRF design, patient recruiting, data collection, and statistical analysis. It reveals that CRF design and data collection were performed within ArchiMed in all cases. Patient recruiting was performed within ArchiMed in all but three cases, in which existing patient cohorts of earlier trials could be reused. ArchiMed's analysis component was typically used in combination with external applications for statistical analysis of trial data.

[Insert table 1 about here]

5 Discussion

Table 2 compares the functionality of ArchiMed with that of the Trial/DB system; the two applications have a number of similarities. The advantages of ArchiMed lie in its broader support for client platforms, a more advanced form design tool, and its integrated analysis component. Trial/DB has the advantage of being operable via the internet (web-based forms, SSL support), and thus provides an efficient platform for the management of multicenter clinical trials.

[Insert table 2 about here]

ArchiMed provides physicians with a high level of flexibility for data management in clinical trials. Nevertheless, it must be made clear that a corresponding organisational structure is an absolute requirement for the use of a flexible system of this type. Since clinicians are given the technical ability to directly introduce adaptations into the system 'on site' (e.g. by defining new or modifying existing variables and forms), the associated organisational structure must also be on site. That is, an authorised clinician should be present in each department to play a coordinating role.

6 Conclusion

The ArchiMed system aids medical research by providing tools that support the management clinical trial data. ArchiMed's main strengths are the following: (1) The system is customised for clinical end users, allowing them to autonomously manage their clinical trials without having to consult computer experts. (2) The system provides a predefined central database to store all trial data, based on a generic data model. Thus, the end users need not set up and manage a database for each individual trial. The database also features an interface to the HIS,

which allows the user to include data from routine patient care in clinical trials. (3) ArchiMed offers a sophisticated tool for an interactive design of electronic CRFs. (4) For the analysis of trial data, the system provides a powerful research component that offers a multitude of predefined statistical methods, including temporal operators. Based on the generic data model, it is possible to predefine database queries that address data from all trials within the system.

Limitations of the ArchiMed system exist in the following points: (1) ArchiMed is currently not entirely suitable for use in multicenter clinical trials. Two tasks remain to be solved: (a) The system should be equipped with an http-based interface to the internet, which will allow remote data collection by means of web forms. (b) The security system must be extended to allow safe operation via the internet. So far, the system has been operating behind a firewall. (2) At present ArchiMed does not provide built-in randomisation functionality.

Extension of the system in the near future will mainly focus on its ability to handle multicenter clinical trials. The reason is that there is a great demand for such a system at the University of Vienna Medical School, where numerous multicenter trials are being conducted and coordinated on an ongoing basis. For optimal coverage of multicenter clinical trials, a randomisation component is currently being integrated in the Form Designer tool. This will allow clinical end users to generate randomisation forms. We are also implementing an http-interface to the system, which will allow web-based entry of data. Work on the randomisation component and the web interface is scheduled to be completed by the end of 2002.

Acknowledgements

Financial support for the project described in this paper was provided by the ARGE/AKH - VAMED. The authors thank Gerald Boldt, Peter Sachs, Wilhelm Tensch and Thomas Wrba, who contributed to the execution of the ArchiMed project. We are also grateful to Jan Bialy, Martin Bodingbauer, Gerhard Prager, Harald Puhalla, Susanne Rockenschaub and Martin Schindl for informing us of their strategies of clinical trial data management.

References

- [1] Tai, B. C., and Seldrup, J., 2000, A review of software for data management, design and analysis of clinical trials. *Ann Acad Med Singapore*, 29, 576-581, 5.
- [2] Oracle, 2000, *Oracle Clinical Version 4.0 Application-wide Activities*, (Redwood City, CA; Oracle Corporation).
- [3] Dorda, W., Wrba, T., Duftschmid, G., Sachs, P., Gall, W., Rehnelt, C., Boldt, G., and Premauer, W., 1999, ArchiMed: a medical information and retrieval system. *Methods Inf Med*, 38, 16-24, 1.
- [4] Brandt, C. A., Nadkarni, P., Marengo, L., Karras, B. T., Lu, C., Schacter, L., Fisk, J. M., and Miller, P. L., 2000, Reengineering a database for clinical trials management: lessons for system architects. *Control Clin Trials*, 21, 440-461, 5.
- [5] Grabner, H., 1985, *WAMIS Wiener Allgemeines Medizinisches Informations-System*, (Berlin-Heidelberg-New York-Tokyo; Springer Verlag).
- [6] Wolf, W., Dorda, W., Sachs, P., and Graber, E., 1981, WAMAS: On line evaluation of a medical database under consideration of time-relations. *Annals of the World Association for Medical Informatics*, 231-235.
- [7] Laminger, B., Dorda, W., and Reichetzeder, C., 1986, WAMASTAT - ein System zur Benutzerfuehrung bei der Auswertung klinischer Studien. *EDV in Medizin und Biologie*, 17, 1-5, 1/2.
- [8] Dorda, W., 1989, WAREL: a system for retrieval of clinical data considering the course of diseases. *Methods Inf Med*, 28, 133-141, 3.
- [9] Pollak, M. N., 1983, Computer-aided information management systems in clinical trials. A physician's perspective. *Comput Programs Biomed*, 16, 243-251, 3.
- [10] Koval, J. J., Kwarciak, L. M., Grace, M. G., and Lockwood, B. J., 1987, A comprehensive database management system for a variety of clinical trials. *Methods Inf Med*, 26, 24-30, 1.
- [11] d'Hollosy, W., Walrave, M. H., Hendriks, B. T., Debruyne, F. M., and Wijkstra, H., 1995, Semi-automated database design by the end-user. *Methods Inf Med*, 34, 266-271, 3.

- [12] Nadkarni, P. M., Brandt, C., Frawley, S., Sayward, F. G., Einbinder, R., Zelterman, D., Schacter, L., and Miller, P. L., 1998, Managing attribute-value clinical trials data using the ACT/DB client-server database system. *J Am Med Inform Assoc*, 5, 139-151, 2.
- [13] Cincom, 1993-2001, *VisualWorks Application Developer's Guide*, (Cincinnati, Ohio; Cincom Systems Inc.).
- [14] Friedman, C., Hripesak, G., Johnson, S., Cimino, J., and Clayton, P., 1990, A generalized relational schema for an integrated clinical patient database. *Proc Annu Symp Comput Appl Med Care*, Washington D. C. 1990, pp. 335-339.
- [15] Nadkarni, P. M., 1997, QAV: querying entity-attribute-value metadata in a biomedical database. *Comput Methods Programs Biomed*, 53, 93-103, 2.
- [16] Huff, S. M., Haug, P. J., Stevens, L. E., Dupont, R. C., and Pryor, T. A., 1994, HELP the next generation: a new client-server architecture. *Proc Annu Symp Comput Appl Med Care*, Washington, USA. 1994, pp. 271-275.
- [17] Gall, W., Sachs, P., Duftschmid, G., and Dorda, W., 1999, A retrieval system for the selection and statistical analysis of clinical data. *Med Inform Internet Med*, 24, 201-212, 3.

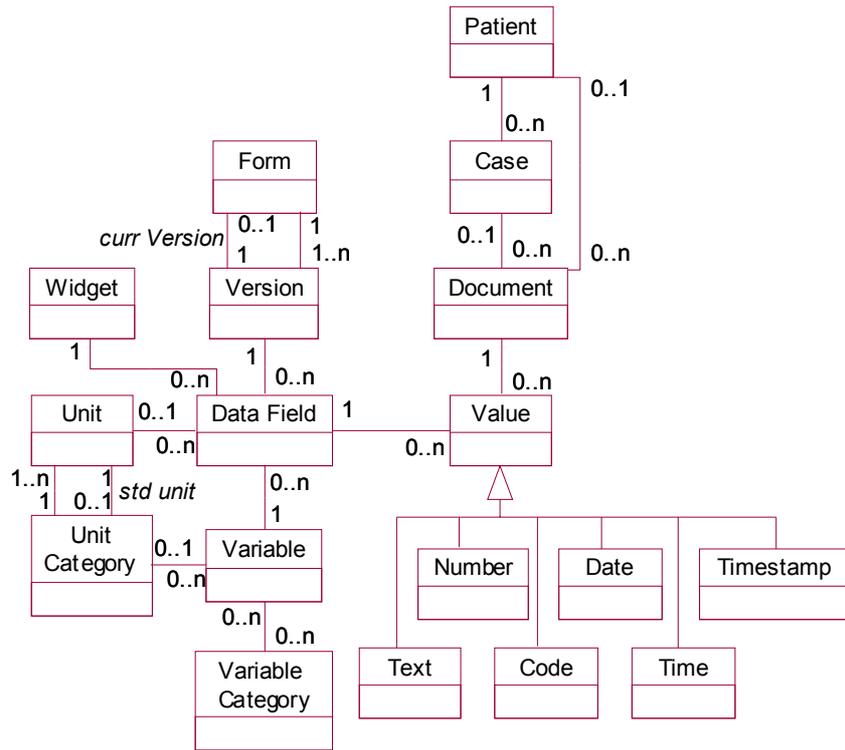


Figure 1

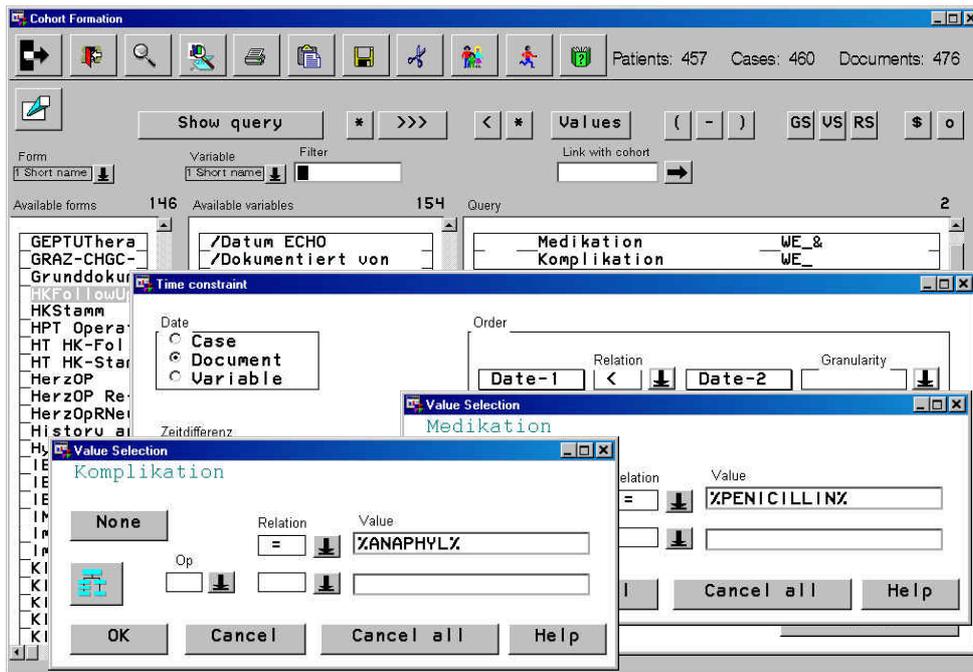


Figure 2

Lebermetastasen (96495428)

Document Edit Table Special Help

Time 06-09-2001 19:38

<no case selected> AC Leber

Leberresektion Zeitpunkt

synchron
 synchron-Intervall
 metachron
 Wochen (bei Intervall)

Operationsdatum: 06-09-2001
 Operationsindikation: elektiv

OP Therapie: Resektion

OP - Zusatztherapie: **Kryotherapie**

Port i.a.:
 ja
 nein

Operationsdauer: 440

Pringle Manöver:
 ja - komplett
 ja - selektiv
 intermittierend
 nein

Vasculäre Exclusion:
 ja
 ja - mit Präservation
 nein

Resektionsart:

- Segmentresektion IV
- Subsegmentresektion VIII

Simultanoperationen:

- Cholecystektomie
- Splenektomie

andere:

- Alkoholinstillation
- Kryo+Alkohol
- Pfortaderhypertrophie
- Methylenblauinjektion

Alter bei C:

Primärerkrankung: Lebermetastasen
 Lebermeta: Chirurgie
 Lebermeta: Histologie
 Periop. Verlauf
 Aktuelles Follow-Up

MODIFY

Figure 3

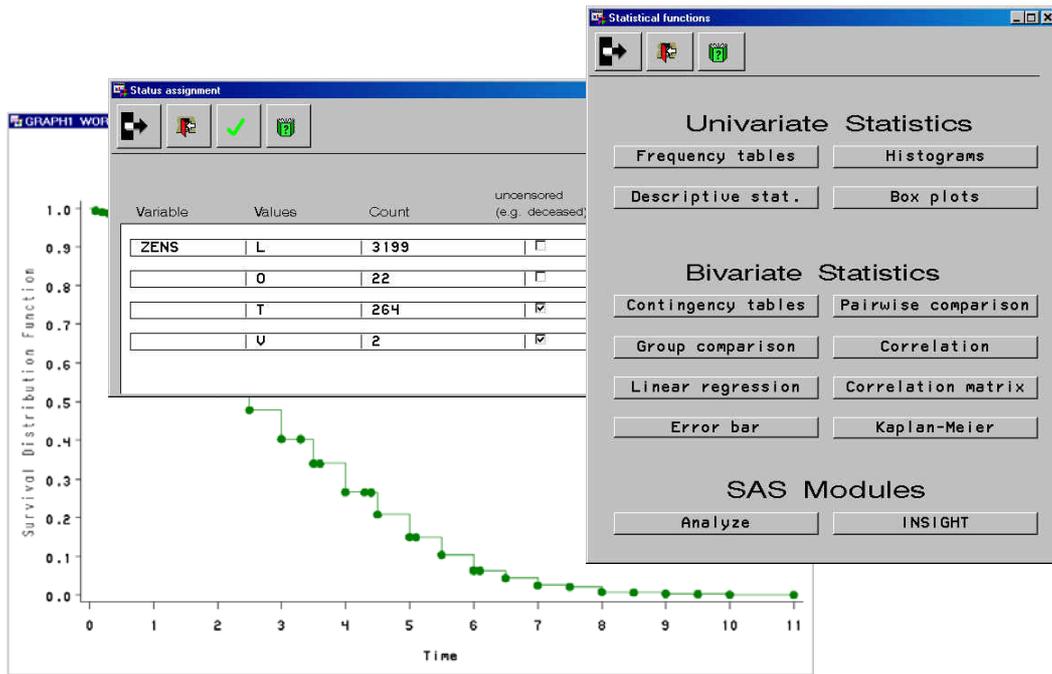


Figure 4

Table 1: Recent trials managed within the ArchiMed system ⁱ

Research subject of the trial	Data management steps performed within ArchiMed			
	CRF design	Patient recr.	Data coll.	Analysis
Department of Transplantation				
Prognostic factors in the course of resecting hilar cholangiocarcinomas	✓	✓	✓	Φ
Evaluation of various aspects of the Eurotransplant Senior Program	✓	✓	✓	✓
Tacrolimus vs. cyclosporine microemulsion in liver transplantation	✓	✓	✓	Φ
Preservation solutions for transplantation	✓	✓	✓	Φ
Living donation in kidney transplantation	✓	✓	✓	Φ
Bile duct complications after liver transplantation	✓	✓	✓	Φ
Hep C recurrence in liver transplant patients	✓	✓	✓	Φ
Perioperative chemotherapy in liver transplant patients with hepatocellular carcinoma	✓	✗	✓	✓
Ureter complications after kidney transplantation	✓	✓	✓	Φ
Department of General and Cardiothoracic Surgery				
Performance of different types of heart valve replacements	✓	✓	✓	✓
Resection techniques of liver metastases in connection with colorectal cancer	✓	Φ	✓	✗
Effectiveness of minimally invasive open parathyroidectomy in an endemic goiter area	✓	✗	✓	✗

ⁱ (✓, 'Φ', '✗' stand for 'fully', 'partly', 'not' performed within ArchiMed)

Table 2: Comparison of the ArchiMed and Trial/DB systemsⁱⁱ

	ArchiMed	Trial/DB
General properties		
System architecture	Client-server	Client-server
End user	Clinician	Clinician
Server-platform	UNIX / Oracle	UNIX / Oracle
Client-platform	Windows, Mac, UNIX, Sun	Windows
Security	User/password, access and function privileges, firewall	User/password, access and function priv., firewall, SSL
Web support	✘	✓
Database design		
Predefined, central DB	✓	✓
Data model	Entity-Attribute-Value	Entity-Attribute-Value
Design of electronic CRFs		
Reuse (CRFs, variables, ...)	✓	✓
Interactive form design	✓	✘
Form → database	Direct mapping	Intermediate 'flat' table (client)
Recruiting patients		
Retrieval component	✓	✓
Collection of data		
Data validation at entry	✓	✓
Dating	Valid time / Document	Valid time / Value
Statistical analysis		
Analysis component	✓	✘

ⁱⁱ ('✓', '✘' stand for 'supported', resp. 'not supported')

Figure 1: ArchiMed core data model

Figure 2: The retrieval component allows the user to specify conditions that will help identify patient cohorts. In the session shown here, all patients who suffered anaphylactic shock after receiving penicillin are to be retrieved.

Figure 3: Data entry through electronic CRFs in the ArchiMed system.

Figure 4: The ArchiMed system's analysis component provides a selection of statistical functions in a predefined form (right window). The depicted session shows the computation of a Kaplan-Meier plot (left window) for the survival analysis of selected patient cohorts; the documented values for the status variable can be assigned to the groups censored, uncensored or to be ignored (centre window).