Report on ObTiMA
as a GCP conformant
Software Application

Project Number: FP6-2005-IST-026996

Deliverable id: D2.6

Deliverable name: Report on ObTiMA as a GCP conformant software application

Submission Date: 23/07/2010

COVER AND CONTROL PAGE OF DOCUMENT	
Project Acronym:	ACGT
Project Full Name:	Advancing Clinico-Genomic Clinical Trials on Cancer: Open Grid Services for improving Medical Knowledge Discovery
Document id:	D 2.6
Document name:	Report on ObTiMA as a GCP conformant software application
Document type (PU, INT, RE)	PU
Version:	1.0
Submission date:	23/07/2010
Editor: Organisation: Email:	Norbert Graf UdS graf@uks.eu

Document type PU = public, INT = internal, RE = restricted

ABSTRACT:

The intention of this deliverable is to report on the GCP criteria to make ObTiMA GCP conformant and to build the basis for certification of ObTiMA to use in GCP conform Trials. Requirements for GCP conformant software applications and tools are explained and an overview of the functionalities of ObTiMA including the Trial Outline Builder (TOB) is given. The steps towards a GCP conformant ObTiMA are listed and described. The certification of ObTiMA is not part of ACGT. Only important documents are listed how to proceed in ongoing projects to fulfil these criteria.

KEYWORD LIST: ObTiMA, GCP, Data Management system, Trial outline builder, ACGT platform, Certification

MODIFICATION CONTROL			
Version	Date	Status	Author
0.1	10/02/2010	Draft	Norbert Graf
0.2	02/05/2010	Draft	Holger Stenzhorn
0.5	30/06/2010	Draft	Holger Stenzhorn
1.0	23/07/2010	Final	Norbert Graf

List of Contributors

- Norbert Graf, UdS, Homburg, Germany
- Holger Stenzhorn, UdS, Germany
- Fatima Schera, IBMT-FhG, St. Ingbert, Germany
- Micke Kuwahara, UHok, Saporro, Japan

Contents

1 Executive Summary

The intention of this deliverable is to report on the GCP criteria to make ObTiMA GCP conformant and to build the basis for certification of ObTiMA to use in GCP conform Trials. Requirements for GCP conformant software applications and tools are explained and an overview of the functionalities of ObTiMA including the Trial Outline Builder (TOB) is given. The steps towards a GCP conformant ObTiMA are listed and described. The certification of ObTiMA is not part of ACGT. Only important documents are listed how to proceed in ongoing projects to fulfil these criteria.

2 Requirements for GCP conformant software applications and tools

Introduction

The term Good Clinical Practice, or GCP for short, refers to a set of internationally recognized and accepted rules which follow established ethical and scientific quality standards for mandating how to design, conduct, monitor, audit, analyze and report clinical trials in which human subjects participate and which are intended to be submitted to some regulatory authorities. Having their origin in the Declaration of Helsinki, the main focus of those rules lies on the protection of the rights, integrity, and confidentiality of all subjects within a trial as well as on the quality, credibility and accuracy of the results gained.

The GCP forms part of several so called GxP guidelines for Good Work Practice for developing and manufacturing medical and pharmaceutical products or devices. GCP encompasses not only rules but also provide guidelines giving an in-depth description of the various roles of all parties involved in a clinical trial. Also the documents are specified which are necessary during a clinical trial, like the protocol or the investigator's brochure. In order to protect the trial subjects, the GCP further defines how the subjects' informed consent is established and how to react if unexpected side effects or serious adverse events arise.

The rules of the GCP are not merely propositions that can be followed but are now also mutually accepted as a unified standard in the countries of the European Union, Japan and the United States and have become part of binding, enforceable jurisdiction and regulations. The World Health Organization as supra-national organization is further promoting the acceptance and compliance to the set of GCP rules in other parts of the world.

To reach the goal of the GCP, both quality assurance and control have to be applied: Via the former systematic actions are planned and established to ensure that a clinical trial is performed and its data are generated, documented (recorded), and reported in compliance with the GCP and the applicable regulatory requirements, and the latter represents the operational techniques and activities undertaken within the quality assurance system to verify that the requirements for quality of the trial-related activities have been fulfilled.

In this deliverable, we are concerned with those parts of the GCP rules which relate to using software applications and tools in clinical trials. Specifically, we describe how the ObTiMA (Ontology based Trial Management Application) itself fulfils those rules making it compliant with the GCP. For example, we show how all records describing the methods, conduct, and results of a trial are safely and securely stored. We also show how ObTiMA supports the user in adhering to the GCP requirements. For example, the system keeps track of all the essential trial documents that permit evaluation of the trial conduct and the quality of the produced data, like that all information is recorded and stored to allow accurate reporting, interpretation and verification of the trial.

Basic Considerations

In the following section we describe some of the fundamental issues that are necessary to be considered when trying to make ObTiMA compliant to the rules of the GCP.

Essential Documents

The essential documents of a clinical trial are the set of all documents that make it possible for e.g. the regulatory authorities to evaluate the trial conduct and the quality and integrity of the data produced in it. They make also possible to document the compliance of all involved parties with the GCP rules and with all applicable regulatory requirements and thus greatly assist in the successful management of a trial by the investigator, sponsor and monitor.

Those documents encompass documents collected before, during and after running the clinical trial. The trial's master files are created at the beginning of the trial at all participating parties. A clinical trial can only be formally closed when those files have been successfully monitored and the existence and appropriate placing of all documents has been confirmed.

Of particular importance is the trial protocol describing the objective(s), design, methodology, statistical considerations, and organization of a trial usually together with is background and rationale. (Amendments to the protocol are possible to in written form to introduce changes or formal clarification.) A second highly important document is the trial report holding detailed clinical and statistical descriptions, presentations, and analyses of the trial and highlighting any therapeutic, prophylactic, or diagnostic agent used on human subjects.

Informed Consent

In order for a potential trial subject to be able to decide whether to participate in a clinical trial or not, he or she must voluntarily confirm the willingness to take part in that trial, after having received all relevant information about the various trial aspects. Informed consent is documented in a written, signed and dated form and has to be obtained from every subject prior to clinical trial participation. During the conduct of a trial, informed content also regulates the sharing of all gained data and information, i.e. who besides the treating physician can see a patient's medical data within a trial.

Electronic Trial Data

There are some particular issues that deserve consideration when electronic systems are employed for (remotely) entering and handling data, such as in the case of ObTiMA. As these issues are the basis of this deliverable, the following is explained in more detail below:

- Validate and document that ObTiMA conforms to the established requirements for completeness, accuracy, reliability, and consistent intended performance

- Develop and maintains Standard Operation Procedures (SOP) for using this system

- Enable an audit trail within ObTiMA that ensures that all changes to data are documented and that there is no deletion of entered data

- Create a security system that prevents unauthorized access to any of the data

- Maintain a list of the individuals who are authorized to make data changes

- Create systematic and adequate backups of all data

- Make sure that subject blinding is enabled whenever data is entered or processed

- Set-up the possibility to compare data if it is processed with original data

- Create an unambiguous identification codes for all trial subjects allowing the identification of all the data reported for each subject

Good Software Engineering

To ensure that the above points are fulfilled, a proper process for good software engineering needs to be in place. This includes planning, verification, testing, traceability, configuration management, and many other aspects, to support the validation of the application. Computer applications that are used to create, modify, and maintain electronic records and to manage electronic signatures are subject to the validation requirements. They must be validated to ensure accuracy, reliability, consistent intended performance, and the ability to discern invalid or altered records.

Software for the above applications may be developed in-house or under contract. However, software is frequently purchased off-the-shelf for a particular intended use. All production and/or quality system software, even if purchased off-the-shelf, need to have documented requirements defining its intended use, and information against which testing results and other evidence can be compared, to show that the software is validated for its intended use.

Also, a proper life cycle and risk management for the development process needs to be included in this as well. Based on the intended use in clinical trials of ObTiMA, the above points represent a highly critical issue, and thus a strong effort combining several techniques and approaches (following the GCP guidelines) have to be applied. The software validation and verification activities are to be conducted throughout the entire software life cycle, i.e. not only during the main development but also already during its design or testing phase.

3 Overview of Functionalities of ObTiMA including TOB

Introduction

Clinical Trial Management Systems (CTMS) promise to help researchers in hospitals and biotechnology/pharmaceutical companies to better manage the tremendous amounts of data involved when conducting clinical trials. Their goal is to simplify and streamline the various aspects of clinical trials, such as planning, preparation, performance, and reporting, by providing functionalities, like automatic deadline tracking for legal or regulatory approval, progress report issuing, keeping participant information up-to-date, or import/export data from/into other clinical information systems. For example, it is still a common yet tedious and error-prone practice to collect data at each trial site on paper-based Case Report Forms (CRF) and then to enter them manually into the trial database at the trial center. CTMSs are supposed to avoid this by providing user interfaces that blend into clinical work settings and shield users from underlying data and system complexity.

But as standardized, commercial CTMSs are not yet widely deployed, trial databases and their entry interfaces are often developed in-house specifically for a given trial and therefore not readily reusable in other trials. This issue causes an additional reimplementation burden and makes it difficult to compare or integrate data between different trials. But even if CTMSs are used, the following issue remains unresolved: Those systems allow a user to freely define the CRF items and structures without the need of any informatics skills. But although this is very desirable, it can create the same interoperability problems. If a database is derived from the trial-specific CRF definitions, the database in turn is again also trial-specific and data reuse in further research stays problematic. Thus, our work focuses on solving this interoperability issue through an approach based on ontology and semantic (data) mediation.

Components

Trial Builder

The Trial Builder is one of ObTiMA's two main components (cf. Fig. 1) and enables the user to specify the various aspects of a trial. The trial outline and metadata can be defined in a master protocol based on templates for describing the trial goals and its administrative data, like start/end date. Treatment plans can be graphically designed to guide clinicians through the treatment of individual patients and particular treatment events, like chemotherapy or surgery, can be defined with all necessary information. The particular order of treatments for each patient can be defined by placing them on a timeline. Also, treatment stratifications and randomizations to be applied for a patient can be described. For each stage on the treatment plan a CRF can be assigned to collect the data documenting the treatment.

Ontology-based CRF Creation

The creation of CRFs marks the core functionality of the Trial Builder. In a graphical user interface, the user can define the content, layout, and navigation of the CRFs which are used to capture all patient data during a clinical trial, like the patient's history, medical findings, diagnostic data, or genomics data.
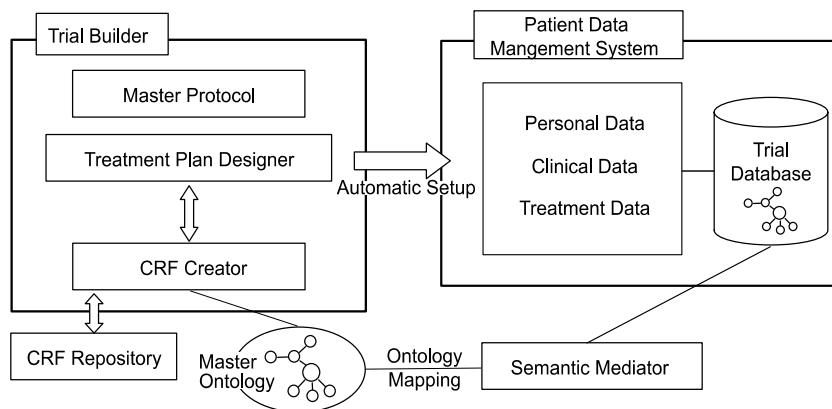
Figure 1 – ObTiMA System Components

It is important that all information can be defined here which are necessary for the data integration, i.e., each CRF item is described based on ontology concepts together with metadata, like data type and measurement unit, to set-up the trial data-base. However, the internal CRF (data) representation is not the focus of clinicians but their "user interface" (layout) and their adaption and integration into the specific workflow of the planned trial: clinicians are not to be bothered with the under-lying aspects of the trial database or the ontological metadata. Thus all these aspects are made transparent to the user through a graphical user interface which hides the actual complexity yet gathers all required information for automatically creating the trial database. This interface is derived also automatically from the content and structure of the Master Ontology but presents a simplified ontology view, adapted to the task of creating items (cf. Fig. 2). It comprises the following sections:

In the Ontology View (1), the user selects concepts from the ontology to create a CRF item. The interface tries to bridge the gap between clinical practice and the logical representation of ontology concepts: Although the ontology provides natural language descriptions for its concepts/relationships (in addition to the logical definitions), those often do not fully mirror the needs of practical or clinical perception of reality. In order to meet this need, we do not present the full Master Ontology here but rather a simplified clinical view which contains a trial-independent basic classification of CRF contents from a clinician's viewpoint.

It is by intention that the clinical view is far less detailed as the actual Master Ontology and since this allows the possibility to provide a much easier entry point for the user. The interface of the clinical view is implemented as a tree always that starts at node of the concept "Patient" as focus of any clinical study (and hence CRF) and only presents those concepts that are directly reachable from this concept, like "Weight" or "Tumor" (indicating a patient's tumor). Only when a concept is selected then also the concepts directly reachable from this one are shown, such as "Laterality" in the case "Tumor" was initially chosen (indicating the laterality of the patient's tumor).
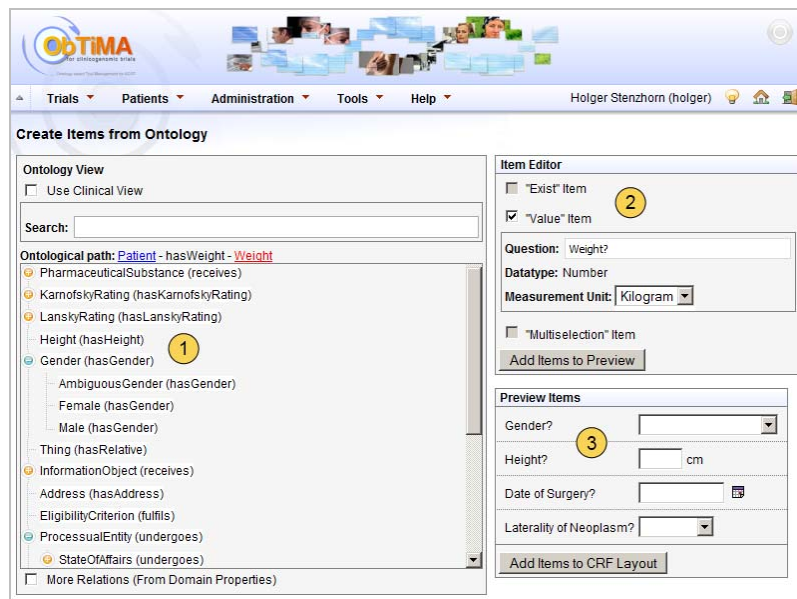
Figure 2 – Ontology Viewer while Creating CRF Items

When a concept is chosen in (1) then a corresponding item is automatically created and shown in the Item Editor (2) together with its attributes determined automatically based on the chosen concept, such as label, data type, or answer possibilities, and which can be manually adopted. For example, the concept "Weight" has a numerical data type and a list of suit-able measurement units attached. So, when the CRF with this item is used in a clinical trial then the measurement units are offered as selection possibilities (in a drop down menu). The specified value (entered into a text field) is automatically tested to be of numerical type and also to be non-negative (since a weight cannot be negative). Finally, Preview Items (3) presents all created items in the order in which they are intended to appear on the CRF. Single items can be reordered by simple drag and drop and subsequently transferred to the interface where the overall layout of the CRF is then defined in turn.

CRF Repository

Revisiting the reuse and interoperability issue discussed in the introduction, in many trials similar or equal data are collected, yet stored differently because of different data(base) definitions. Applying the Master Ontology already improves this situation through using standardized concepts when creating CRFs. Going a step further, the situation would be further improved by partial or complete reuse of existing CRF in case similar data is collected. This idea realized by creating a unified CRF Repository as crucial part of ObTiMA. This repository allows the storage and retrieval of entire ontology-based CRFs and single CRF items or components for reuse and adaption in subsequent trials: When setting-up a clinical trial, fit-ting CRFs can either be directly reused or new ones quickly created by "plugging together" existing CRF items and components. This in turn fosters the standardization of CRFs even more, since CRFs can now be compared not only on the level of single items (through their basis on ontological concepts) but also on the level of larger components or in their entirety.

Patient Data Management System (PDMS)

The PDMS supports clinicians when conducting a clinical trial and is automatically set-up based on the master protocol and CRFs defined in the Trial Builder. The PDMS guides the clinicians through the actual treatment of patients according to their individual treatment

plans and provides a graphical user interface to fill in the CRFs relevant to the patient's current treatment situation. The interface is adjusted to everyday clinical needs: As with the Trial Builder, the complexity of the underlying ontology is hidden from the user, yet its logic-based concept definitions are used to provide direct validity checking when CRF are filled in. The basic look of the data entry interface corresponds to section (3) on Fig Y with each input element providing on-the-fly feedback about its current state based on the just mentioned checking, i.e., in case a negative value is specified for a weight then this error is immediately highlighted along with an explanation of the error.

Data Export

To integrate ObTiMA into real-world clinical settings, the system must be capable to interface with other existing CTMS and be able to exchange data in a format they understand. To meet this requirement, ObTiMA allows to import and export trial metadata, CRF descriptions and patient data through an extended version of the CDISC Operational Data Model (ODM) format [8]. This platform-independent, quasi-standard for exchanging and archiving clinical trial data is supported by many current CTMSs. Observing CDISC's extension guidelines, we enriched this format by allowing the additional inclu-sion of (metadata) descriptions based on Master Ontology concepts. In the case other CTMSs want to import data generated by ObTiMA, they can chose to interpret the supplemental descriptions but if this is not feasible the resulting data is still "ODM complete" and can sensibly be used by those systems.

Administration, Security and Pseudonymization

To administer multicentric clinical trials, ObTiMA contains several advanced facilities for managing the multitude of institutions, researchers, and patients usually participating in such trials. An elaborated and fine-grained security architecture has been implemented to handle the rights and roles that can be attached to the system's users in order to guarantee that they can only perform the tasks which they are fully authorized for. It is also straightforward to dynamically react to changes with-in a running clinical trial, since new institutions and users can always be added or extra security roles and rights be defined.

It is also indispensable that ObTiMA, as a system holding real patient data, securely stores all of the data which could possibly identify some patient to non-authorized persons in pseudonymized and encrypted form. To foster security even more, such personal data is physically separated from the actual clinical research data through the use of two distinct database servers: One server holds the database for storing the personal data of the patients, such as their names and addresses (which must never be shared, e.g., via the Semantic Mediator). The protection of this database strictly follows all current legal regulations for data protection in clinical environments. The other server hosts the database that contains the actual research data collected in a clinical trial (through the use of the CRFs). It is possible in the Trial Builder to mark CRF items as personal which results in this data being stored in the database for personal data and not in the one for research data.

Trial Outline Builder

A highly visualizing software that on every plane of a clinical trial's existence will aid and guide any type of user who needs to participate in the trial at any level and at any period of the clinical trial process.

This software has been given the name Trial Outline Builder (TOB) and aims to do a number of things, where just a few is to simplify the building and planning of a clinical trial, ease the pains involving the collection of data on patients being treated within the trial and not least supporting the researchers who need to analyze the trial data in order to increase the understanding of a trial's outcome and further improve future treatments.

TOB is a Web application, running within the borders of any optional Internet Browser. This allows it to have full access to any data storage available online as well as being available itself from any computer anywhere in the world. No hassle with installations or data moving. The only need for TOB to work is that the Microsoft Silverlight free-ware plug-in is installed, which can be achieved within a minute by a simple mouse click.

TOB considers everything to be an interactive object, an object that can be dragged and dropped, whose attributes can be altered, which can be combined, saved, restored and much more. The most apparent objects in TOB for a first-time user would probably be the Medical Event objects together with the work area windows and the timeline.

TOB is originally designed to be used in four different distinctive purposes in mind, all depending on user types and the specific times of the trial lifespan. These four begin with the design level (the level 1), where a system manager designs different treatment event objects and treatment elements, followed by the Trial Outline Building level (the level 2) where a trial chairman construct a new trial with its treatments and time plan. Third comes the treatment level (the level 3) where the patients are treated and the CRF data are collected by the local physician, and finally the data analysis level (the level 4) where researchers can analyze the trial data and visualize them in multiple ways.

Level 1 – Design and Admin Level

If a user activates the TOB in level 1 mode he/she will then be allowed to add new medical events or edit existing medical events. This process includes everything from drawing its appearance to configuring its inner structure. A medical event is exactly what it sounds like, something that is medical in nature and might happen to a patient during a trial. These events can be planned or they can be accidental. They are used to describe the path of a patient's treatment during the trial.
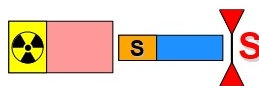

Figure 3 – Small range of Medical Event Objects

A medical event could for example be radio therapy treatment, surgery, chemo therapy or trial specified events like randomization or stratification. A medical event is described by simple properties like the color and the name but also by more complex ones like the duration, drug type, dosage and distribution by time. These complex properties are considered objects in themselves and are referred to as treatment elements. The process of creating these medical events is a simple matter of drag and drop in combination with attribute manipulation via self explanatory forms and mouse maneuvers.

Other features at this level for those with the proper rights to enter TOB in this mode are to configure, for example, the database connection, where all input data will be stored while working with TOB, and to add or block features available in other levels of the application.

In this level TOB also allows the user to add completely arbitrary objects, to alter or improve the user experience while using TOB (as explained below).

Level 2 – Trial Chairman Level

The trial chairman uses TOB at this level to structure a treatment plan, meaning to configure a set of different type of treatment combinations over a certain period of time. Two windows are available to the chairman, one working as a menu or repository of medical events and

another working as a work- or drawing board. The user drags medical events from the repository and places them on the drawing board in order to build the treatment time plan.

Also visible on the drawing board is a relative timeline that is automatically updated for each newly placed medical event to show the full course of the trial in number of days.
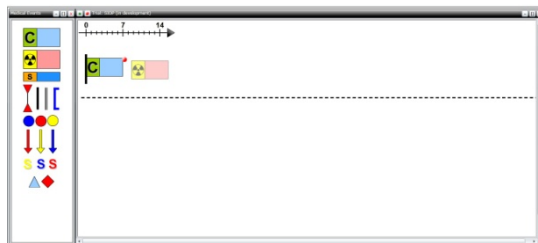

Figure 4 – Building a treatment plan

All medical events can be arranged and rearranged all to the users will, events or a group of events can be duplicated or deleted. A red marker guides the user to available positions in the treatment tree for adding a new event. Every change is recorded and saved automatically to the main trial database online for safety and convenience. Each event can have its attributes/properties edited; changing colors, adding labels, altering durations and time offsets and most frequent settings of the treatment elements.
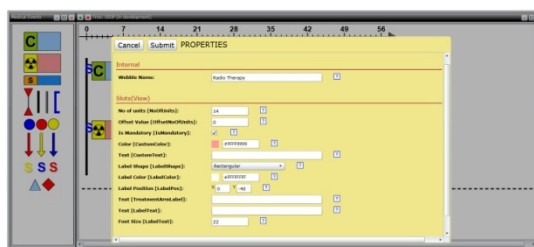

Figure 5 – Property form where a medical event can be edited

Some medical events acts like forks in the treatment flow where the treating physician is supposed to make a choice for each patient which path or treatment arm to take, others are grouping several treatments together to be executed in parallel. Some treatment elements lack duration but work more like milestones or additional data collectors.

Data collection is done as usual by filling in CRFs. Each medical event can have one or more CRF associated with it. In this level of TOB the trial chairperson is supposed to setup the connection between certain CRFs and certain events. These CRFs may be found in the CRF repository or he/she can create new CRFs, either from scratch or by editing existing ones. The process of associating CRFs with medical events is just a simple matter of few mouse clicks and the process of creating a CRF is just performed by using a simple CRF creation editor which is a part of the TOB experience.
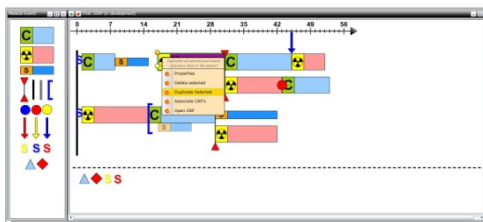
Figure 6 – Associating CRFs to a medical event and
an example of how a treatment plan could look

Some events are not supposed to be planned but should only be available for the treating physician in case of some problem or deviation from the original plan. SAE (Severe Adverse Events), SUSAR (Suspected Unexpected Severe Adverse Reaction) events, time delay reports or emergency surgeries could be some examples of such events. They will be placed below the actual treatment plan but otherwise will be prepared and edited like other events.

When a Treatment plan is finished in the level 2, its user can then activate the trial, allowing clinicians to run it for treating each patient and for acquiring his or her treatment data. An activated trial cannot be edited in level 2 anymore for obvious reasons.

Level 3 – Treating Physicians Level

As soon as a trial has been activated it is available in TOB's level 3 mode where physicians carry out treatment to patients and that treatment is recorded in each event's CRF.

In this level the TOB view shows two windows, one with the original treatment plan as created by the trial chairperson, and the other showing the actual treatment for a specific patient. The patient is selected at the time of entering level 3 but can easily be changed to other patients in this trial treated by the current physician. Patient data are protected in the database and can only be seen by the authorized medical staff for the particular patient.

The Timeline in the treatment plan window is relative as in the case of the level 2, while the one in the patient window is absolute, showing the actual date for each medical event. The red vertical line in the patient window shows the current date and therefore also indicating which medical event the patient is supposed to undergo at the moment.

Another difference of the treatment plan in the level 3 mode is that some events are faded. If an event is faded in white it means that this event might be a future event for the specific patient, while if it is faded in black, means that this event will never be a part of this patient's treatment path. Non-faded events are events which the patient has passed or is passing through. All these functions are to give the user a quick overview of the treatment path for a specific patient compared to the original plan.

Figure 7 – Clinical trial treatment plan with a patient that has not yet begun any treatment

The physician will select the first medical event the patient is about to undergo in the treatment plan view and activates it for the patient. It will then appear in the patient's window where it can be manipulated. The main manipulation, and most often the only one is the filling out of the CRFs associated with each event. The physician clicks a patient event to open a Case Report Form and then answers the questions accordingly.
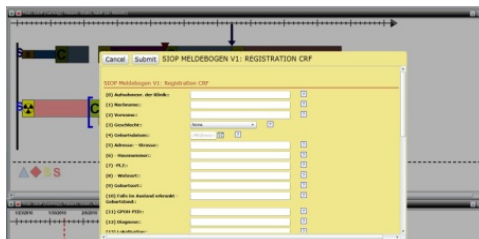

Figure 8 – Ready to fill in a CRF

Each event is marked with a small triangle which indicates the status of the CRF, Red indicating that the CRF is empty and green indicating it is done. Orange means on-going and black means no CRF available (the later probably not that common).

The Physician can also edit the patient specific treatment path if something unplanned occurs or if some medical event could not start on time or had to start earlier. This is easily done via the property form mentioned in level 2 and/or by adding patient specific events from the lower part of the treatment plan, for example a SAE/SUSAR event. Except for patient specific events optionally added by the physicians, the events which correspond to the actual events in the treatment time plan cannot, quite naturally, be moved or rearranged since that would render the trial data for that patient fairly useless in comparison with other patients.
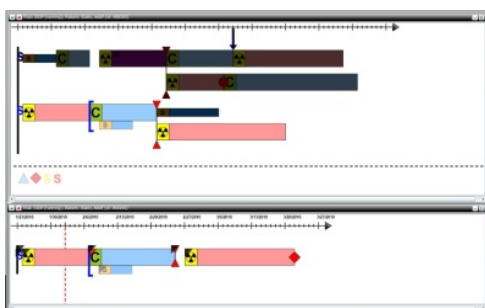

Figure 9 – Example for patient treatment path

Level 4 – Investigating Researchers Level

When TOB is running in level 4, where the user is probably a researcher of some kind or a statistician for clinical trial analyses, its goal is to help the user to visualize all the data collected so that he or she can make reliable conclusions and analyses which hopefully can lead to new and more effective treatments.

In this mode the user sees two windows to start with, one shows the original treatment plan and the other shows the data analyzing and querying view using interactive parallel coordinates. The researcher clicks on the medical events of interest and opens the CRF related with that event. In the CRF viewer the user then selects those questions (i.e. CRF attributes) which should be analyzed. When a question is activated as described above, a parallel-coordinate value-range object is displayed in the query window for that question. Every patient who is subject to that question in the specified CRF of the specified event is

marked in the corresponding parallel coordinate and forms a line from this parallel coordinate to the other parallel coordinates.
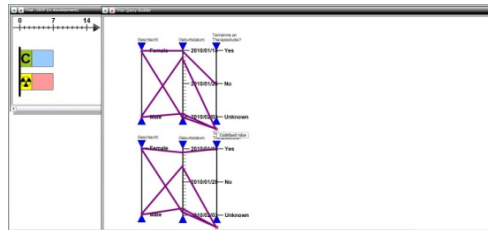

Figure 10 – Small example of data visualization of a tiny trial and treatment plan

The parallel coordinates can be manipulated easily with the mouse to filter patients, or order them according to the user's visualization need. Each treatment path query is grouped together and can be moved around so that they can be compared in various and informative ways. For this purpose colors of patient-representing lines can be edited and group opacity can be changed.

If the user creates a query whose visualization he/she finds to be especially valuable, it can easily be stored for later use or presentation. The data visualized in each parallel coordinate can also be combined with and presented in various charts. Users can also mirror them back to the original treatment plan by overlaying diagrams showing for example treatment success rates and other valuable views in relation to each treatment arm.
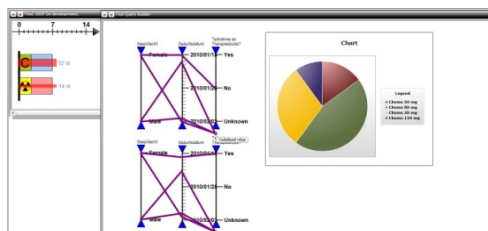

Figure 11 - Additional visualization features in TOB at level 4

4 Steps towards making ObTiMA GCP conformant

Audit Trail

The term audit refers to a systematic and independent examination of all activities and documents that relate to a clinical trial for determining whether all activities were conducted, and data recorded, analyzed and reported according to the protocol, sponsor's standard operating procedures (SOPs), the GCP, and the legal requirements.

It is a documentation that makes possible the reconstruction of the course of events and is implemented as software-based process where all user activities are continuously monitored and recorded. First, this allows truthfully reconstructing the interaction of the users with the system and, second, facilitate the recovery of data in the case of errors or exceptions. For example, an audit trail makes possible the targeted testing and correction of possibly erroneous user inputs.

The result of an audit trial is an audit report which combines all collected data in written form that can be used, for example by the sponsor's auditor to verify the results of some given clinical trial. The investigator or institution should permit monitoring and auditing by the sponsor, and inspection by the appropriate regulatory authorities.

In the concrete case of ObTiMA, we too record every change performed to the data in the system: So any change or correction to a data are dated, marked with the changers username, the actual change performed, and (if necessary) explained. This does not obscure the original entry or the changes that have been performed before. The result of the audit trail can be viewed, given the necessary user privileges (like for a trial chairman) are given in form of a tabular view.

Data Security and Privacy

Data security is the means of ensuring that data is kept safe from corruption and that access to it is suitably controlled. Thus data security helps in ensuring privacy and in protecting personal data

Data privacy is the relation between collecting and disseminating data and the legal issues surrounding them. Privacy concerns exist wherever personally identifiable information is collected and stored in digital form or otherwise. Improper or non-existent disclosure control can be the root cause for privacy issues. The challenge in data privacy is to share data while protecting personally identifiable information.

To avoid such issues is obviously of highest importance for clinical trials. So the goal is to keep all private data in system confidential and prevent disclosing the identity of a trial subject to other than authorized individuals. The confidentiality of records that could identify subjects must always be protected, respecting the privacy and confidentiality rules in accordance with the applicable regulatory requirements.

Therefore only people that have received appropriate permission to examine, analyze, verify, and reproduce any records and reports that are important to evaluation of a clinical trial, are allowed to do so. It needs to be specified in the trial protocol or other written agreement that the investigators and institutions provide direct access to source data/documents for trial-related monitoring, audits, reviews, and regulatory inspections. Any party with such direct access should take all reasonable precautions within the constraints of the applicable regulatory requirements to maintain confidentiality. It must be verified that each trial subject has consented, in writing, to direct access to the original medical records for trial-related monitoring, audit reviews, and regulatory inspection.

Since ObTiMA is based on remote data entry and uses the Internet to create the communication between the actual server and its remote client, we are using HTTPS. Hypertext Transfer Protocol Secure (HTTPS) is a combination of the Hypertext Transfer Protocol with the SSL/TLS protocol to provide encryption and secure identification of the server and are used for sensible transactions (such as payments) over the Internet and for sensitive transactions in corporate information systems.

Roles and Rights

A user role defines the tasks, properties and rights of a particular user in a software application. User roles are used so that the aforementioned settings do not to have to be set individually for each user. Instead of directly assigning rights to users, a user role is defined, which in turn can then be assigned to many users. This facilitates the rights management especially since after changes in the rights' structure only the rights of the user's role must be adapted. Roles are a conceptual development of user groups. A user can have multiple roles and their rights arising subsequently by the union of the rights of all roles. Modern software systems provide the user with a customized graphical user interface to its users' roles.

ObTiMA is accessed by many different users with heterogeneous capabilities and responsibilities and thus best care must be given to managing the security administration. In our system we therefore employ a role based access control, as just described, as it is a powerful mechanism, yet reduces the complexity of security administration. Access to all operations within ObTiMA is regulated by this mechanism. Each user is assigned to one or more roles characterizing their purpose and function within ObTiMA. Permission to access to several sections of the system might be granted or not according to the role of each user. ObTiMA's access control mechanism contributes in complying with all applicable statutory and regulatory requirements, ensuring the security of personal data held in the database.

Figure 12 shows the roles that are available within the system and which can be attached to a user of the system. Figure 13 then shows part of the list of possible rights which can be combined together to form a role.

Name	Description
ROLE_ADMINISTRATOR	Administrator
ROLE_STATISTICIAN	Trial Statistician
ROLE_TRIAL_STUDY_NURSE	Trial Data Manager / Study Nurse
ROLE_REFERENCE_SURGEON	Reference Surgeon
ROLE_REFERENCE_PATHOLOGIST	Reference Pathologist
ROLE_REFERENCE_RADIOLOGIST	Reference Radiologist
ROLE_REFERENCE_RADIOTHERAPIST	Reference Radiotherapist
ROLE_REFERENCE_MOLECULAR_BIOLOGIST	Reference Molecular Biologist
ROLE_DSMC_MEMBER	Member of the DSMC (External Data Security and Management Committe)'
ROLE_TRIAL_MONITOR	Monitor of the trial
ROLE_LOCAL_PHYSICIAN	Local Physician caring for a specific Patient
ROLE_LOCAL_PHYSICIAN_ON_DUTY	Local Physician on duty
ROLE_LOCAL_STUDY_NURSE	Local Management / Study Nurse
ROLE_LOCAL_MOLECULAR_BIOLOGIST	Local Molecular Biologist
ROLE_CONSULTANT	Consultant

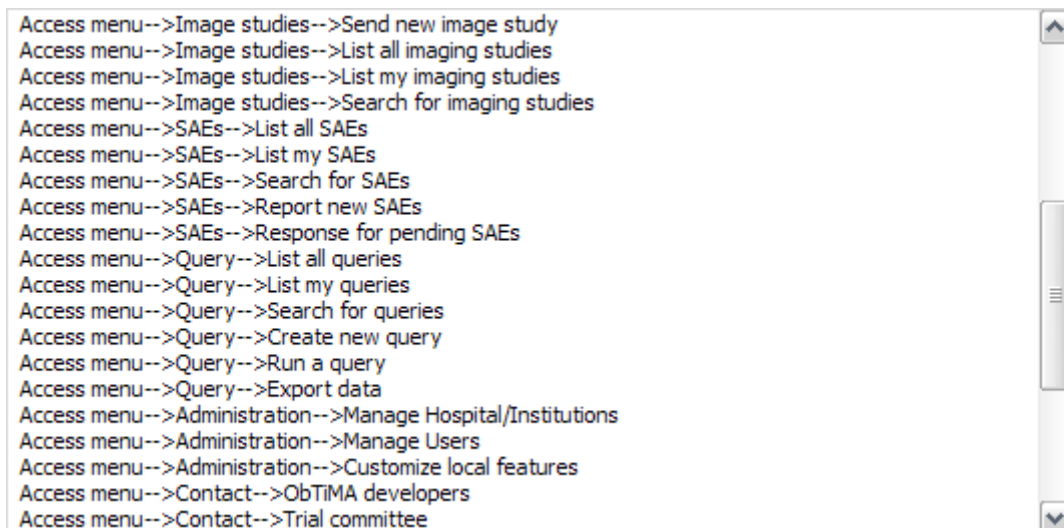Figure 12 – Roles used in a regular clinical trial

Figure 13 – Rights that can be assigned to a role

Data Encryption

Encryption is the process of transforming information using some algorithm to make it incomprehensible to anyone except those who have special key to decrypt it again. In ObTiMA, all patient-related data are fully encrypted based on the strong encryption "Secure Hash Algorithm", more accurately its SHA-256 version. The key that is used to encrypt and decrypt is highly protected and safely stored away from the data: Thus even if the (encrypted) data would be "stolen" then they would be of no value because they could not be decrypted into clear text values.

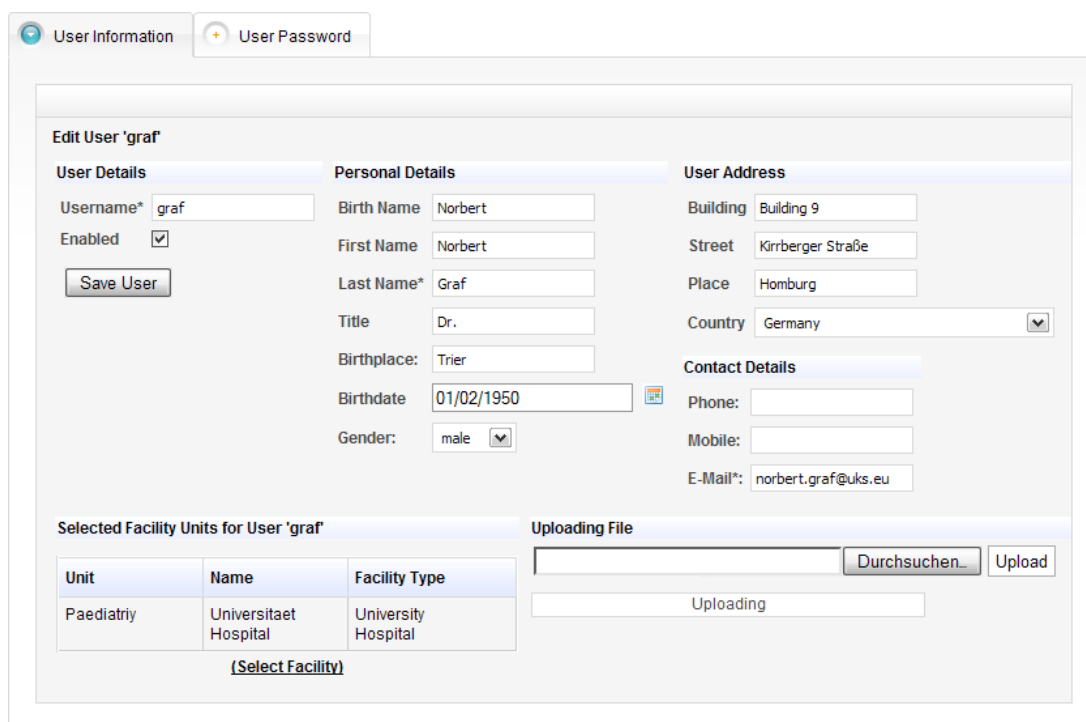Figure 14 – Personal information page

The same applies basically to the passwords of the ObTiMA users. They are even not only encrypted but "digested" meaning that the clear text of them cannot be reverted. (When authorizing a user to the system only the digests of the passwords are compared and never the actual passwords.) Also nowhere on the user interface (cf. Fig. 15) a password appears in clear text: an administrator can only autogenerate and (automatically) send via e-mail a new password to a user, the user him/herself then can change the password – thus only the user ever knows his or her own password. (As side note, the password must fulfil certain criteria towards e.g. minimal length or contained characters).

When the user logs into ObTiMA, then also digest authentication mechanism is employed, meaning that never the clear text password crosses the internet.
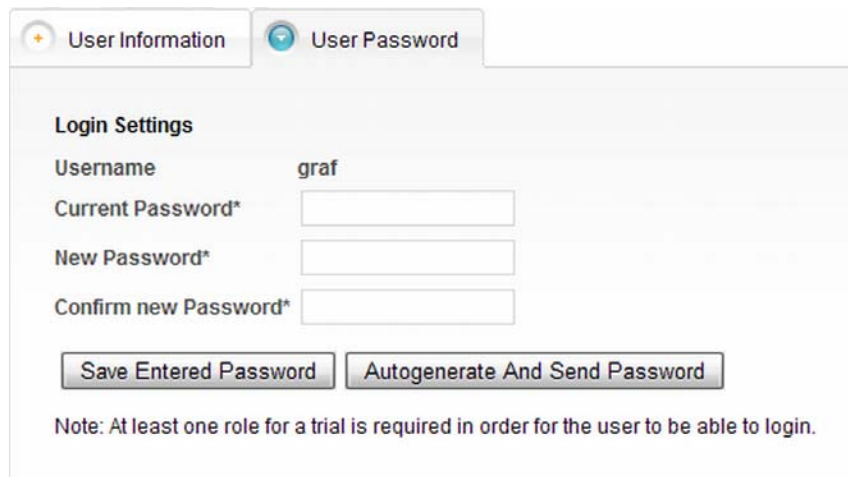
Figure 15 – Password setting/generation

Pseudonymization, Anonymization

Pseudonymization is the procedure in which the fields within a data record that could identify some person are replaced by some artificial identifier. There can be a single pseudonym for a collection of replaced fields or a pseudonym per replaced field. The overall purpose is to render the data record less identifying and therefore lower objections to their use and thus render it suitable for extensive analytics and processing.

The choice of which data fields are to be pseudonymized is partly subjective but should include all fields that are highly selective, such as a patient's name or address. Less selective fields, such as birth date are often also included because they are usually available from other sources and therefore make a record easier to identify. But pseudonymizing these less identifying fields removes most of their analytic value and should therefore be accompanied by introducing new derived and less identifying forms, such as year of birth.

The actual core research data (which are less identifying), such as clinical data, are usually not pseudonymised. It is important to realise that this is because too much statistical utility is lost in doing so, not because the data cannot be identified.

Protecting statistically useful pseudonymized data from re-identification requires a sound information security base as well as controlling the risk that the analysts, researchers or other data workers cause a privacy breach

The difference between pseudonymization and anonymization lies in the fact that pseudonym allows tracking back of data to its origins, which distinguishes pseudonymization from anonymization, where all person-related data that could allow backtracking has been

purged. For example, pseudonymization is an issue in patient-related data that has to be passed on securely between clinical centers in a clinical trial.

In a clinical trial, a trial subject is identified by a unique subject identification code which is assigned by the investigator to each trial subject to protect the subject's identity and used in lieu of the subject's name when the investigator reports adverse events and/or other trial related data.

All of the above points are fully accomplished by ObTiMA. If no special privileges arranged then only the persons having the appropriate rights can see all of a patient's master data. All other parties in the trial only see the pseudonym of that patient. The mentioned privileges can be given via informed consent which means that a patient has to specifically allow sharing the master data with other than the local physician. ObTiMA also contains a mechanism to enable such informed consent and "de-pseudonymization".

Data integrity checks

It is possible within ObTiMA to enable data integrity checks. This means that it is possible to define dependencies between entry fields for different values, e.g. XXX, or to give warning message if some values are not possible, e.g. a negative blood pressure value.

Hardware related requirements

The criteria that we have listed so far to be compliant to the GCP rules have concerned themselves only on the actual software side. But there are also some requirements that the hardware used must fulfil to be GCP compliant.

First of all, there must be a physical separation of database and application server onto distinct machines. This means that the actual data that is collected within a clinical trial is stored on one server and the application (logic) that processes this data is running on another computer machine. The communication between those machines must of course be fully and strongly encrypted to avoid some man-in-the-middle attacks.

The database that holds the patient master data and then again the actual research data can also be divided onto two machines. Thus, in addition to the encryption of the master data as explained above, an additional layer of security is introduced. The connection between the two databases happens only based on using the patient's pseudonym. For the application, the data is then combined "on-the-fly" on the application server to be presented and worked on by the user.

Software Validation Process

In each environment, software components from many sources may be used to create the application (e.g., in-house developed software, off-the-shelf software, contract software, shareware). In addition, software components come in many different forms (e.g., application software, operating systems, compilers, debuggers, configuration management tools, and many more). The validation of software in these environments can be a complex undertaking; therefore, it is appropriate that all of these software validation principles be considered when designing the software validation process. The resultant software validation process should be commensurate with the safety risk associated with the system, device, or process.

To have a proper validation process for the ObTiMA system, there actually need to be some requirements fulfilled first already during and even before the actual software development

process. Namely, a documented software requirements specification provides first the baseline for the development of the software and then also for both validation and verification. The software validation process thus cannot be completed without an established software requirements specification.

Software validation activities and tasks may be dispersed, occurring at different locations and being conducted by different organizations. However, regardless of the distribution of tasks, contractual relations, source of components, or the development environment, the device manufacturer or specification developer retains ultimate responsibility for ensuring that the software is validated.

To build a case that the software is validated requires time and effort. Preparation for software validation should begin early, i.e., during design and development planning and design input. The final conclusion that the software is validated should be based on evidence collected from planned efforts conducted throughout the software lifecycle.

Defect Prevention

Software quality assurance needs to focus on preventing the introduction of defects into the software development process and not on trying to "test quality into" the software code after it is written. Software testing is very limited in its ability to surface all latent defects in software code. For example, the complexity of most software prevents it from being exhaustively tested. Software testing is a necessary activity. However, in most cases software testing by itself is not sufficient to establish confidence that the software is fit for its intended use. In order to establish that confidence, software developers should use a mixture of methods and techniques to prevent software errors and to detect software errors that do occur. The "best mix" of methods depends on many factors including the development environment, application, size of project, language, and risk.

Software Validation after a Change

Due to the complexity of software, a seemingly small local change may have a significant global system impact. When any change (even a small change) is made to the software, the validation status of the software needs to be re-established. Whenever software is changed, a validation analysis should be conducted not just for validation of the individual change, but also to determine the extent and impact of that change on the entire software system. Based on this analysis, the software developer should then conduct an appropriate level of software regression testing to show that unchanged but vulnerable portions of the system have not been adversely affected. Design controls and appropriate regression testing provide the confidence that the software is validated after a software change.

5 Concrete steps for certification of ObTiMA

Within the ACGT project, ObTiMA has been developed basically as a proof-of-concept application to highlight the possibilities of ontology based creation and managing of clinical trials. As said in the preceding sections, the rules and guidelines to make ObTiMA compliant to the GCP have been closely followed.

Within the timeframe of the ACGT project we have not undertaken any concrete steps to apply for a formal certification of the software application. But since the development of ObTiMA is being continued in further projects funded by the European Commission, the application and fulfilment of all necessary certification steps together with an accredited certification organization is planned.

Important documents related to certification of ObTiMA

To use ObTiMA prospective clinical trials with real data GCP conformance is necessary and it needs to be certified that ObTiMA fulfils all these requirements. The following link (http://portal.scdm.org/?q=certification-resources) provides different documents as resources for certification.

The following link provides further information to make clinical software GCP compliant: http://www.scdm.org/gcdmp/.

One important document is deliverable D10 "GCP-compliant data management in multinational clinical trials" released from the EUROPEAN CLINICAL RESEARCH INFRASTRUCTURES NETWORK - TRANSNATIONAL WORKING GROUPS ECRIN-TWG funded in FP6-2005-Life Sciences and Health LSH-2005-3-4. The final version is dated as from the 15th of September 2008.

Besides Europe the US Department of Health and Human Services has published a document in 2007 called "Guidance for Industry. Computerized Systems used in Clinical Investigations" (http://www.fda.gov/OHRMS/DOCKETS/98fr/04d-0440-gdl0002.PDF). Further document are listed by the FDA (U.S. Food and Drug Administration) on: http://www.fda.gov/RegulatoryInformation/Guidances/UCM122046 and http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/default.htm.

The Society for Clinical Data Management (SCDM: http://www.scdm.org/gcdmp/) is a further source to help in simplifying the certification process of ObTiMA.

Appendix 1 - Abbreviations and acronyms

CDISC	Clinical Data Interchange Standards Consortium
GCP	Good Clinical Practice
ODM	Operational Data Model
TOB	Trial Outline Builder
URL	Uniform Resource Locator