Author's personal copy

COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE 114 (2014) 206-217





journal homepage: www.intl.elsevierhealth.com/journals/cmpb

Five-year experience with setup and implementation of an integrated database system for clinical documentation and research



Kerstin A. Kessel^{a,*}, Christian Bohn^b, Uwe Engelmann^b, Dieter Oetzel^a, Nina Bougatf^a, Rolf Bendl^c, Jürgen Debus^a, Stephanie E. Combs^{a,d}

- ^a Heidelberg University Hospital, Department of Radiation Oncology, Im Neuenheimer Feld 400, 69120 Heidelberg, Germany
- ^b CHILI GmbH, Friedrich-Ebert-Str. 2, 69221 Dossenheim, Germany
- ^c Heilbronn University, Department of Medical Informatics, Max-Planck-Str. 39, 74081 Heilbronn, Germany
- ^d Technical University of Munich (TUM), Department of Radiation Oncology, Ismaninger Straße 122, Munich, Germany

ARTICLE INFO

Article history: Received 8 October 2013 Received in revised form 30 January 2014 Accepted 6 February 2014

Keywords:
Documentation system
Data management
Clinical trials
Data capture
Information systems
Electronic analysis

ABSTRACT

In radiation oncology, where treatment concepts are elaborated in interdisciplinary collaborations, handling distributed, large heterogeneous amounts of data efficiently is very important, yet challenging, for an optimal treatment of the patient as well as for research itself. This becomes a strong focus, as we step into the era of modern personalized medicine, relying on various quantitative data information, thus involving the active contribution of multiple medical specialties. Hence, combining patient data from all involved information systems is inevitable for analyses. Therefore, we introduced a documentation and data management system integrated in the clinical environment for electronic data capture. We discuss our concept and five-year experience of a precise electronic documentation system, with special focus on the challenges we encountered. We specify how such a system can be designed and implemented to plan, tailor and conduct (multicenter) clinical trials, ultimately reaching the best clinical performance, and enhancing interdisciplinary and clinical research.

© 2014 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

In the age of intelligent information systems, data sharing in a medical environment remains a challenging objective [1]. The aim lies in ensuring that system architecture, communication protocols and usable procedures facilitate the interaction of data for any use, regardless of the point of origin of the information. This communication refers to the reuse

of data by other systems in the same department, or health-care networks (e.g. for telemedicine consultations and clinical referral), or collaborative research projects.

In the past, data from various information systems, which included both paperless and paper-based documentation, were used parallel within the clinical setting. Recently, information availability has become more elaborate and wide spread, and treatment decisions are based on a multitude of factors including imaging, molecular or pathological markers,

^{*} Corresponding author. Tel.: +49 6221 56 8202; fax: +49 6221 56 5353. E-mail address: Kerstin.Kessel@med.uni-heidelberg.de (K.A. Kessel).

surgical results and/or patient's preference. To avoid double documentation, loss or mix-up of data, and to provide a fast and reliable basis to collect all relevant data, interconnected information systems are developed. Relying on various quantitative data information becomes a strong focus, as disease management steps into the era of modern personalized medicine [2], thus involving the active contribution of multiple medical specialties. Gathering relevant data is therefore critical for reaching the best clinical performance, and enhancing interdisciplinary and clinical research – ultimately leading to optimizing treatment concepts, adjusting them, and developing new ones.

The achievement of building a new system uniting all these specifications is a challenging task from both a technical and non-technical point of view. Certainly, protecting patient privacy on all levels is of key importance and security mechanisms are required. Further, significant technical and architectural focus must lie on a vendor independent [3] and IHE (Integrating the Healthcare Enterprise) [4] complying concept with innovative methods and tools, thus, providing increasing flexibility and performance for the future.

Conducting clinical evaluations especially with large groups of patients is rather difficult in radiation oncology. Heterogeneous, distributed, voluminous amounts of data arise. This creates high complexity and involves considerable time and effort for analysis. As a result, it requires a precise methodical design to address these aspects in particular. The proposed method using an analysis system has the main goal to reduce time and effort for future clinical evaluations, ensuring high-quality data at the same time. Clinical experience during treatment will be transferred efficiently with the goal to improve therapy concepts. With advanced methods and tools complex problems can be processed in an understandable, transparent and reproducible way.

This interdisciplinary work contains two key aspects: a highly IT-focused setting and the application within an indepth medical context. First task is designing a detailed concept, which includes the implementation of an analysis system. This system must be realized and gradually integrated into the clinical routine. Further steps are the set-up of an adaptable analysis workflow to electronically assess clinical research questions. Eventually, several tools for analyzing radiotherapy treatments are connected. The validation of the concept of the analysis and documentation system is performed by conducting multiple clinical evaluations. Using the analysis system, it will be possible to get sophisticated evaluations faster and with less effort. New intelligent infrastructures and analysis procedures in medical diagnostics and therapy will be enabled by integrating the analysis system into the clinical routine. The resulting adjustment and optimization of therapy concepts will improve clinical practice and patient care and, ultimately, shape modern medicine.

2. Background

The intention of clinical trials is to test new and promising diagnostic and therapeutic methods with the objective to improve existing standard treatments and diagnostics. Scientific relevant evaluations are conducted retrospectively or with prospective clinical trials. With focus on current guidelines these new treatment concepts can identify prognostic and outcome-relevant parameters. Particularly in oncology, where treatment concepts are elaborated in interdisciplinary collaborations, handling large heterogeneous amounts of data efficiently is essential for an optimal treatment of the patient as well as for research itself – as Reboussin et al. [5] stated "good science requires good data".

The last decades showed enormous technical advances in radiation oncology, for instance introducing particle therapy with protons and carbon ions into clinical routine [6,7]. Especially radiation oncology is a highly image intensive medical specialty. Diagnostic and therapeutic data acquisitions are acquired throughout the course of treatment and during follow-up. Hence, not only heterogeneous and large amounts of data must be evaluated, it is also spread across various information systems within several involved departments in a large variety of documentation styles [8,9]. Involved systems are the Hospital-, Laboratory- and Oncology Information System (HIS, LIS, OIS), Picture Archiving and Communication System (PACS) and Record & Verify System. Researchers need assistance in reusing the terabytes of invaluable information collected routinely in all separate information systems. They hold treasures that are hidden in the deep [10].

Therefore, to date, retrospective clinical analyses, especially with large groups of patients, involve an immense time effort [11–13]. However, these evaluations provide important information about the efficacy of therapy, side-effect profiles or data for clinical quality assurance. Accordingly, these analyses are highly relevant [14].

Combining patient data from all involved systems is essential to prepare unstructured data for the analysis of retrospective and also prospective clinical trials. This demands special coordination in data management [11,15]. Such centralized repositories are non-existent in the medical enterprise [16,17]. Thus a documentation and data management system integrated in the clinical environment for electronic data capture needs to be introduced. This approach is a challenging task, but with researchers willingness for improvement, it offers many advantages regarding data collection, monitoring as well as analyzing and validation [11,18-20]. Key goal of the approach is to add an additional, built-in possibility to use the documentation system for immediate analysis of the collected data [21]. To establish such a documentation and analysis system, necessary workflows must be characterized and technical as well as clinical requirements regarding the subsystems must be defined.

The combination of medical image data with all other relevant documentation parameters or trial documentation, as well as the integrated analysis possibilities distinguishes the presented approach from other documentation systems and allows an improved outcome analysis for the future. The solution differs from other systems, which either only manage and organize patient treatment within a single department or other numerous strategies only focused on electronically documenting a single clinical trial (see Table 1). Our approach combines both: On the one hand, a common platform was created, that allows the coordination of clinical trials in radiation oncology even across departments and country borders. On the other hand, the system was linked

Author	Publication date	Institution	Approach summary
Edwards et al. [22]	2007	University of Utah, Salt Lake City, USA	Documentation system for a single, multicenter epidemiological trial in Alaska
Pavlovic et al. [23]	2007	University of Ljubljana, Ljubljana, Slovenia	Web-based documentation system for a single clinical trial about electrochemotherapy
Syed-Mohamad [24]	2008	Universiti Sains Malaysia, Kelantan, Malaysia	Web-based documentation system for a single clinical trial about children with malnutrition
Maass et al. [25]	2008	Satakunta Central Hospital, Pori, Finland	Web-based documentation and management system in s single department.
Angelow et al. [26]	2008	Institute for Community Medicine, Greifswald, Germany	Data management system for biosamples in a single three-center clinical trial
Buchsbaum et al. [27]	2009	Columbia University, NY, USA	Web-based documentation system for a single clinical trial about neuromuscular disorders
Heinemann et al. [28]	2009	University Hospital Freiburg, Freiburg, Germany	Documentation and patient management system in a single department in radiotherapy
Meyer et al. [29]	2010	Institute for Community Medicine, Greifswald, Germany	Data management system for epidemiology trials
Pozamantir et al. [30]	2010	Mount Sinai School of Medicine, NY, USA	Web-based documentation system for multicenter neuroscience research
Durkalski et al. [31]	2010	Medical University of South Carolina, Charleston, USA	Web-based documentation system for a single clinical trial (multicenter) about depression
Yamamoto et al. [32]	2012	Kyoto University Hospital, Kyoto, Japan	Documentation system for clinical trials and routine documentation

to the mandatory information systems to manage a complete treatment overview with more detailed information that might be relevant in retrospect.

In this paper we discuss our concept and five-year experience of a precise electronic data collection and analysis system in radiation oncology. We specify how such a system can be designed and implemented to plan, tailor and conduct (multicenter) clinical trials and evaluations, ultimately leading to the answer to clinically relevant questions via a defined and systematic process [14]. A special focus lie on the issues we encountered.

3. Design considerations

To guarantee the success of the methodology the following performance principles were included.

3.1. System objective

The need exist for an integrated electronic data collection and analysis system which allows simple database queries on structured data, and offers the ability to conduct extensive and complex analyzes. The system should support researcher in time-consuming evaluation tasks to get meaningful results faster and with less effort.

3.2. Data consolidation

Gathering relevant data for analyses is crucial. Information from different systems in the hospital must be combined to enhance interdisciplinary and clinical research. This wish always existed in our department and beyond, however, has never been realized before.

3.3. Clinical trial documentation

For data capture of clinical trials (retrospective and prospective) an infrastructure to organize and maintain each trial individually is of high importance. The ability to define and manage electronic CRFs is normally not available in hospital information systems. Therefore, a system must be installed which offers the possibility to cope with any clinical trial or research question.

3.4. Web-based approach

A web-based platform for joint clinical research is proposed. In addition the system is used as an international referring system for partners and referring physicians, as well as for transnational research collaborations. Especially in health-care, it is crucial to have all patient information on hand – even on mobile devices [33,34] – particularly in radiotherapy where imaging information always must be considered. However, because of strict medico-legal restrictions and lack of sufficient security and privacy mechanisms most medical applications have been restricted to in-house Intranets [35].

3.5. Communication standardization

Ensuring interoperability within a hospital itself and even in a wider collaborating health network, certain standards for export and import of data need to be applied and considered in the architecture and development phase of any health information system. In the field of radiation oncology being a multivendor system environment leading independent communication standards are DICOM, HL7 and IHE. Thus, information systems must be compliant with these technical specifications to provide flexibility and performance.

The goal is to exploit all of the latest technologies in the field regarding data management and image processing to form one single central system for clinical trials, which stores and exchanges, visualizes and analyzes any data on demand.

4. Description of methods and system

In the last decades, not only in the general sense we saw a transformation from paper-based to electronic documentation in the clinical routine, for instance introducing the electronic health record. Also clinical trial or rather research documentation supported by innovative soft- and hardware solutions evolves in the medical enterprise [9,32,33,36]. However, it is still not unusual for clinical documentation to be achieved with collections of paper-based case report forms (CRFs), excel sheets and local copies of medical images [18,37]. It is not necessary to explain the disadvantages of such unstructured and distributed documentation.

For the first time in our department we approached a novel strategy: a professional database supported documentation system for data capture, as well as for data merging from different sources. This potential way of addressing the discussed issues offers a better ability to distribute and access information throughout the department. Even beyond that, the platform is also created for multicenter research. Furthermore, the system is planned to be extendible by several analysis tools.

Clinical requirements as well as medical informatics aspects together with data structures and security requirements are considered (see Table 2). Data collection includes demographic patient information and medical history, therapeutic, biological and physical data including treatment planning, dose distribution, molecular and pathological specifications as well as outcome and follow-up information. This way, the system will offer the unique possibility to document specific clinical protocols or rather trials retrospectively as well as prospectively. In particular with regard to large-scale analysis large amounts of data will be analyzed. Specific requirements demand a flexible infrastructure, hence, not an off-the-shelf system was considered, but a system customizable as needed.

4.1. System design

After an intensive workflow analysis we previously published [38], we considered the emerging requirements (see Table 2) and designed an overall documentation and evaluation system offering the following characteristic

- Single central documentation system incorporated into the clinical environment with a web-based access, hence, platform independent
- 2. Full DICOM-RT support; DICOM-RT viewer
- Internal as well as external access possibility, meeting all necessary privacy and security guidelines as well as

Table 2 – Requirements for the central documentation system.				
Requirement	Details			
Reduction of time and effort for	Currently it takes considerable time and effort to work through patient records to gather all			
clinical trial analysis	necessary data			
Electronic support of the clinical evaluation process	Many research analysis tools already exist. These should be integrated in the new system			
Transparent research and	With a researcher leaving the department data collections should not get lost, hence, they will be			
reproducibility	reusable for similar purposes. Subsequent evaluations can build up on previous results and evaluations are comparable over a period of time			
Efficient data management	A new system should support an efficient way of data management through easy administration and maintenance, and support correct and complete data capture; duplicated information should be avoided. For multicenter evaluations new centers need to be integrated easily into the system			
Data ownership and export	Data should be accessible and useable only by authorized users; people are protective of their data, hard acquired information and results and, consequently, do not want to give away their control; export possibility of data for statistical analyses needs to be ensured			
Accessibility at all times	Data should be accessible wherever and whenever needed			
Central storage and access possibility	Data should be stored centrally to facilitate the collection of data especially if several researchers work together as a group; it should improve the current way of having to exchange continuously and merge Excel spreadsheets			
Use of computer resources of the department	Data should be accessible from every PC at the department; researchers may be working in their office or at a workstation in a lab			
Data security and confidentiality	A system accessible for many users in a clinical environment must allow for several levels of security and patient confidentiality; because of the high sensitivity of the data anonymization and pseudonymization need to be maintained as well as a secure data transfer			
Usability	Physicians need assistance in clinical trial management and documentation and a user-friendly application, not occupy oneself with technical details or learn special computer skills. Users should easily learn how to use the system with one training session			
Visualization of RT relevant data	Diagnostic and therapeutic imaging should be connected and viewable within the new system			
Regular and automatic backup mechanism of all data	Even with a careful handling of personal storage devices, data acquisitions get lost or overridden for different reasons; backup and recovery features are needed to protect against data loss			
Automatic import of data	Data from other systems should be imported automatically or accessible from within the new system whenever possible			

- pseudonymization and anonymization, allowing for multicenter research
- Professional database and data management system, for efficient administration and maintenance, handling a large variety of voluminous datasets from various information systems
- 5. In the long-term, framework for unlimited clinical trial management
- Connection of existing clinical information systems; data exchange over standard interfaces via DICOM and HL7
- 7. Integration of analysis tools for evaluation.

4.2. System architecture and integration

An open source relational database based on the data management system PostgreSQL builds the basis. It comes with three additional components: a PACS system, a DICOM-RT viewer and a web-based telemedicine record functioning as a clinical trial documentation user interface.

The CHILI GmbH, with whom we maintain a strong cooperation, established this flexible data model and robust infrastructure. It is consistently refined and adapted to our needs for clinical trial management and documentation. With our approach any type of data, especially DICOM-RT, can be stored, processed, exchanged and visualized. We published previously, see Kessel et al. [21], a detailed overview on the system design and architecture and characterize the main features as follows.

Several interfaces for DICOM and HL7 are implemented and allow for interoperability within the clinical environment [21]. Connected systems are the Hospital Information System (HIS) or rather the electronic patient record (eHR), three Picture Archiving and Communication Systems (PACS) and the Oncology Information System (OIS). A professional DICOM-RT viewer (Class IIb; according to the European Medical Devices Directive) visualizes not only common DICOM data, but also DICOM-RT plans for reviewing irradiation. It displays dose distributions and calculates dose volume histograms on demand, and gives an overview on dose statistics and plan parameters right within the documentation system. The underlying components are compliant with the IHE framework and have been tested at five European IHE Connectathons. The data model itself can be dynamically extended with additional data structures instantaneously.

Patient demographic data are imported and updated from the HIS via HL7-ADT messages. The HL7 event types A01, A02, A04 (patient visits) create a new patient in the database; an A08, A31 (patient update) represent changes in the patient demographic data, which are automatically updated in the system. In addition, patients are merged because of an A34 or A40 message. The OIS supplies radiation information via HL7-DFT messages. The HL7-mapping, automatically updates fields in the treatment overview of each patient. This way, the first and last day of irradiation, radiation technique, and ICD-10 diagnosis code are automatically imported without manual data entering. A performance code of each message indicates the radiation method (for example tomotherapy, IMRT, brachytherapy). Further, a Query/Retrieve (Q/R) mechanism is triggered with each DFT message to import all corresponding DICOM files (including DICOM-RT) from each connected

PACS and map it to the corresponding patient as soon as it is available. This is processed at least two times, at the beginning of irradiation and updated again at the end of treatment, because DICOM objects such as verification images are created not until treatment delivery.

The web-based graphical user interface allows for a platform independent usage. Data import can be done manually in prepared entry modules, or using the mentioned DICOM and HL7 interfaces, or as file attachments. These are uploaded using a Java applet functioning in any Internet browser, which is able to read and write DICOM-CDs, receive and send data with DICOM C-Store, and moreover anonymize DICOM during import.

Export of imaging data from one PACS and importing it into another is not a comfortable way, as these days telemedicine features exist to support transfer and sharing. DICOM-Email [39] is one option we realized with one external partner.

It is crucial to offer several privacy and security mechanisms to meet all requirements in a health network. To assure that, we use an encrypted https data transfer for exchanging data between the central documentation system and any other system. For the external access, we installed an SSL Server certificate and generate individual SSL client certificates to provide more host-to-host security. Additionally, an intermediate application gateway in the demilitarized zone (DMZ) and several firewalls allow for the optimum data protection. Certainly, a user needs an account and password to access the system. Further, a roles-and-rights concept defines in very detail user access rights for any client or trial respectively. User authorization can change dynamically and vary depending on the involvement of a user, for example being the leading investigator for a trial, or being the main person documenting his/her own evaluation conducted within the system. Further, as a final technical privacy feature, pseudonymization of patient demographic data is done by a PID-Generator (patient identifier) by the TMF (Technology, Methods, and Infrastructure for Networked Medical Research) [40].

Export of data is possible in CSV (comma-separated values) format. This is supported by a web-based query builder and query management functionality fully integrated in the documentation system. These queries for up-to-date reports can be saved and executed any time.

4.3. Clinical trial management

Data collection includes demographic patient information and medical history, therapeutic, biological and physical data including treatment planning, dose distribution, molecular and pathological specifications as well as outcome and follow-up information. This way, the system offers the unique possibility to document specific clinical protocols or rather studies retrospectively as well as prospectively. In particular with regard to large scale analysis large amounts of data will be analyzed.

Each client, or rather clinical trial, is stored in a separate database, thus having different data maintenance and user management possibilities. Creating and configuring a new client is achieved via the web-based administration system. A new client can be derived from an existing client, which is particularly useful when trials are similar. In the clinical

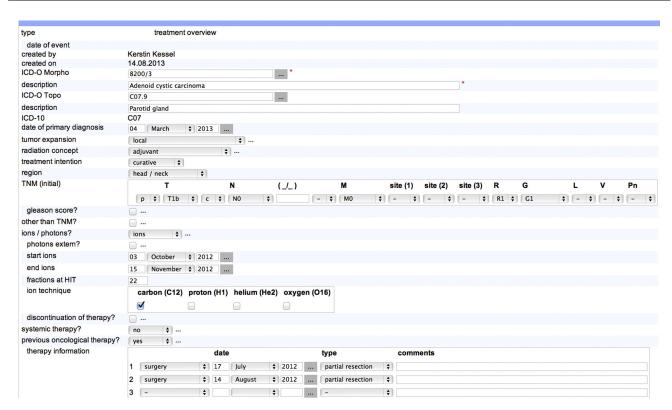


Fig. 1 - Entry form "treatment overview" with essential information of treatment and previous therapies.

workflow, patients can be moved from one client/trial to another, as at the beginning of the documentation it is often unclear if the patient is going to enroll in a clinical trial. However, moving patients is only possible as long as all clients have the same data structure, in this case having the same patient demographic data.

Wanting basic documentation for all patients has led to developing three different kinds of entry modules, which can also be referred to as eCRF (electronic case report forms). Documentation modules are either used for all patients, or unique for individual evaluations. We designed and set up basic entry modules for overall documentation – the minimal information for each patient entered in the database. This even allows defining database queries, which address data from all trials in the system. Additionally, specific modules for each trial are separately designed and generated, regardless if it is for a current prospectively conducted trial of the department to document parameters required by the trial protocol or a retrospective evaluation, for example done by a PhD student.

The system has an advanced admin web-front-end to manage the whole documentation system, hence all clinical trials. This way nearly all configurations can be done in a comfortable way and not on a server level, which, however, is sometimes necessary or faster in some cases, for instance changing detailed HL7-mapping configurations. Each client/trial has its own database. The entry modules are written in XML language specifying each data field regarding data type, range and dependencies. Configurations are interpreted by the server, thus creating several tables in the database. Dependencies are for example that some modules are only selectable once for each patient (e.g. screening form), or a data

field is only visible depending on other fields (e.g. Gleason score for prostate cancer only in male patients), or some fields are only editable and visible for certain user groups.

We embedded three standardized international medical dictionaries for tumor documentation. An especially implemented search module for ICD-O (International Classification of Diseases for Oncology, third revision) encoding helps to save oncological diagnoses in a standard way. It is based on Ajax technology. This is used in the basic documentation "treatment overview" for each patient in the system, shown in Fig. 1. Among others, diagnosis, radiation dates and previous oncology therapies are entered. The TNM classification (tumor (lymph) nodes metastasis) and WHO grade (world health organization) encode cancer stages. Toxicity and side effects, for example during treatment, are coded according to CTCAE (Common Terminology Criteria for Adverse Events, version 4.03) guidelines. We predefine a side effects profile to be filled out depending on the tumor region, see Fig. 2. This has been established in cooperation with the professionals and adapted to radio-oncology typical needs.

5. Integration process and current status

Many others have already said there is no "one-size-fits-all" solution for web-based documentation of clinical trials or patient data per se [5]. Or as Kristianson et al. [41] stated "each clinic is a 'kingdom' of its own" acting in its peculiar independent way. The large number of requirements and circumstances for a documentation system demand an individual approach.

type	clinical examination
date of event	05 June
created by	Kerstin Kessel
created on	25.06.2013
type of visit	follow-up 💠 *
imaging modality	MRT • *
date of imaging MRT	04 June \$ 2013
max. tumor diameter [mm]	23
response	SD 🛊
karnofsky performance score [%]	80 🛊
extended examination?	☑
region	brain / base of skull \$
pain	CTCAE characteristic
	1 \$
skin / subcutaneous tissue disorder	
	1 🛊
nausea	
	0 \$
vomiting	0 \$
fatigue	
	1 🕏
lymphedema	0 \$
motor neuropathy	0 \$
sensory neuropathy	
	0 \$
cognitive disturbance	0 \$
alopecia	
dioposia	2 \$
mucositis oral	
	- •
ear and labyrinth disorder	- \$
trismus	
uisiilus	- -

Fig. 2 – Entry form for follow-up documentation showing a part of the toxicity/side effect profile for brain/base of skull tumors.

We started in 2009 with the idea of initiating a documentation system in the department and clinical routine. Nearly five years later we conducted 37 analyses: 15 retrospective evaluations [42–46,47] and 22 clinical prospective studies [48–51], most of them are still ongoing with recruitment or follow-up. One of the retrospective analyses is a multicenter glioblastoma trial together with eight national clinical centers. Moreover, for the referral of (international) patients, referral clients were prepared for external partners.

The concept and workflow analysis phase, assessing requirements as well as choosing, purchasing and launching a basic web-based documentation database took one and a half years until May 2010. In the following year, we connected the mandatory information systems of the hospital, and set up the necessary communication interfaces for DICOM and HL7. Implementation of the HL7 interface took nine months until February 2011. The challenging part for connecting the existing PACS was to get permissions from the local data protection supervisor; therefore it lasted until April 2012 for all PACS systems to be completely incorporated.

The design of a detailed rules-and-rights concept was followed by the first launch of the first clinical trial in May 2011. Even though the DICOM connection was not yet finished we

started with our first clinical trial documentation. The strategy was to test as we go along. Initial concepts and design on how certain entry modules are configured and what needs to be documented developed with experience over time.

The development of the DICOM-RT viewer (see Fig. 3) took about two years and was implemented in strong cooperation with the CHILI GmbH. It is an additional component, as the basic DICOM viewer existed already. Since April 2012 it is fully incorporated, and especially radiation plans can be reviewed at any computer in the hospital without having to access a clinical PACS application or TPS (treatment planning system) client workstation.

The external access was realized very fast and is ready for use since January 2012. However, we needed more time to get permission for an SSL Server Certificate from the authorities. During that time we implemented the pseudonymization feature. In early 2013, we started our first multicenter clinical trial with eight cooperating centers.

As of October 2013, about 7500 patients are in the system and daily new ones are added. Since 2013 we document all patients treated in the department, even though they are not enrolled in a clinical trial (see Fig. 4). After performing several evaluations, it became apparent that our documentation

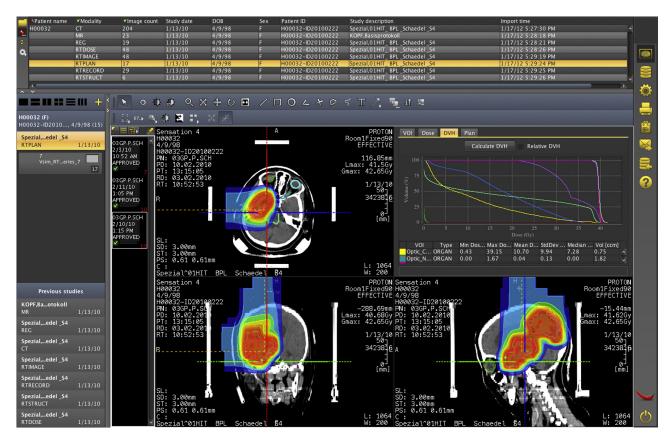


Fig. 3 – DICOM-RT viewer showing a radiation plan; all approved radiation plans from one RT-Plan series are visualized in the light box on the left side; displayed are the transversal (top left), coronal (bottom left) and sagittal (bottom right) view; DVH and statistics tab is shown in the top right.

system is a helpful tool in clinical research. It also serves as an archive especially for large-scale trials possibly conducted in the future.

The time it takes for an experienced computer specialist, familiar with the system and database platform, to allocate a ready-to-enter trial infrastructure for web-based data capture, including all entry modules and accounts, is about two days. However, this implies that the trial design is already made, for instance as paper based CRF.

At the moment, entered data can be exported in CSV format via SQL queries, which vary in their complexity. Compared

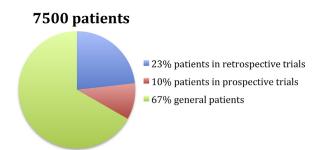


Fig. 4 – Distribution of all patients in the database according to trial participation. However, not all clinical trials are yet documented in the database; the actual percentage for patient in prospective trials is higher.

to manual data collection and preparing data for statistical analysis and publication, which takes several weeks, executing a complex SQL query takes only several minutes. It contains various joins of database tables and complex calculations. However, both depend highly on the number of patients included. Above all, a complete data export requires all data to be entered into the database prospectively and continuously.

Particularly, in use during prospective documentation the query functionality is heavily used to monitor upcoming patient visits and to ensure complete and correct documentation, as the reports are produced on live data. For instance, case managers start their day with getting the current list of patients who will finish their irradiation, to enter missing key values about treatment; or radiographers are interested in the weekly list of patients who request special care, for example children who need anesthesia or risk patients with MRSA or pacemaker.

6. Experience within the clinical routine

Different trials have different requirements. Since we began this project we tested and evaluated our system in several scenarios. We analyzed prospectively patients grouped by disease and tumor location [51,48], and performed large-scale analyses (for example of patients with pancreatic cancer [43,44,42]). One of these includes extended and integrated analysis tools for image registration and dose calculation. Compared to paper-based CRFs, data is right "there" and available for overview, analysis and quality checks during trial process.

We started the first user test documentation with one case manager and soon afterwards with two trained study nurses. They immediately reported issues and suggested improvements for the documentation process. The positive response led us to recruit documentation specialists and include all radiographers and case mangers for a continuous documentation. With the beginning of the first clinical trials resident and attending physicians started documenting CRFs and clinical examinations into the system.

6.1. Multicenter application

Our first multicenter clinical trial is currently under way. It would not even be feasible solely on a paper-based level, because further information is necessary for significant analyses, such as radiation plans, CTs, MRI. Particularly is this context, the fully integrated DICOM-RT viewer with its dose viewing tools (see Fig. 3) is a highly used feature for reviewing any external radiation plan during evaluation process. Data capture runs as efficiently as with the single-center trials.

6.2. Scientific use

Clinical evaluations conducted and based on the documentation system are comfortable to perform. Any researcher can query the database anytime, assuming having the user right. However, a computer scientist must create advanced SQL queries, which require more IT knowledge. After complete data collection, generating and tailoring such SQL queries to extract data for further advanced statistical analysis (e.g. using SPSS) takes about 6 h to 2 days, but depends especially on how familiar one is with the underlying data structure. Most of the time is used to get accurate specifications from the physician/researcher. Normally, their first request is too vaguely phrased. Redefining of queries is time consuming, besides the time loss for previous versions that prove to be incomplete. The understanding of how detailed a computer scientist must now what exactly the physician/researcher wants as a result, can be hard to teach.

6.3. Productivity

For example, in our analysis of 260 patients with brain tumors and tumors of the skull base [51], we built an SQL query in about two days. This time was necessary because 20 database table joins needed to be configured. The final query consists of 90,000 characters. Involved personnel gathered all data prospectively. Executing the query with calculations, such as progression free survival, overall survival and side effects group for acute and late appearing, takes almost 20 min. According to our previous experience and comparison with standard way of conducting clinical evaluations [38], data collection and preparation for these patients would have been at least 1.5 h per patient. Of course work time depends highly on the necessary data for analysis and the experience of the

researcher. However, it results in more than eight weeks of constant work, assuming one single person would analyze these patients full-time (9 h a day = 6 patients a day) in a retrospective research project.

Analyses are only as good as the data captured in the database. Therefore, the key goal is to ensure high data quality. We developed several features, such as edit checks or the direct access of further clinical information systems, to support the work for documentation specialists as mentioned previously [21,43]. However, subsequent difficulties and our experience are discussed in detail in the following.

7. Lessons learned

We have established a central information system collecting all relevant clinical data including imaging, pathology reports, treatment reports and radiation oncology planning information, as well as detailed information on follow-up. Having a central documentation system provides many advantages as mentioned previously. However, the difficulties of transitioning from paper-based to electronic data capture should not be underrated. There are major challenges that must be overcome in a clinical enterprise, some of which only developed during the process of development through experience with use and application of the new system.

7.1. System maintenance

If the system is used in a dynamic environment with expanding study concepts, various research questions as well as researchers and caregivers are involved. Clinical trials constantly vary in size and data fields. Besides, no study runs exactly as initially planned. Therefore, a computer specialist focused solely on ongoing maintenance and customization of the database is required. Being a complex and dynamic process, it becomes its own profession to manage and coordinate clinical trials [11] having not only strong communication skills, but also multiple abilities such as knowledge about clinical processes, workflows, underlying infrastructures and clinical trials, as well as a strong scientific interest and IT background. This includes also administration of service contracts with the company of origin and continuous system/feature upgrades over time.

7.2. Data protection

One major problem we did not anticipate as being the most difficult one was to get permissions from the data protection supervisor. Safety issues regarding clinical data are of utmost priority within a hospital. Therefore, all systems collecting and/or connecting clinical data must be approved on a medico-legal level.

Our system fulfills all official requirements and is verified by official institutions as a medical product. In the beginning, however, despite this characteristic of the system, connection with existing clinical databases was performed only hesitantly because of the fear that data protection might be disrupted. Several quality-checks as well as internal approval mechanisms must be passed positively before connection of the system was performed, which should not be underestimated for future applications. Thus, in spite of the given medicolegal approval of the database backbone, individual security mechanisms of the institutions need to be followed and respected.

7.3. Multicenter application

All aspects described come even more into focus for multicenter application. Anonymization and pseudonymization of data is necessary, data safety must be explained and followed in detail, and web-access input for data should be easy to use for the documenting person. The set up of for example an external DICOM-Email communication can be a huge effort, not from a technical point of view, but rather incorporating a new procedure into a current workflow and determining how, when and who is responsible to process it.

Detailed instruction manuals and strong interpersonal communication, motivation and organizational skills, even more than in trials for single-center, are necessary to make use of the multicenter approach of such a system. Again, online tools or instructions cannot replace a contact person, since some problems at or from the distant centers are only be solved with individual procedures.

7.4. Documentation

As soon as the system is up and running, life comes to the database only with sound and solid data, which are entered on a regular basis. To secure correct data input, an experienced and qualified person is needed. Thus, investment cost does not only include hard- and software, as well as computer scientists to set-up the database, but also continuous documentation by responsible clinical documentation specialists, case managers, radiographer. We found that the best solution is to share the workload between clinical staff and document as a team. Only this way, a continuous documentation can be financed, at least in our department. However, investing in good documentation in a clinical environment will benefit all.

7.5. Acceptance

At the beginning, our physicians were not enthusiastic about the introduction of the new documentation system, because their role and perspective changed from being the active and independent data manager to the dependent end user. Furthermore, the usual format and data presentation changed from what physicians were familiar with. They preferred their old way documenting patients in Excel spreadsheets, especially as they are now reliant on an informatics specialist to set up the necessary data structures in the documentation system before they can start. First, this responsibility shift seemed to be too complicated. However, when physicians and researchers got to know the system and realized the full potential we could gain their participation. Human issues can slow down the undertaking of such a system [52], yet success comes only with commitment between all players. This demands consideration of all individual researchers' needs and respect of their data ownership.

7.6. Scientific use and output of data

The real value lies beyond implementation and input of data. Analysis and evaluation of data is the goal and scientific rationale for such a system, providing productivity gain by needing less time and effort. However, each query or research question is based on individual data fields, connection of information from various subareas (such as imaging, follow-up, treatment plan). To furnish a data export including detailed data regarding different research questions is not a simple task and exceeds the informatics knowledge of most physicians and researchers. Again, a specialized computer specialist to generate queries and output data collections is needed to exploit the full scientific value of the system.

In our database approach we learned all these lessons either beforehand, or during the generation process. All issues could be solved, however, continuous effort on a technical as well as personal basis is required to include a new and elaborate system within an existing infrastructure. For multicenter use of such a common database, the foundation must be strong interpersonal communication between centers. This also requires convincing other centers regarding data protection and safety, as well as ease to use the web-access applications.

With accepting the change how to enter and retrieve clinical data for research, and understanding advantages and limitations, users will join the new direction of clinical data management [25].

8. Future plans

A platform for joint clinical research was created. The first multicenter trial has already been initiated. In addition the system will be used as an international referring system for partners and external physicians, as well as for transnational research collaborations. Intra- and interdepartmental comparisons of treatment concepts are planned as well as treatment techniques across specialized institutions.

The Laboratory Information System (LIS) is the next system we plan to connect. It sends two types of HL7-ORU messages with lab results to the hospital network, with either single-value or textual findings. We want to display these results and map them to the patient, as we did with the HL7-DFT messages. However, the number of laboratory values can vary per finding, hence, a dynamic data infrastructure is necessary.

More automatic functionality is a major task for the future. Among others, we want to add an automatic import of radiation plan parameters and statistics to the database, thus arithmetic calculations can be easily done. All DICOM files are already in the system and can be visualized with the handy DICOM viewer, but we need a way to extract the parameters immediately for analysis. For now, we use a manual tool to calculate the values and write them back into the database (see [43]). The step for doing that automatically needs still to be accomplished. Connection of further analysis tools holds immense benefits that for now are still a vision.

Conflict of interest statement

There is no conflict of interest to report for this article.

Acknowledgments

The project was co-funded by the European Commission within the Framework Program 7 Capacities Specific Program, under Grant Agreement Number 228436. We gratefully thank Eva Meyerhof for her diligent work and documentation for two years.

REFERENCES

- H.K. Huang, PACS and Imaging Informatics, John Wiley & Sons, Hoboken, New Jersey, 2010.
- [2] M.L. Maitland, R.L. Schilsky, Clinical trials in the era of personalized oncology, CA Cancer J. Clin. 61 (2011) 365–381.
- [3] C.H. Elsner, M. Egbring, H. Kottkamp, T. Berger, S. Zoller, G. Hindricks, Open source or commercial products for electronic data capture in clinical trials? A scorecard comparison, Comput. Cardiol. 30 (2003) 371–373.
- [4] Integrating the Healthcare Enterprise, (IHE), http://www.ihe.net (accessed 06.09.13) (n.d.).
- [5] D. Reboussin, M.A. Espeland, The science of web-based clinical trial management, Clin. Trials 2 (2005) 1–2.
- [6] S.E. Combs, O. Jäkel, T. Haberer, J. Debus, Particle therapy at the Heidelberg Ion Therapy Center (HIT) – integrated research-driven university-hospital-based radiation oncology service in Heidelberg, Germany, Radiother. Oncol. 95 (2010) 41–44.
- [7] J.S. Welsh, Basics of particle therapy, Am. J. Clin. Oncol. 31 (2008) 493–495.
- [8] M.Y.Y. Law, B. Liu, L.W. Chan, Informatics in radiology: DICOM-RT-based electronic patient record information system for radiation therapy, Radiographics 29 (2009) 961–972.
- [9] A. Mandal, A.K. Asthana, L.M. Aggarwal, Development of an electronic radiation oncology patient information management system, J. Cancer Res. Ther. 4 (2008) 178–185.
- [10] H.U. Prokosch, T. Ganslandt, Perspectives for medical informatics. Reusing the electronic medical record for clinical research, Methods Inf. Med. 48 (2009) 38–44.
- [11] Z. Lu, J. Su, Clinical data management: current status, challenges, and future directions from industry perspectives, Open Access J. Clin. Trials 2 (2010) 93–105.
- [12] E.L. Eisenstein, R. Collins, B.S. Cracknell, O. Podesta, E.D. Reid, P. Sandercock, et al., Sensible approaches for reducing clinical trial costs, Clin. Trials 5 (2008) 75–84.
- [13] B. Ene-Iordache, S. Carminati, L. Antiga, N. Rubis, P. Ruggenenti, G. Remuzzi, et al., Developing regulatory-compliant electronic case report forms for clinical trials: experience with the demand trial, J. Am. Med. Inform. Assoc. 16 (2009) 404–408.
- [14] M.A. Ebert, J. Blight, S. Price, A. Haworth, C. Hamilton, D. Cornes, et al., Multicentre analysis of treatment planning information: technical requirements, possible applications and a proposal, Australas. Radiol. 48 (2004) 347–352.
- [15] V. Steil, F. Röhner, F. Schneider, F. Wenz, F. Lohr, G. Weisser, Current requirements for image management in radiotherapy, Strahlenther. Onkol. 188 (2012) 499–506.
- [16] B.S. Elger, J. Iavindrasana, L. Lo Iacono, H. Müller, N. Roduit, P. Summers, et al., Strategies for health data exchange for

- secondary, cross-institutional clinical research, Comput. Methods Programs Biomed. 99 (2010) 230–251.
- [17] J.A. Hernandez, C.J. Acuna, M.V. de Castro, E. Marcos, M. Lopez, N. Malpica, Web-PACS for multicenter clinical trials, IEEE Trans. Inform. Technol. Biomed. 11 (2007) 87–93.
- [18] I. Pavlovic, T. Kern, D. Miklavcic, Comparison of paper-based and electronic data collection process in clinical trials: costs simulation study, Contemp. Clin. Trials 30 (2009) 300–316.
- [19] K. El Emam, E. Jonker, M. Sampson, K. Krleza-Jerić, A. Neisa, The use of electronic data capture tools in clinical trials: web-survey of 259 Canadian trials, J. Med. Internet Res. 11 (2009) e8.
- [20] O. Ratib, M. Swiernik, J.M. McCoy, From PACS to integrated EMR, Comput. Med. Imaging Graph. 27 (2003) 207–215.
- [21] K.A. Kessel, N. Bougatf, C. Bohn, D. Habermehl, D. Oetzel, R. Bendl, et al., Connection of European particle therapy centers and generation of a common particle database system within the European ULICE-framework, Radiat. Oncol. 7 (2012) 115.
- [22] R.L. Edwards, S.L. Edwards, J. Bryner, K. Cunningham, A. Rogers, M.L. Slattery, A computer-assisted data collection system for use in a multicenter study of American Indians and Alaska Natives: SCAPES, Comput. Methods Programs Biomed. 90 (2008) 38–55.
- [23] I. Pavlovic, D. Miklavcic, Web-based electronic data collection system to support electrochemotherapy clinical trial, IEEE Trans. Inform. Technol. Biomed. 11 (2007) 222–230.
- [24] S.-M. Syed-Mohamad, Development and implementation of a web-based system to study children with malnutrition, Comput. Methods Programs Biomed. 93 (2008) 83–92.
- [25] M.C.M. Maass, P.P. Asikainen, T.T. Mäenpää, O.O. Wanne, T.T. Suominen, Usefulness of a Regional Health Care Information System in primary care: a case study, Comput. Methods Programs Biomed. 91 (2008) 175–181.
- [26] A. Angelow, M. Schmidt, K. Weitmann, S. Schwedler, H. Vogt, C. Havemann, et al., Methods and implementation of a central biosample and data management in a three-centre clinical study, Comput. Methods Programs Biomed. 91 (2008) 82–90.
- [27] R. Buchsbaum, P. Kaufmann, A.I. Barsdorf, R. Arbing, J. Montes, J.L.P. Thompson, Web-based data management for a phase II clinical trial in ALS, Amyotroph. Lateral Scler. 10 (2009) 374–377.
- [28] F. Heinemann, F. Roehner, M. Schmucker, G. Bruggmoser, K. Henne, A.-L. Grosu, et al., Department and patient management in radiotherapy. The Freiburg model, Strahlenther. Onkol. 185 (2009) 143–154.
- [29] J. Meyer, S. Ostrzinski, D. Fredrich, C. Havemann, J. Krafczyk, W. Hoffmann, Efficient data management in a large-scale epidemiology research project, Comput. Methods Programs Biomed. 107 (2012) 425–435.
- [30] A. Pozamantir, H. Lee, J. Chapman, I. Prohovnik, Web-based multi-center data management system for clinical neuroscience research, J. Med. Syst. 34 (2010) 25–33.
- [31] V. Durkalski, Wenle Zhao, C. Dillon, J. Kim, A web-based clinical trial management system for a sham-controlled multicenter clinical trial in depression, Clin. Trials 7 (2010) 174–182.
- [32] K. Yamamoto, K. Yamanaka, E. Hatano, E. Sumi, T. Ishii, K. Taura, et al., An eClinical trial system for cancer that integrates with clinical pathways and electronic medical records, Clin. Trials 0 (2012) 1–10.
- [33] M. Modahl, Tablets Set to Change Medical Practice, QuantiaMD, 2011, pp. 1–10.
- [34] K. Kiser, 25 ways to use your smartphone. Physicians share their favorite uses and apps, Minn. Med. 94 (2011) 22–29.
- [35] C. Costa, J.L. Oliveira, A. Silva, V.G. Ribeiro, J. Ribeiro, Design, development, exploitation and assessment of a Cardiology

- Web PACS, Comput. Methods Programs Biomed. 93 (2009) 273–282
- [36] Z. Lu, Electronic data-capturing technology for clinical trials: experience with a global postmarketing study, IEEE Eng. Med. Biol. Mag. 29 (2010) 95–102.
- [37] N.R. Anderson, E.S. Lee, J.S. Brockenbrough, M.E. Minie, S. Fuller, J.F. Brinkley, et al., Issues in biomedical research data management and analysis: needs and barriers, J. Am. Med. Inform. Assoc. 14 (2007) 478–488.
- [38] K.A. Kessel, D. Habermehl, C. Bohn, A. Jäger, R.O. Floca, L. Zhang, et al., Database supported electronic retrospective analyses in radiation oncology: establishing a workflow using the example of pancreatic cancer, Strahlenther. Onkol. 188 (2012) 1119–1124.
- [39] G. Weisser, U. Engelmann, S. Ruggiero, A. Runa, A. Schröter, S. Baur, et al., Teleradiology applications with DICOM-e-mail, Eur. Radiol. 17 (2007) 1331–1340.
- [40] K. Pommerening, M. Reng, P. Debold, S.C. Semler, Pseudonymisierung in der medizinischen Forschung – das generische TMF-Datenschutzkonzept, GMS Med. Inform. Biom. Epidemiol. 1 (2005), Doc17.
- [41] K.J. Kristianson, H. Ljunggren, L.L. Gustafsson, Data extraction from a semi-structured electronic medical record system for outpatients: a model to facilitate the access and use of data for quality control and research, Health Inform. J. 15 (2009) 305–319.
- [42] S.E. Combs, D. Habermehl, K.A. Kessel, F. Bergmann, J. Werner, I.C. Brecht, et al., Intensity modulated radiotherapy as neoadjuvant chemoradiation for the treatment of patients with locally advanced pancreatic cancer: outcome analysis and comparison with a 3D-treated patient cohort, Strahlenther. Onkol. 189 (2013) 738–744.
- [43] K.A. Kessel, D. Habermehl, A. Jäger, R.O. Floca, L. Zhang, R. Bendl, et al., Development and validation of automatic tools for interactive recurrence analysis in radiation therapy: optimization of treatment algorithms for locally advanced pancreatic cancer, Radiat. Oncol. 8 (2013) 138.
- [44] D. Habermehl, K.A. Kessel, T. Welzel, H. Hof, A. Abdollahi, F. Bergmann, et al., Neoadjuvant chemoradiation with

- gemcitabine for locally advanced pancreatic cancer, Radiat. Oncol. 7 (2012) 28.
- [45] S.E. Combs, N. Schall, K.A. Kessel, J. Rieber, M. Bischof, W. Behnisch, et al., Outcome after radiation therapy in pediatric and adult patients with rhabdomyosarcoma treated with modern radiation therapy, in: Proc. DEGRO, Berlin, 2013.
- [46] K.A. Kessel, A. Jäger, R.O. Floca, D. Habermehl, R. Bendl, J. Debus, et al., Image guided and adaptive radiotherapy (IGRT and ART) for locally advanced pancreatic cancer considerations for dose adaption to changes in GTV volume and organ motion, in: Proc. DEGRO, Berlin, 2013.
- [47] D. Habermehl, I.C. Brecht, F. Bergmann, T. Welzel, S. Rieken, J. Werner, et al., Chemoradiation in patients with isolated recurrent pancreatic cancer – therapeutical efficacy and probability of re-resection, Radiat. Oncol. 8 (2013) 27.
- [48] S.E. Combs, T. Welzel, D. Habermehl, S. Rieken, J.-O. Dittmar, K.A. Kessel, et al., Prospective evaluation of early treatment outcome in patients with meningiomas treated with particle therapy based on target volume definition with MRI and 68Ga-DOTATOG-PET, Acta Oncol. 52 (2013) 514–520.
- [49] S.E. Combs, K.A. Kessel, K. Herfarth, A. Jensen, S. Oertel, C. Blattmann, et al., Treatment of pediatric patients and young adults with particle therapy at the Heidelberg Ion Therapy Center (HIT): establishment of workflow and initial clinical data, Radiat. Oncol. 7 (2012) 170.
- [50] H. Hauswald, S. Rieken, S. Ecker, K.A. Kessel, K. Herfarth, J. Debus, et al., First experiences in treatment of low-grade glioma grade I and II with proton therapy, Radiat. Oncol. 7 (2012) 189.
- [51] S.E. Combs, K.A. Kessel, D. Habermehl, T. Haberer, O. Jäkel, J. Debus, Proton and carbon ion radiotherapy for primary brain tumors and tumors of the skull base, Acta Oncol. 52 (2013) 1504–1509.
- [52] L. Teixeira, C. Ferreira, B.S. Santos, User-centered requirements engineering in health information systems: a study in the hemophilia field, Comput. Methods Programs Biomed. 106 (2012) 160–174.

Provided for non-commercial research and education use. Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/authorsrights