Medical Informatics Europe 2002
G Surján et al. (eds)
IOS Press 2002

A Web-based Central Diagnostic Data Repository

Thomas H. MÜLLER, Klaus ADELHARD

Institute for Medical Informatics, Biometry and Epidemiology
University of Munich, Marchioninistr. 15, D-81377 Munich, Germany

Abstract. A central repository for diagnostic information about individual patients was created as a service to diagnostic laboratories participating in the Compentence Network for Acute and Chronic Leukemias in order to support health care delivery to patients suffering from leukemia. During the diagnostic phase several specialised laboratories perform different assays on samples from the same patient. The use of these assays in the diagnostic process and during the treatment phase may be improved in terms of both rapid delivery and cost if any one laboratory is aware of preliminary or final results from the assay carried out in other laboratories. In order to support a more efficient communication of these results, a central diagnostic data repository (CDDR) was created and web-based user interface was developed. Currently, the CDDR maintains documents in the form of portable document format® (PDF) files. Several other formats are accepted and converted automatically upon entry. Patient identification is accomplished by pseudonym rather than proper name and the data is held on the CDDR for a limited time interval to accommodate the stringent privacy regulations in Germany. The principle operation of a CDDR may also be applied to the diagnostic or therapeutic process of other diseases.

1. Introduction

In the beginning of 2000 the Competence Network for Acute and Chronic Leukemias was founded as one of several networks of competence, each directed at a specific disease [1,2]. The intent of these organisations, for which initial funding has been provided by the German government, is to improve medical research and healthcare delivery by creating both technical and organisational communication structures for the clinicians and researchers involved.

The development of new diagnostic techniques has significantly expanded the classification of the various forms of leukemia. As a result, the treatment can and must be adapted to the specific form as well and as timely as possible. The rapid completion of the diagnostic phase is particularly urgent in cases of acute myeloic and lymphatic leukemia. The precise diagnosis, which crucially determines the therapeutic strategy, results from a variety of assays spanning classical morphology and cytology, cytogenetics, immunological characterisation of the phenotype as well as molecular genetic techniques such as fluorescence in-situ hybridisation (FISH). For an individual patient these assays are performed by several – usually three or four – specialised laboratories that receive blood or bone marrow samples. The process can be accelerated and the cost reduced by communicating early or even preliminary results among the laboratories involved. In the past this has been done via telephone. An electronic document-based form of communication using a central repository is more efficient because it requires only one action on the part of the sender to distribute the information to all participants and because it allows asynchronous retrieval on the part of the recipient. The internet and world wide

web are an obvious candidate for providing the technical infrastructure supporting such a service due to the high availability and acceptance.

Hence, a web-based means of communicating diagnostic information – the CDDR – was designed and implemented. It receives both preliminary and final reports from each participating laboratory and in turn provides information entered by other laboratories regarding the same patients. The CDDR has been designed in compliance with German privacy regulations and may serve as a model for other diagnostic and therapeutic processes.

2. Methods

The CDDR comprises a relational database containing patient and laboratory identification and a collection of PDF files that contain the actual medical information. It is implemented on a standard personal computer running Linux® and the public-domain database system PostgreSQL.

Patients are identified by pseudonyms rather than their proper names as this is one of the legal prerequisites for maintaining a central database of patient-related medical data in Germany (see also [3]). Pseudonyms are generated by cryptographic encoding of personal information such as surname, given name and date of birth. Hence, the true identity of a patient is not known to the CDDR. Since the generation of the pseudonym should be independent of the CDDR, a pseudonym generator providing central repositories of patient-related data with this service is currently under development by the Fraunhofer Society. Until this service becomes available an identification code derived from the patient's date of birth, sex and the initial letters of surname, given name and the city is used in place of the pseudonym.

The web-based user interface relies on the well known webserver Apache and is implemented using the script language Perl. The interface uses the secure hypertext protocol provided by the Apache's SSL-based implementation. Currently, the basic authorisation mechanism is employed for user identification.

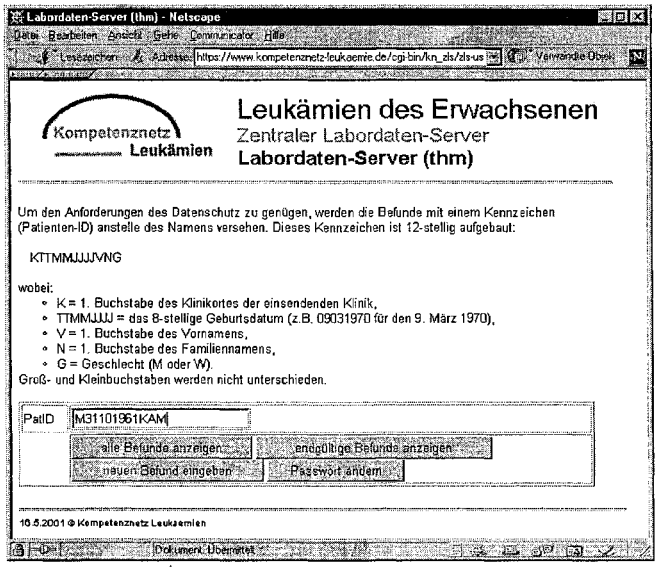

Figure 1: Main screen of CDDR user interface

3. Results

The CDDR and its web-based interface were designed for straightforward and easy use. The technical requirements regarding the client system were kept to a minimum. In particular, no Java engine is necessary. Only an SSL-conformant web browser with JavaScript is needed.

Following the initial user identification dialog, the initial screen (Figure 1) is presented. At this point, the interface accepts a patient identification code of the type described in the methods section. The user can then either:

- enter a new report,
- retrieve previous reports related to the specified patient (currently, two selection options are provided, yielding either final and preliminary reports or final reports only),
- change his/her own password.

Figure 2 shows the report entry form. In addition to the report date and a local identification string the form allows for a brief summarising caption that will help identifying the desired report from retrieval lists. The user's full name and the originating laboratory are added automatically. The actual report document is entered as a local file to be uploaded. In addition to "native" PDF files, the interface currently accepts PostScript®, plain ASCII text and Microsoft® Word documents. These are all converted to PDF. The interface supports a generic mechanism for adding further conversion filters.

The retrieval process is illustrated in Figures 3 and 4. Retrieving reports related to a specific patient will first yield a chronological list of these reports from which an individual report can be selected.

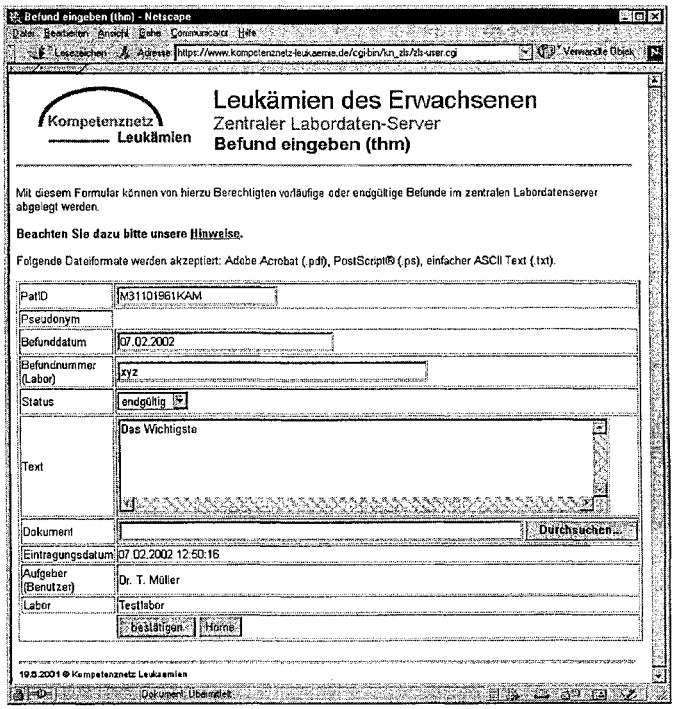

Figure 2: Report entry form.

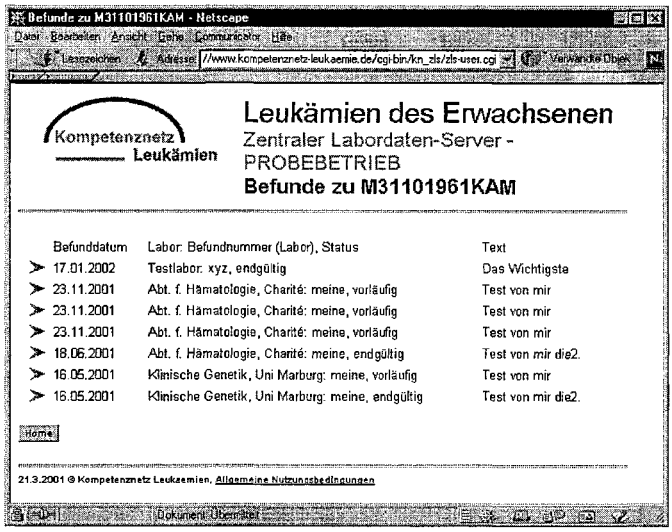

Figure 3: Report retrieval list

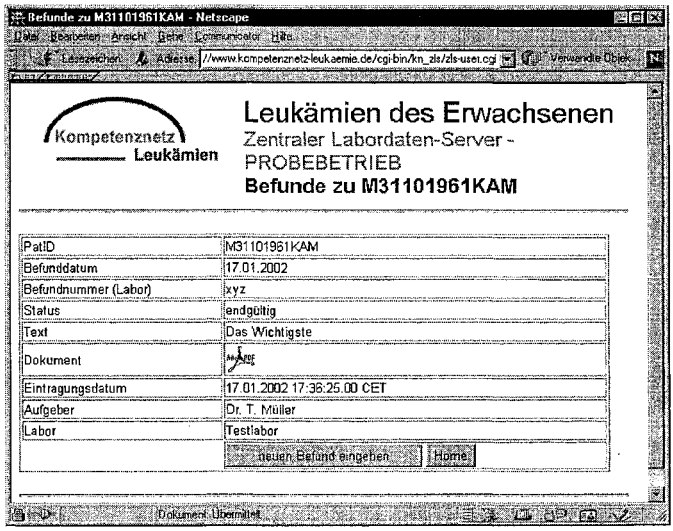

Figure 4: Report retrieval

Due to privacy considerations, the information may not be stored indefinitely in a central database. Therefore, reports are kept available only as long as the concerned patient is treated by one of the participating physicians and are removed when such treatment ends. The treatment is assumed to be terminated when no additional entries are made within six months of the last entry.

4. Discussion

The Competence Network for Acute and Chronic Leukemias provides a framework to improve research and healthcare. Within this context the aim of the CDDR is to provide an efficient means of communicating diagnostic information in order to accelerate the diagnostic process and help reduce the cost of assays. Preliminary and final reports can be communicated rapidly and asynchronously among the participating laboratories.

Due to the severity of this disease and the existing structures of health care delivery, the number of specialised laboratories involved is limited. Furthermore, most of these institutions are part of or associated with university hospitals that provide a modern technical infrastructure, in particualr with respect to reliable internet relay. This has provided a good setting for developing the CDDR as a communication service and addressing technical, organisational and legal problems.

Currently, the CDDR maintains essentially unstructured documents. This is essentially results from the lack or near absence of standardised medical data structures and communication protocols. It is intended to expand the CDDR's capabilities to include structured medical information. The development of corresponding "minimal data sets" specific for leukemia is currently under way. Ultimately, the principle of a central repository for partial or preliminary medical data may be applicable to the treatment of other diseases.

References

[1] Adelhard K, Müller TH. Integrated Healthcare and Research for Leukemia A Concerted Action for Germany. **Proc. AMIA Fall Symp.** 2001.

[2] Kompetenznetz Leukämie [online] http://www.kompetenznetz-leukaemie.de

[3] Schoenberg R, Safran C. Internet based repository of medical records that retains patient confidentiality. **BMJ.** 2000 Nov 11;321(7270):1199-203.