Development of a Virtual Private Database for a Multi-institutional Internet-based Radiation Oncology Database Overcoming Differences in Protocols

Hajime Harauchi, BSc 1), Takashi Kondo, BSc 2), Yu Kumasaki, BSc 2), Masatoshi Ishibashi, BSc 2), Hodaka Numaasaki, BSc 2), Hiroko Kou, MS 2), Yasuhiko Okura, MS 2), Tokuo Umeda, PhD 3), Akihiro Takemura, BSc 4), and Kiyonari Inamura, PhD 1)

1) School of Allied Health Sciences, Faculty of Medicine, Osaka University
2) Graduate School of Medicine, Course of Health Sciences, Osaka University
3) School of Allied Health Sciences, Kitasato University
4) School of Health Sciences, Faculty of Medicine, Kanazawa University

Research Code No.: 220. 3

Key words: multi-institutional database, radiotherapy, clinical record, internet, virtual private database

Abstract

A multi-institutional Radiation Oncology Greater Area Database (ROGAD) was started in 1991 under the direction of the Japanese Society for Therapeutic Radiology and Oncology (JASTRO). Use of ROGAD was intended to allow reflection of results of data analysis into treatment strategy and treatment planning for individual cases, to provide quality assurance, to maximize the efficacy of radiotherapy, to allow assessment of new technologies or new modalities, and to optimize medical decision making. ROGAD collected 13,448 radiotherapy treatment cases from 325 facilities during the period from 1992 to 2001. In 2000, questionnaires were sent to 725 radiotherapy facilities throughout Japan, to further obtain the situation of the radiation oncology database. Workers at 179 facilities replied that “the protocol of my facility is different from ROGAD protocol and I must send data according to the ROGAD protocol”. So, we developed the Virtual Private Database System (VPDS) which is operated as if an
oncologist had a database solely owned by his own facility, in spite of actually operating ROGAD. VPDS réalisés integration of different plural databases, regardless of differences in entry methods, protocols, definitions and interpretations of contents of clinical data elements between facilities.

Received Jan.28, 2002; revision accepted May.27, 2002

1. Introduction

A study group for the “Construction of a multi-institutional Radiation Oncology Greater Area Database (ROGAD)” was begun in 1991 as a part of the activities of the Japanese Society for Therapeutic Radiology and Oncology (JASTRO). ROGAD was intended to allow reflection of results of data analysis into treatment strategy and treatment planning for individual cases, to provide assurance by radiotherapy technologists and physicists, to maximize of the efficacy of radiotherapy, to allow assessment of new technologies or new modalities, and to optimize of medical decision making.\(^1\) - \(^6\).

The data elements were selected to reflect the present status of radiation therapy. Parameters and variables were optimized to reflect real distribution of clinical cases. The number of data elements of data was kept limited and the items were simplified for quick and easy registration. International codes such as those of the International Classification of Diseases for Oncology (ICD-O), UICC and WHO were adopted whenever possible. In all ROGAD had 53 data elements in 13 categories per patient, and had 10 items for follow up.\(^7\),\(^8\) Data collection of radiotherapy records was carried out nine times during the period from 1992 to 2001. Follow-up data collection was carried out six times. ROGAD had 13,448 effective cases of radiotherapy collected using worksheets and by offline data collection software using the ROGAD protocol. Here, “protocol” means an agreement in database formats for receiving and sending data between databases. In other words, the protocol is a definition of the terminology, syntax, semantics, and structure of the database. In total clinical records from 325 institutes in Japan registrated. There were 3,112 effective follow-up data cases from 131 facilities. The number of data converted by data-link software from four facilities was 24,000 cases. Differences in radiotherapy environments as well as the results of radiotherapy between facilities have turned up.\(^9\) - \(^17\).

In 2000, questionnaires were sent out to 725 radiotherapy facilities throughout Japan, to further obtain the situation of the radiation oncology database. There were 242 responses (return rate: 33 %). Ninety percent of the replies were “I need a radiation oncology database”, 69 % of replies were “I employ a radiation oncology database”, and 74 % of replies were “The protocol of my facility is different from ROGAD protocol and I must send data according to the ROGAD protocol”.

The most difficult problem to be solved for effective data collection was differences of protocols of radiotherapy record between facilities. Generally, data collection is done in two ways. One way is data entry with a standard protocol such as ROGAD protocol and the other is with different protocols by facilities where data conversion is necessary.\(^18\) - \(^20\).
In the past, even if a facility’s requests for protocol change were strong, we did not modify the ROGAD protocol because consistency had to be kept before and after modification, and deterioration of reliability had to be avoided.

**Purpose**

Consequently, in this study we developed the Virtual Private Database System (VPDS) which is operated as if an oncologist had a database solely owned by his facility, in spite of actually operating within the multi-institutional database ROGAD.

Also we had to establish a security methodology against leakage of patient ID and data in ROGAD.

The final target of VPDS development and completion of the methodology was to optimize the en-

![Fig. 1 Comparison of VPDS with plain internet-based database system in terms of flexibility and operability.](image)

(a) The Virtual Private Database System
(b) A General Internet-based Database System

(a) An operator can operate the database through the individual HTML as if he is entering clinical records into his own proper database in spite of his entering data into ROGAD.
(b) An operator must operate the database with ROGAD protocol, because on-line data entry through uniform HTML can not respond to the demands of individual facilities.
virement of clinical data collection, to remove the boundary between ROGAD and existing local databases specific to institutions, to increase the number of registered cases, to improve reliability of data contents, to improve speed of circulation of statistics of clinical data, and to enhance operability of data management with the minimum manpower.

Materials/Methods

There are three methods for ROGAD data collection:

(1) Hand-written worksheets are sent back to the ROGAD office from radiotherapy facilities. Until 1995 ROGAD office staff had to enter them into ROGAD with data collection software. That year, an optical character reader (OCR) was introduced. In 2000, a marksheet reader with facsimile was implemented and now hand-written worksheets are automatically read and filed in the database.

(2) Radiotherapists enter clinical data into personal computers in each facility by operating the data collection software of the ROGAD protocol. The generated ROGAD data registration floppy disks which contain clinical records are then mailed to the ROGAD office.

(3) When a facility has its own database in operation, the clinical data are transferred to ROGAD protocol data and filed into ROGAD. This is done automatically by interpretation software developed by us. We call this method "Data-link".

These three methods have been working well. However, because the amount of labor in the ROGAD office to manage clinical data has increased in proportion to the increase of the number of facilities supplying information, a new method of data collection was required. "VPDS" is considered to be one of the best solutions for both the facility and the ROGAD office.

VPDS was developed based on Microsoft Visual Basic 6.0, Microsoft Jet Database Engine, Microsoft Windows 2000 server, and Microsoft Windows 98 2nd edi-
A subscriber can easily customize an entry screen, the protocol, and the definition of data items in accordance with the protocol provided with his own proper database, and can save them in a new individual HTML (Hypertext markup language) on the facility’s client computer, (on-line data entry through uniform HTML can not respond to the demands of facilities) (Fig.1). VPDS is divided into two parts as shown in Fig. 1(a). The part on the facility’s client computer consists of our developed browser, individual HTML, and protocol conversion program. The second part, on the ROGAD server computer, consists of the interface program and each facility’s database. Fig. 2 shows that an operator can easily customize and operate the database through the new individual HTML as if he is entering clinical records into his own proper database in spite of his entering data into ROGAD. VPDS automatically converts clinical data written by different protocols in different facilities to data of ROGAD protocol, and it also converts from ROGAD protocol into the protocol of each institution. Fig. 3 shows the VPDS has expanded flexibility and customization. Fig. 4 describes how the VPDS keeps security against leakage of patient privacy. Patient ID data are saved only in the facility’s client computer, separate from radiation therapy records in the ROGAD server computer. In order to access the ROGAD server computer, an operator must be authenticated by user

![Customized Entry Screen of VPDS](image1)

![Uniform Entry Screen of ROGAD](image2)

**Fig. 3 Comparison of customized entry screen of VPDS and uniform entry screen of ROGAD.**

The data elements of ROGAD were selected to reflect the present status of radiation therapy. Parameters and variables were optimized to reflect real distribution of clinical cases. The number of data elements was kept limited and the items were simplified for quick and easy registration. For example, “Field size” and “Wedge filter” are not included in the data elements of ROGAD as shown in (b).

On the lower part of the entry screen of VPDS shown in (a), new data items are added. The codes list and contents can be displayed with a combo box. The order of entry of the data elements can be changed in accordance with requirements from the facility, too.
Fig. 4 Security enhancement against leakage of patient ID in the VPDS.

Security against patient ID leaks is improved, because patient ID data are kept in the facility’s client computer separately from radiation therapy records in the ROGAD server computer, so that operators at other facilities and hackers cannot see patient ID data. Further, ID number for reference is provided in requests instead of the patient ID which does not exist on the ROGAD server computer.

ID number, password, and the facility’s client computer information. Furthermore, the data flow is managed by keeping a session, and the data which flow on the internet are encrypted with SSL (Secure Sockets Layer), so that operators from other facilities and hackers can not see patient ID data.

Results

One facility’s radiotherapy database had 89 data elements in 14 categories. The protocol of this facility was obviously different from the ROGAD protocol. We applied VPDS to this facility’s radiotherapy database. An operator in this facility would enter radiotherapy data with the VPDS protocol where 36 data elements (89-53 = 36, and 53 are stated in the “Introduction”) were added and 1 category (14-13 = 1, and 13 are stated in the “Introduction”) of the ROGAD protocol was also added. Definitions of 10 data elements in the ROGAD protocol were changed, and suited to the protocol in the facility. At the same time, the entered radiotherapy records were saved in ROGAD with the ROGAD protocol. As a result, the deficit rate for data conversion was 0%. Moreover, even when there were fewer items in the facility’s database than in ROGAD, the number of input items could
be similarly reduced on VPDS.

**Discussion**

The Virtual Private Database System provided improved operability, flexibility, security and serviceability, as described in the following items.

1. Our VPDS was compared with a plain internet-based database system, and the advantage of VPDS in terms of flexibility and operability was clearly found.
2. Circulation and exchange of clinical data without delay was conducted between different facilities of different protocols.
3. Reliability of entered data in ROGAD increased because of automatic checking for errors and emptied entries in ROGAD office.
4. Labor saving and time saving at each facility were realized because of the centralized management of collected data in the ROGAD office.
5. Security against leakage of patient ID was improved by saving patient ID information only on the facility's client computer, and by having an authentication system, data flow management, and SSL.
6. Statistics service with the macro function of Microsoft Excel 2000 was provided, so that the survival rate curve was instantly calculated and displayed on the facility's client computer (Fig. 5).

![Fig. 5 An example statistics service of the Virtual Private Database System (VPDS).](image)

For a facility that adopts VPDS, statistics service with the macro function of Microsoft Excel 2000 is provided, so that the facility's most recent radiotherapy data are downloaded from the ROGAD server computer to the facility's client computer, the survival rate curve is calculated in only 30 seconds, and it is displayed.
methods, protocols, definitions and interpretations of contents of clinical data elements between facilities.

However, there are a few problems which must be solved in the future, as described in the following items.

1) In case a facility’s registration code is macroscopic and differently classified compared with ROGAD, the facility’s registration data are rather difficult to convert to the ROGAD protocol.

2) In case a facility’s client computer is changed for security, software for registration cannot be easily improved without changing the facility’s security environment.

**Conclusion**

Under present conditions, when a facility wants to operate a multi-institutional database, the details of the facility’s protocol and computer environment of the facility may encounter some difficulty in modification.

However, by applying VPDS, an operator can enter and see clinical records from a multi-institutional internet-based radiation oncology database in spite of differences in protocols between institutions. The VPDS has contributed to an increased number of participating institutions, to an increased number of registered cases, and to improved security of patient ID data on the internet.

Infrastructure of linkages with different protocols of databases has been promoted by means of implementing the VPDS. So, standardized assessment of radiotherapy results as well as the environmental situation in radiation oncology in a large area are expected to be realized. Quality control of radiotherapy and enhancement of radiotherapy results will be stimulated by our method of infrastructure construction.

**References**


