ORIGINAL ARTICLES

Standard Protocol for Exchange of Health-Checkups Based on SGML: The Health-Checkups Data Markup Language (HDML)

Hiroki Sugimori; Katsumi Yoshida; Shoichiro Hara; Katsuhiko Furumi; Ikuo Tofukuji; Takeshi Kubodera; Takashi Yoda; and Masaki Kawai

1 Department of Preventive Medicine, St. Marianna University, School of Medicine, 2 Research Information Department, National Institute of Japanese Literature, and 3 Committee of Japanese Association of Healthcare Information Systems Industry (JAHIS), c/o Toranomon TBL Building, 1-19-9, Toronomon, Minato-ku, Tokyo, 105-8070 Japan.

ABSTRACT

This research aims to provide the protocol to achieve efficient information exchange by means of electronic data communication between health-checkup facilities. Joint Working Group of JMHTS (The Japan Society of Multiphasic Health Testing and Service) and JAHIS (The Japanese Association of Healthcare Information Systems Industry) developed a health/medical data interchange model that stood on the markup information structure. Our data encoding language, HDML (Health-checkups Data Markup Language), was based on SGML that has context-free grammar. HDML had the standard DTD which defined anamnesis, physical examination, laboratory examination, summary findings, and judgment of total health status, etc. The laboratory examination contains following items: item's name, method, unit, device, company name, product name, principle, and standard reference value. We take into account the interchangeability of data with HL7 and other international standard protocols. As a preliminary study, we carried out an experimental trial in October 1999, which transferred laboratory data by translating into HDML, from 2 health-checkup facilities to other 2 health-checkup facilities. We have succeeded in transferring almost all laboratory data appropriately by using the HDML protocol between the health-checkup facilities. Moreover, we could convert and standardize the laboratory data properly from the information written in the DTD. We propose the HDML protocol to standardize health/medical data that will make it available for multi-health facilities on the basis of the standardization of data exchange regarding health-checkups. We found this HDML protocol worked effectively in using the actual health/medical data.

Key Words: Health-checkup; SGML; HL7; HDML

INTRODUCTION

In Japan, since regular health-checkups (HCs) in working places are mandatory according to the Japanese Industrial Safety and Health Law, HCs have been widely and frequently carried out in workplace. HCs were originally used as a tool for disease screening contributing to early diagnosis and early treatment of the disease (secondary prevention). Recently, due to a rapidly aging society, the role of HCs has been strengthened and shifted into disease prevention and health promotion in community-health fields in Japan (primary prevention). The disease structure in Japan has changed, and lifestyle-related diseases are becoming a problem as is the national health burden. To manage lifestyle-related disease efficiently, we should use HC data as source for active improvement of individual lifestyles. From the viewpoint for prevention of lifestyle-related disease, accumulation of seamless integration of personal electronic HC data is important. Therefore, we needed to consider effective and efficient usage of HCs data and construct a lifelong health care management system.

The management of HCs data is relatively easy as long as the data is collected, stored and viewed in a closed environment like a single HC facility/company (information island). However, in Japan most of companies conduct employee HCs by outsourcing other HCs facilities such as a Multiphasic Health Testing and Service (MHTS). One HC facilities is often contracted with many companies (multi-clients). Moreover, one company often contracts with several outsourcing HC facilities to meet their requirements.

However, HC data systems vary heterogeneously in Japan; from HC facility to another, and company to company. In a general way, various terms and codes are actually used in both the system of HC (facilities) and the system of health care managements (companies). In addition, the information interchange protocol between these systems varies diversely; even the sequencing of data is different (spaghetti information structure). Thus, as yet, due to different system format, standardization and integration of HCs information among multi-facilities systems is insufficient and requires much effort to exchange HC data.

This research aims to provide the Health-checkups Data Markup Language (HDML) protocol that enables information exchange efficiently by means of electronic communication even between different HC (facilities) and health care management (companies) systems.

MATERIALS AND METHODS

2.1 History

Since 1998, we have organized Joint Working Group of The Japan Society of Multiphasic Health Testing and Service (JMHTS) and The Japanese Association of Healthcare Information Systems Industry vender (JAHIS), in 1999 we worked together to develop a new health data transfer and interchange protocol, that is HDML, and its software.1-3)

2.2 Structure of HDML

Our data encoding language, HDML, is based on Standard Generalized Markup Language (SGML; ISO8879: 1986), which is an open, internationally accepted standard for document interchange. The structure of HDML has the standard Document Type Definition (DTD), which defines anamnesis, physical examination, laboratory examination, diagnosis, summary findings, and judgment of total health status, and so on.
HDML is comprised of two major parts (Fig. 1), that is the header and the body. The header is used for pre-processing across all level of the data. It includes detailed authentication information of the HC service, the HC facility, the client, and so on. The header also provides the default attributes of data items. For instance, the laboratory examination item contains the following attribute information of value: item name, data type, unit, method, device, number of measurements, condition, ingredient, and standard reference value.

HDML description syntax is as follows,

Syntax:

<Tag-Name Attribute-lists> Value </Tag-Name>.

Where, "tag-name" identifies the data item, and "attribute-lists," describes meanings (properties) of the data item, and its "value." Attribute-lists give us sufficient information of data value. We can define attribute information either in the header as default data or in the body. Since the data item is identified by attribute tag information, the sequence of data item is not restricted by the data definition.

Example (low density lipoprotein: LDL)

<rTest c = "S3" n = "LDL" Unit = "mg/dl" Upper = "160" Lower = "50" Data_Type = "NM" Code_Type = "L" > 100</rTest>

We have also taken into account the interchangeability of data with Health Level Seven (HL7), Medical Markup Language (MML), and other standard protocols. For the compatibility of standards, HDML is designed based on the OBX data segment of HL7, and several attributions have been added in accordance with the discussions in the Joint Working Group.

2.3 HDML exchange software (CSV to HDML and HDML to CSV)

We have also developed an exchange software which converts HC data from CSV format into HDML format and visa versa. Before exchanging into the HDML transfer file, the HC facilities prepares both their health data (CSV format) and their original code table which describes the attribute information of value such as unit, method, reference values. By converting this original code table information into JAHIS standard code table, the sender’s original health data can be standardized. This exchange software has the user-friendly interface and sender/receiver facilities can easily define their own original attribute information. Figure 2 shows a sample of the software display that determines the attribute information of value data in the original system. Figure 3 shows an outline of transferring HCs data from the sender facilities system to the receiver company system via HDML.

2.4 Feasibility study (Pilot implementation)

As a preliminary step, we carried out a feasibility study in October 1999, which transferred laboratory data by translating it into HDML from two health-checkup facilities to two other health-checkup facilities. First, we translated laboratory data (sender’s original CSV format) into HDML. Second, we translated inversely HDML into laboratory data (receiver’s original CSV format). Our purpose of this feasibility study was to confirm the appropriate transfer with the accurate sequence of laboratory data and to confirm the appropriate storage in the receiver’s own health management system.

RESULTS

We succeeded in appropriately transferring almost all the laboratory data using the HDML protocol between the health-checkup facilities. We were also able to convert and standardize the labora-
DISCUSSION

We propose the HDML protocol to standardize HCs data making health data transfer available between multi-health facilities. In the context of the standardization of data exchange, we found this HDML protocol worked effectively for the actual HCs data.

Since HL7 is most successful among many standards for its syntax and an authoritative standard for medical informatics exchange between healthcare providers, standardization of patient’s electronic data records by HL7 has been proposed in Japan. Therefore, we have taken into account the interchangeability of data with HL7 on exploiting HDML. Both HL7 and HDML have similar characteristics derived from SGML, such as a context free grammar, a flexibility in describing any kinds of text. However, our HDML has several different characteristics from HL7 in its exploiting concepts.

1. HDML is characterized by its light protocol that focuses only on individual HC data transfer, and it is easy to implement and to use. HL7 is a heavy protocol to implement and to modify, because HL7 focuses on complex hospital patient records interchange.

2. HDML is focusing on users in HC facilities which deal with mass health data (a batch of HC data). In this situation, most of health data properties (attributes) are uniform; i.e., same method, same unit, same device, same device company, same standard reference value. Therefore, HDML can handle default information in the HDML header and facilitates the data transfer by eliminating redundancy, which is HDML expansion compared to HL7.

Furthermore, in the context of standardization, we also take into account the concept of “common unit” in HDML. This index is calculated as follows,

\[
\text{common unit} = \frac{(\text{original laboratory data} - \text{standard median}) \times 100}{\text{standard SD}} + 100
\]

This concept is similar to the idea of Z-score or standard deviation index (SDI), and this will enable multiple laboratory numerical data measured by different methods (e.g. different reagent) to be integrated into “common unit” and be interpreted appropriately. Moreover, this will enable the display of different items (such as aminotransferase: AST and lactate dehydrogenase: LDH) into same single standard (i.e. common unit). Then most laboratory numerical data can be displayed in the same single median (100) and the same single range (from 80 to 120), and the patient/subject can easily understand the level of his/her own data by avoiding medical language and technical terms (i.e. the user-friendly interface). This may contribute to informed consent of patient/subject in the future.

Our proposal of HDML and the software have several following important advantages.

1. By introducing our HDML protocol exchange software, both the sender (MHTS) and the receiver (company) can handle HC data with one tool.

2. Since our software can intermediate CSV format, we do not have to compel health providers to change their present systems and they can utilize present resources.

3. With standardized HC data at the sender’s stage, individual personal health data that is examined in different facilities (multi MHTS facilities) will be standardized easily and will be able to accumulate seamless integration of personal electronic HC data (lifelong health care management system).

Finally, these advantages may contribute to the prevention of lifestyle related diseases and aid in cost-efficiency (Fig. 4).

We still have several challenging tasks for the future regarding HDML implementation. The syntactical and semantic mapping between different heterogeneous systems will be the main issue for automated data exchange in the long term. Other than the numerical data, such as qualitative data, namely anamneses or categorical data, we have to develop new syntax for precise exchange of data. In addition, we need security standards that ensure an individual’s health information remains confidential and is protected from unauthorized or inadvertent disclosure, alteration, or destruction. Security standards are especially important because electronic health records make information accessible to multiple users in multiple locations.

CONCLUSION

1. HDML will be very useful to transfer HC data among HC facilities and companies.

2. HC data resources may be used effectively by using the standardized transfer protocol, HDML.

3. HDML provides applicable situations on present health/medical resources by using our HDML protocol exchange software (i.e. CSV converter). We have succeeded in the feasibility study and could utilize present resource in MHTS.

4. Lifelong health management may be possible by sharing HC data by HDML

ACKNOWLEDGMENTS

This research is supported by The Medical Information System Development Center (MEDIS-DC, Tokyo, Japan).

REFERENCES


