MEDINFO 2017: Precision Healthcare through Informatics
A.V. Gundlapalli et al. (Eds.)

doi:10.3233/978-1-61499-830-3-516

Case Report Form Reporter: A Key Component for the Integration of Electronic Medical Records and the Electronic Data Capture System

Yasushi Matsumura[a, b], Atsushi Hattori[b], Shiro Manabe[a, b] Daiyo Takahashi[b], Yuichiro Yamamoto[a, b], Taizo Murata[c], Akito Nakagawa[a], Naoki Mihara[a], Toshihiro Takeda[a]

[a] Medical Informatics, Osaka University Graduate School of Medicine, Osaka, Japan
[b] MKS ltd, Osaka, Japan
[c] Division of Medical Informatics, Osaka University Hospital, Osaka, Japan

Abstract

To improve the efficiency of clinical research, we developed a system to integrate electronic medical records (EMRs) and the electronic data capture system (EDC). EDC is divided into case report form (CRF) reporter and CDMS with CRF receiver with data communication using the operational data model (ODM) . The CRF reporter is incorporated into the EMR to share data witth the EMR. In the data transcription type, doctors enter data using a progress note template, which are transmitted to the reporter template. It then generates the ODM. In the direct record type, reporter templates open from the progress note and generate narrative text to make record in the progress note. The configuration files for a study are delivered from the contents server to minimize the setup. This system has been used for 15 clinical studies including 3 clinical trials. This system can save labor and financial costs in clinical research.

Keywords:

Electronic Health Records; Clinical Study; Data Systems

Introduction

We need a method for transcribing the data recorded in a patient's medical record to a predetermined form for submission to a party outside of the hospital. In the era of managing medical records via a paper medical chart, the data written in the record needed to be transcribed by hand to another piece of paper; however, in the era of electronic medical records (EMRs), we now need a way to fill out forms without redundant data input. This function is particularly important in clinical research and would also be useful when sending a report to external organizations when for example following up a patient in collaboration with other medical facilities.

In clinical research, such as clinical trials, observational studies and registries, data for pre-determined items are systematically collected for patients who meet the entry criteria. In clinical trials, source data verification (SDV) by a monitor is required to confirm the consistency of the data in the case report form (CRF) with those in the medical record. This increases the costs associated with a study. In observational studies and registries, data are often entered without the support of clinical research coordinators (CRCs); as such, redundant entry in the medical record and electronic data capture system (EDC) is a major obstacle to promoting these studies. As registries become more popular, we may encounter instances where one case needs to be submitted to multiple registries, which results in redundant data input.

Interoperability between medical care and clinical reserach would be advantage, but has not been achieved in a wide scale [1]. The Clinical Data Interchange Standard Consortium (CDISC) Electronic Source Data Interchange Group identified issues that may inhibit adoption, explored the values and benefits of implementing standards for data acquisition, exchange and archive of eSource [2]. We developed and already reported on a system to achieve these issues [3]. In our model, the EDC is divided into a CRF reporter, which functions as a data input and transmission unit, and a clinical data management system (CDMS) with CRF receiver, which functions as a data reception and storage unit. Data communication between the CRF reporter and the CRF receiver is achieved via an operational data model (ODM) developed by CDISC [4]. The CRF reporter is incorporated into an EMR system that can automatically retrieve the data recorded in the EMR and output narrative text generated from the data included in the CRF reporter to the medical record. In order to automate setup, a template master, an ODM and a study configuration file are distributed via a network from a content server to the hospitals participating in the study (Figure 1). Since the previous report, we have added several functions and refined the system further. In this paper, we describe our system, with focus on the CRF reporter, which is a key component for integration of EMRs and EDC.

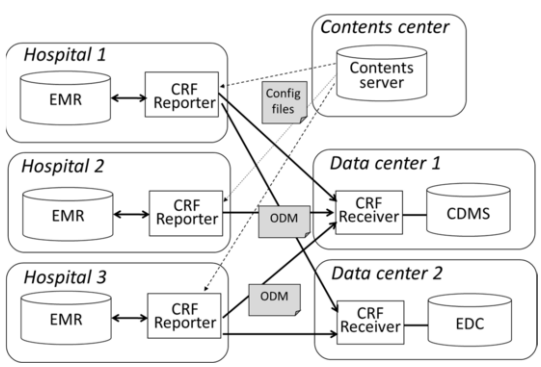

Figure 1 – Components of the System.

Methods

System Overview

There are two types of CRF reporters: a data transcription type, and a direct record type. The functions of each are outlined below.

Data Transcription Type

When recording progress notes, the data used for clinical research are input using a template. The CRF reporter is an application independent of the EMR. In the CRF reporter, after selecting a study and a subject, an event list is displayed. One event consists of several forms. When a form is selected, the corresponding template is displayed, along with a data retrieval button. By clicking this button, the data entered by the template for the progress note or data (e.g. laboratory test data) stored in the EMR database are retrieved and used to populate the items in the template. When the input of the template set corresponding to a study event is completed, the input data are mapped to the ODM, which is sent to the data center (Figure 2).

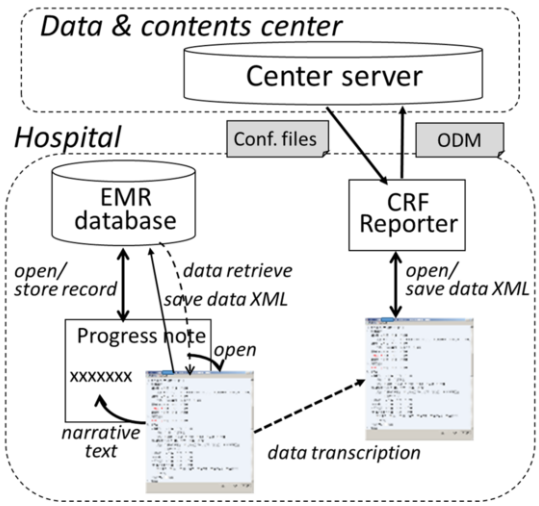

Figure 2 – Configuration of Data Transcription Type.

Direct Record Type

The start button of the CRF reporter is placed in the progress note. Clicking this button brings up a list of studies in which the patient and the user are participating. After selecting a study, a list of events with forms is opened. Selecting a form displays the corresponding template. After data input is completed, the narrative text is generated and added to a frame in the progress note. The subsequent process is the same as that of the data transcription type (Figure 3).

Input Template Function

The input template is a key module for this system. We developed "dynamic template" to aquire medical records in structured form [5]. If any abnormalities are noted in the observation items, a detailed description of the abnormality is required; this allows the medical description to be presented in a hierarchical tree structure. Therefore a template where the next input item changes according to the selected value is suitable for medical description.

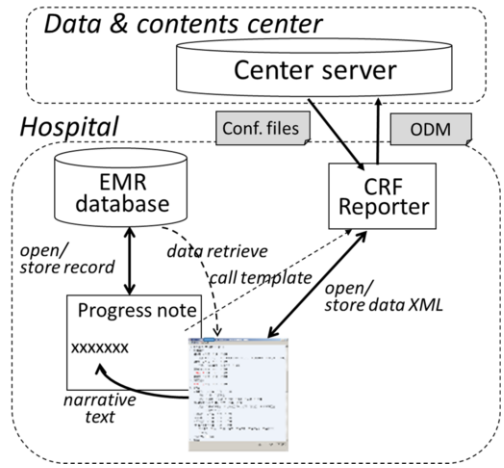

Figure 3 – Configuration of Direct Record Type.

To make it easier for the user to read the data entered by the template, natural narrative text is generated from the input data using natural language generation technology. The template outputs a medical record displayed with descriptive expressions, which is easier for the user to read. It also outputs the data as an XML file (data XML). When modifying the data, the input state is reproduced using the data XML.

The content of the template is controlled by the template master, which is also saved in the XML file (template XML). Therefore, when a different template XML is input into the template engine, different templates are displayed.

Data Retrieval Interface

In the EMR database, data such as patient characteristics (height, weight, etc.) and laboratory test results are stored. These data frequently need to be collected into templates. When retrieving laboratory test results or similar, an assigned test item code is required in order to obtain the value for a designated test item. However, the test item code differs among hospitals. Therefore, a conversion table (correspondence table between a universal keyword and the local code) is created at each hospital. This table also has a multiplied number to convert the value in local units to standard units and has a correspondence table to convert local enumerated values to standard ones.

In the template master, the parameters for data retrieval, such as the test item code, search period, and selection method, are described. In the template master, the test item code is described in the notation generally conforming to the CDISC-controlled terminology. The template engine obtains the local code for the test item and the information for converting the value into the standard form from the conversion table. Then it passes the parameters including the local code, search period, and selection method to the interface module. The selection methods are for example selecting one latest value, one oldest value or user's selection. If there are several corresponding data in a given search period, the data picker module presents them all to the user for selection.

The template can also retrieve values entered in other templates in the past by designating the element code with the template code of the source template or the concept code if it is written to the corresponding element.

Data Mapping to the eCRF

In our system, the ODM was used as the eCRF form outputted from the CRF reporter. After the completion of input, the template engine generates a dataset in a tree structure, where each leaf corresponds to a value and each node corresponds to an item. The value of a node is the narrative text generated from the data under the node. For each ItemData of the ODM, the corresponding value of the leaf or node can be mapped. For example, when a value of "the date of occurrence" is entered by selecting the year and month from a combo box on the template, the leaf value is year = 2010, month = 10, and the node value is "October 2010". Thus, if the ItemData value corresponding to "the date of occurrence" is requested in a string, "October 2010" can be mapped to it.

ODM Transmission

First, we must enter subjects who meet the inclusion and exclusion criteria. At this point, the subject key is determined on the hospital side or issued on the server side; in our system, both are possible. The CRF reporter performs ODM transmission to the CDMS/EDC in the data center. To check the contents of the ODM, a readable case report form is also generated. After checking the contents, the ODM is transmitted to the center server.

Originally, the CRF reporter sent the ODM to one CDMS; however, it must now be able to send the ODM to any CDMS/EDC equipped with an ODM import function in datacenters, depending on the study. Because the communication procedures of ODMs differ, CRF reporters need to be able to switch transmission modules according to the destination center server.

User Management

The CRF reporter needs to manage local users as well as center server users linked to local users. When logging into the CRF reporter, a local user table is checked, but when creating and sending the ODM, the center server user is used instead. The user ID for the center server is recorded in the AdminData element in the ODM. When connecting to the center server, authentication is performed using the center server user.

Setup Automation

Occasionally in clinical trials, there are cases where several subjects are handled at a given hospital. Requiring a system engineer to set up this system at each hospital can be costly, resulting in the operation using our system becoming more expensive than that with a regular EDC. Therefore, we developed a function to automate setup using a template master, an ODM and a study configuration file, all of which are distributed via a network from a content server to the hospitals participating in the study. The only operation required by each hospital is user management and the maintenance of the correspondence table between the universal keyword and the local code. Before starting a study, the operation staff on the hospital side need only to confirm that the ODM is correctly created using test subjects.

User Operation

Data Transcription Type

The CRF reporter is a system independent of the EMR. After logging into the system, it shows the studies for which the user has been given operation authority. When a study is selected, the subjects participating in the study are displayed. When a new subject is entered, patient identification information such as the patient ID, name, gender, date of birth, etc., are acquired from the EMR database added to a correspondence table between the subject key and the patient ID. The patient identification information is then displayed in the subject list to help the user select the correct subject. These data, however, are not sent to the data center.

When a subject is selected, the event list for the subject is displayed. Each event includes several forms. After selecting a form, the corresponding template is displayed. By clicking on the data retrieval button in this template, the fields of the template can be populated by the data if they have already been entered by the progress note templates.

After entering all of the required items for an event, the status of the event changes to "completed", and ODM creation becomes possible. The user checks the contents of the ODM. If no amendment is needed, they click the send button and transmit the ODM to the center server in the data center. When the center server receives the ODM correctly, the data reception status is returned to the CRF reporter.

Direct Record Type

The direct record type is currently under development and set to start operation from April 2017. The functions scheduled to be included are shown in Figure 4.

When a patient is selected in the EMR, the patient's medical record screen is opened. On opening the progress note input screen from this point, an activation button for the CRF reporter is displayed. If the patient is a subject or a subject candidate and the user has been given operation authority, this button flashes. Clicking on this button displays the event list for the study. When a form for an event is selected, the corresponding template is displayed. The user then enters data in accordance with the items on the template. For items such as laboratory test results, the corresponding data are automatically retrieved and populated. By clicking the send buttom, the ODM is created and sent to the center server in the data center.

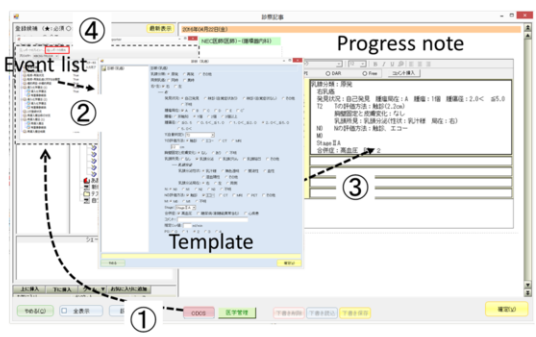

Figure 4 – Operation of Doctors when using Data Record Type
1. Open the event list of the study for the paitnet by clicking activation button for the CRF reporter
2. Select a form to open the corresponding template
3. Output the narrative text to record in progress note
4. Create the ODM and send it to the center server

Results

Dynamic templates have been incorporated into the EMR system developed by NEC since 2000. Thus far, more than one hundred hospitals have employed this system, and more than seventy thousand templates have been created and used in these hospitals. For example, at Osaka University Hospital, 2182 template masters have been registered, and 123,191 templates per month are used, resulting in 1,195,831 data items per month being entered.

At present, only the data transcription type CRF reporter is being used. Three hospitals have implemented this CRF

reporter in their EMR systems; however, at other hospitals, the CRF reporter is installed on a single PC for use. While the CRF reporter cannot cooperate with the EMR system in such cases, it can still be used as a terminal for usual EDCs. The ability to create CRFs offline makes this approach convenient compared with normal EDCs. Using this system, we have carried out 12 observational studies and 3 clinical trials so far. Six of these studies consisted of more than 10 events. These clinical studies have successfully collected data, and six of them have already been completed. The maximum number of the entry subjects were 1119.

Discussion

This system is a special type of EDC comprising a CRF reporter and a CDMS with a CRF receiver. The CRF reporter is built into the EMR. This model is similar to the Retrieve Form for Data Capture (RFD) model proposed by Integrating the Healthcare Enterprise (IHE) [6]. The RFD model consists of a Form Manager, Form Filler, Form Archiver, and Form Receiver. Our model consists of a content management system corresponding to a Form Manager, templates in the CRF reporter corresponding to Form Filler, a local server of the CRF reporter corresponding to a Form Archiver, and a CDMS with a CRF receiver corresponding to a Form Receiver. However, our model is not completely RFD-compliant, as we used a unique template module instead of XForms for the form filler and REST communication instead of SOAP for communication between Form Filler and Form Receiver. However, several challenges to the EDC and EHR interoperable systems have been reported, and a completed system based on the RFD model has not yet been developed.

Regarding the integration of EHRs and EDCs, one report described a system that retrieves the data elements of patinet care-centric template in use in the EHR to pre-populate the CRFs, but only 13.4% of the data elements in the CRFs were able to be retrieved [7]. In our model, even with the data transcription type of CRF reporter, a template used for patient care is purposefully created to allow users to enter data for a given study; as such, almost all data collected in the study can be retrieved. In addition, since narrative expression-style text generated from the data are entered as a record of the progress note and doctors can freely add supplementary text to the progress note, the note can be treated as a medical record as well.

For registries, various patient data such as the symptoms, findings, laboratory test results, image examination reports, and treatments, are entered. In the data transcription type, a template for the progress note can be created freely at each hospital. When registering a single case in multiple registries, if a template with unified items in the case report forms for multiple registries would be created, with the initially entered data then retrieved by the templates for the respective CRFs and the completed CRFs sent to the corresponding data centers.

For clinical and prospective observational studies, certain forms must be input for each visit, and some forms are used across several visits. If such studies use data transcription type CRF reporters, the template corresponding to each visit is not shown and data input cannot be guided. Another problem is that retrieving the data for the same form from different visits using the corresponding CRF reporter form is complicated. In contrast, when using the direct record type of CRF reporter, the events and forms to be entered are displayed and their statuses shown, guiding doctors to input data into the proper template for each visit. The preparation of the CRF reporter for the direct record type is easier than that for the data transcription type. As such, for prospective clinical studies, the direct record type is more suitable.

This system saves labor through the simultaneous entry of the medical record and the CRF as well as the automatic retrieval of the data recorded in the EMR. The former is an essential function of this system, but the latter depends on the interface function between the template and EMR. The more easily various data can be retrieved from the EMR database, the more efficient clinical research becomes. Details of this function of our system were reported in MIE in 2016 [8].

When integrating EDCs and EHRs, locus of responsibility should be clear [9]. With this system configuration, the CRF reporter is a investigator side system. Therefore, we believe that the validation for clinical research is required within the scope of using a CRF reporter. For data transcription type CRF reporters, since the system automatically retrieves the data and applies it to the template of the CRF reporter, the mechanism of correct retrieval must be validated. For the direct record type as well, we must also validate the automatic retrieval function for laboratory test data and similar values; however, the data retrieval range is narrow compared to the data transcription type. For the direct data record type, since data is output to the medical chart, we must validate whether or not this output is correctly recorded in the EMR. From the viewpoint of the locus of responsibility, the investigator side should be responsible for the CRF reporter and for directly transferring data between the CRF reporter and EMR, regardless of data transcription type or direct record type.

For the data transcription type, the doctor may describe the record in free text in the progress note, and the CRC manually transcribes the data to the template of the CRF reporter. Even when doctors input data using the template in the progress note, after the CRC transcribes the data into the CRF, the doctor may change the data using the template, which may result in inconsistencies between the CRF and EMR record. Therefore, interaction with the data transcription type CRF reporter requires an SDV, as in ordinary EDCs. In contrast, in the direct record type, the contents entered by the CRF reporter are forcibly added to the medical record, so it matches the contents of the medical record. Also, when revising, since the template of the CRF reporter is opened and modified, inconsistency does not occur. Therefore, an SDV is unnecessary when the direct record type CRF reporter is used. This further underscores the suitability of the direct record type over the data transcription type for clinical trials.

The integration of EDCs and EHRs relies heavily on the unification of terms. However, using the direct record type of reporter essentially creates a system closed to the EDC, so there is no need to match the term code of ItemData with the EMR side. In contrast, when using the data transcription type of reporter, the item code of the template in the progress note must be designated in the field of the template of the CRF reporter in order to retrieve data. We assume that a template for the progress note is additionally created for clinical research, and doctors purposely select that template if a patient is participating in a given study. It is therefore not absolutely necessary to unify item codes of both templates. However, if the CRF reporter would be set independently from the creation of the templates in the progress note, a universal keyword must be wrote as a concept code of a given element of the template in the progress note.

It has often been said that a unified vocabulary for medical care and clinical research is prerequisite for the interoperability of EHRs and EDCs [10;11]. However, compiling a complete dictionary of terms in both medical care and clinical research is an endless task. Thus, if the creation of this dictionary was premised, the integration of EHRs and EDCs would be

realizable. In our system, while establishing a list of unified keywords would be useful, it is not necessary. We feel that this element makes our system practical for clinical research support.

Because the CRF reporter is an independent conponent of the EMR, it can be incorporated in any EMR system. We have so far incorporated it in the EMR developed by NEC and another by IBM. With accordance of the communication procedure between CRF reporters and the receiver, a different type of CRF reporter comes into effect. Fujitsu is trying to develop another type of CRF reporter which can communicate with the same center server. By incorporating the CRF reporters in the EMRs in the hospitals, we are creating a hospital cluster for networked clinical research, which will consist of more than 16 hospitals in Osaka in the following years.

Conclusion

We developed a CRF reporter to integrate EHRs and EDCs. The CRF reporter incorporated in an EMR system can retrieve data recorded in the EMR, generate an ODM, and transmit the ODM to the CDMS. With the data transcription type of CRF reporter, doctors enter data using a template for the progress note, and the data are then transmitted to the template for the CRF reporter. In the direct record type of reporter, templates of the CRF reporter open from the progress note, generating narrative text to make a record in the progress note. This system can save both labor and financial costs in clinical research.

Acknowledgements

This research is supported by the Project of ICT infrastructure establishment for clinical and medical research from Japan Agency for Medical Research and Development, AMED.

References

[1] C. Ohmann and W. Kuchinke, Future developments of medical informatics from the viewpoint of networked clinical research. Interoperability and integration, *Methods Inf Med* **48(1)** (2009), 45-54.

[2] CDISC Electronic Source Data Interchange (eSDI) Group, Leveraging the CDISC standards to facilitate the use of electronic source data within clinical trials version, November 20, 2008. < http://www.cdisc.org/stuff/contentmgr/files/0/2f6eca8f0df7caac5bbd4fadfd76d575/miscdocs/esdi.pdf >

[3] Y. Matsumura, A. Hattori, S. Manabe, T. Takeda, D. Takahashi, Y. Yamamoto, T. Murata, and N. Mihara, Interconnection of Electronic Medical Record with Clinical Data Management System by CDISC ODM, *Stud Health Technol Inform* **205** (2014), 868-872.

[4] Clinical Data Interchange Standards Consortium (CDISC) , 15907 Two Rivers Cove, Austin, Texas, USA; ©2008. Standard available from: http://www.cdisc.org/models/odm/v1.3/index.html (cited Apr 20, 2009).

[5] Y. Matsumura, S. Kuwata, Y. Yamamoto, K. Izumi, Y. Okada, M. Hazumi, S. Yoshimoto, T. Mineno, M. Nagahama, A. Fujii, and H. Takeda, Template-based data entry for general description in medical records and data transfer to data warehouse for analysis, *Stud Health Technol Inform* **129** (Pt 1) (2007), 412-416.

[6] ITI Technical Committee. IHE IT Infrastructure Technical Framework Supplement Retrieve Form for Data Capture (RFD) Trial Implementation. August 19 (2011), <http://www.ihe.net/technical_framework/upload/ihe_iti_suppl_rfd_rev2-2_ti_2011-08-19.pdf>

[7] A. El Fadly, B. Rance, N. Lucas, C. Mead, G. Chatellier, P.Y. Lastic, M.C. Jaulent, and C. Daniel, Integrating clinical research with the Healthcare Enterprise: from the RE-USE project to the EHR4CR platform, *J Biomed Inform* Dec;**44** Suppl 1(2011), S94-102.

[8] Y. Matsumura, A. Hattori, S. Manabe, T. Tsuda, T. Takeda, K. Okada, T. Murata, and N. Mihara, A strategy for reusing the data of electronic medical record systems for clinical research, *Stud Health Technol Inform* **228** (2016), 297-301.

[9] J.D. Mestler, C. Celingant, S. Bishop, R. Perkins, and M. Rocca, EHR/CR Functional Profile Working Group. EHR/CR Functional Profile, Electronic Health Records/Clinical Research. Committee Level Ballot Release 1, March (2008), http://www.eurorec.org/files /filesPublic/ehrworkshop/2008/HL7%20EHRCR%20 Functional%20Profile% 20Rel%201.pdf

[10] W. Xu, Z. Guan, J. Sun, Z. Wang, and Y. Geng, Development of an open metadata schema for prospective clinical research (openPCR) in China, *Methods Inf Med* **53(1)** (2014), 39-46.

[11] M. Dugas, P. Neuhaus, A. Meidt, J. Doods, M. Storck, P. Bruland, and J. Varghese, Portal of medical data models: information infrastructure for medical research and healthcare, *Database (Oxford)* 2016 Feb 11, 1-9.

Address for correspondence

Yasushi Matsumura MD, PhD

2-15, Yamada-Oka, Suita, Osaka 565-0871, JAPAN

Phone: +81-6-6879-5900, Fax: +81-6-6879-5903

e-mail: matumura@hp-info.med.osaka-u.ac.jp