The Importance of Health Informatics in Public Health during a Pandemic
J. Mantas et al. (Eds.)

doi:10.3233/SHTI200517

Implementation of a Secured Cross-Institutional Data Collection Infrastructure by Applying HL7 FHIR on an Existing Distributed EMR Storages

Katsuya TANAKA [a,1] and Ryuichi YAMAMOTO [b]
[a] *IT Integration Support Center, National Cancer Center, Tokyo, Japan*
[b] *Medical Information System Development Center, Tokyo, Japan*

Abstract. A secure data collection infrastructure was implemented in which Health Level-7 (HL7) Fast Healthcare Interoperative Resources (FHIR) was applied according to standardized electronic medical record (EMR) storage. This infrastructure aims to facilitate rapid secondary use of EMR data in cross-institutional analyses on the basis of the Standardized Structured Medical Information eXchange (SS-MIX), Japan's domestic standard for EMR export. Existing EMR storage comprised enormous numbers of HL7 v2 messages; therefore, the user interface and database structure are basically defined according to the HL7 v2 data types and message structures, causing difficulty in rapid extraction for researches by clinicians. To solve this problem, we are trying to enhance some user interface with HL7 FHIR, which needs SS-MIX items to be mapped through HL7 FHIR resources. We investigated definition gaps between HL7 FHIR and SS-MIX messages, and report a developed user interface with HL7 FHIR to query against existing storage.

Keywords. Privacy, Data Collection, Electronic Medical Records, HL7 FHIR

1. Introduction

Secondary use of electronic medical records (EMR) items over a public cloud is common around the world. During the process of extraction, integration, and anonymization of such data, privacy invasion is not allowed. To promote secure use of EMR items over a public cloud environment, we previously developed a cross-institutional, secure data collection system [1]. Key requirements of this system are (1) secure integration of extracted EMR items, (2) standardized EMR data sources, (3) privacy risk assessment of extracted data set, and (4) traceability of data transfer for patients. As a data source, we adopted the Standardized Structured Medical Information eXchange (SS-MIX) [2], which is the domestic standard for exporting whole EMR data to external storage in Japan. SS-MIX2, which is the next version of SS-MIX, is also based on Health Level-7 (HL7) v2 message files. This standardized storage is widely used for backing up data, regional collaboration, disease registries, and other research purposes. In this standard, we can represent definite states of 37 clinical events for each patient in a hospital,

[1] Corresponding Author, Katsuya Tanaka, IT Integrating and Support Center, National Cancer Center, Tsukiji 5-1-1, Chuo-ku, Tokyo, 104-0045, Japan; E-mail: katstana@ncc.go.jp.

including admission–discharge history, medication history, and laboratory test results, using ADT/RDE/OUL HL7 v2 messages. The directory structure for stored records is based on patient ID, clinical date, and clinical event type. This structure hampers using exported data for cross-patient analyses such as epidemiological studies. Moreover, the data types and message structures are not familiar to clinical researchers, and some common model of sharing data is needed.

In view of this situation, we are trying to improve the user interface of the developed secure data collection system for querying EMR items in an SS-MIX2 storage. HL7 Fast Healthcare Interoperability Resources (FHIR) [3] is a strong candidate for building an intuitive user interface to handle the EMR items from existing EMR storage. However, at this time, sufficient materials for backward compatibility against HL7 v2 messages have not been provided, and it is difficult to apply already accumulated clinical information to this new standard. We describe the results of an investigation of EMR item mapping between SS-MIX2 and HL7 FHIR and the performance of the experimental system.

2. Methods

2.1. System Overview

Figure 1 depicts an overview of the developed system. EMR items can be exported to an SS-MIX2 storage as HL7 v2 message files according to a domestic standard. These message files are stored as blob records to the PostgreSQL v9.5 database system and converted into relational database tables in real time. By deploying a web interface that uses HL7 FHIR, queries can be processed using FHIR resources, and we tried to develop a query user interface for data collection. The web interface is implemented using HAPI FHIR Java library [4], which enables conversion of FHIR-based model objects and HL7 v2 message items.

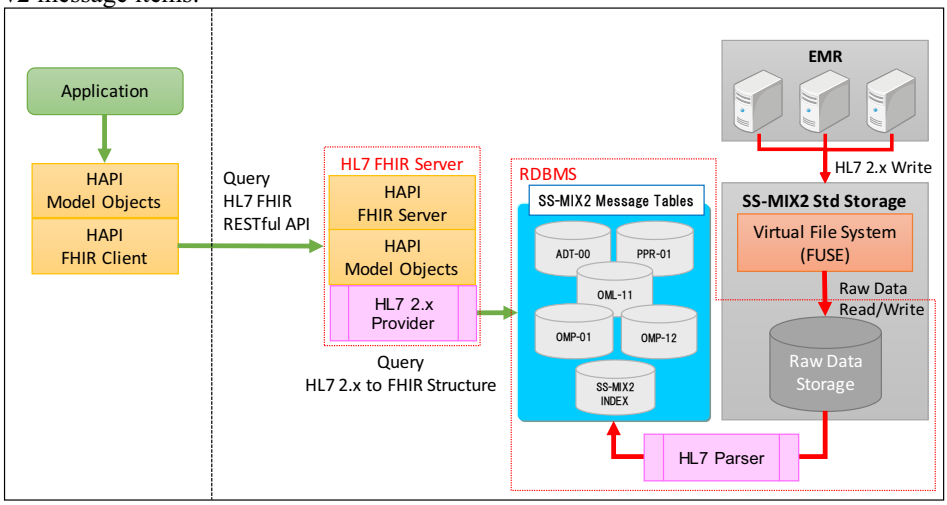

Figure 1. Overview of developed standardize storage using HL7 FHIR and SS-MIX2

Mapping EMR items between SS-MIX2 and HL7 FHIR for Intuitive User Interface HL7 FHIR is providing the resource item mappings for HL7 v2 messages on the HL7 US website. SS-MIX2 comprises 30+ HL7 v2 event messages, and the messages are

coded under the domestic standard. The FHIR mapping table has some deficiencies; some items need to be complemented by checking both coding definitions. We have performed item mapping complementation on the FHIR resources and we also developed a user interface for querying SS-MIX2 storage using HL7 FHIR resources according to the item mapping described above.

3. Results

3.1. Mapping EMR Items

There was little need to supplement the Observation resource. Table 1 shows the result of mapping ratio by FHIR resources. Although the mapping ratio scores are low, it was confirmed that this is partly due to the scattering of unused items on the SS-MIX2 side.

Table 1. Mapping Ratio of properties from FHIR Resources to SS-MIX2 Message Fields

No.	FHIR Resource	Property	Mapped	Ratio(%)
1	Patient	24	14	58%
2	Encounter	40	11	25%
3	Condition	20	12	60%
4	AllergyIntolerance	22	13	59%
5	Observation	33	20	60%
6	Specimen	27	7	25%
7	ServiceRequest	33	17	51%
8	MedicationRequest	40	20	50%
9	MedicationDispense	28	16	57%
10	Dosage	15	8	53%

3.2. User Interface

An intuitive user interface for querying existing SS-MIX2 standardized storage items have been developed. Query items can be chosen by FHIR resource vocabulary. The user interface has an area for specifying search conditions and output items, respectively.

3.3. Performance

Queries yield the expected response if the number of records is approximately 100,000. With a larger number (e.g., >10,000,000), the query response is poor owing to the database structure of the original SS-MIX2 message tables. This was caused by the database performance of joined SELECT across different segment tables.

4. Discussion

This study showed that the data extraction function for clinical research with an intuitive interface using HL7 FHIR vocabulary can be constructed with the standardized storage system widely deployed in Japan, indicating a cost-effective reconstruction of data

extraction or data exchange infrastructure without wasting previous capital investments for existing SS-MIX2 standardized storage items.
Herein, we conducted mapping of EMR items in only several HL7 FHIR resources. The remaining resources and HL7 v2 segments remain to be investigated. By nature, mapping differences persist because HL7 FHIR resources are defined according to HL7 v3 reference information model (RIM); thus, complementing the mapping gaps between HL7 v2 and v3 may remain. To achieve our minimum goal, we need to represent HL7 v2 messages as an FHIR resource, including domestic description standards and EMR item mappings for HL7 v2.

Improvement in database performance is needed when the number of data is large. Further, for our secure data collection, a system based on Bloom filter [1] needs elements of patient identification, such as insurance number, date of birth, and gender, to eliminate duplication of patient data in cross-institutional data extraction. We will continue to develop a patient deduplication mechanism. Moreover, since the developed correspondence table between SS-MIX2 and FHIR can be used for automatic generation of conversion modules, we started development of such a deployment mechanism.

5. Conclusions

This paper reports the initial performance results related to a user interface and query system with HL7 FHIR for secure data collection. We examined the mapping of EMR items between existing HL7 v2-based standardized EMR storage and HL7 FHIR definitions. We also proposed an implementation method for using existing HL7 v2–based storage and deploying HL7 FHIR message structures for the user interface.

Acknowledgments

This study has been approved by the Research Ethics Committee of National Cancer Center Japan (Permission number: 2019-122) and was supported by JST CREST Grant No. JPMJCR1404, Japan.

References

[1] Tanaka K, Yamamoto R, Nakasho K, Miyaji A, Development of a Secure Cross-Institutional Data Collection System Based on Distributed Standardized EMR Storage, Stud Health Technol Inform 255 (2018), 35-39.
[2] Kimura M, Nakayasu K, Ohshima Y, Fujita N, Nakashima N, Jozaki H, Numano T, Shimizu T, Shimomura M, Sasaki F, Fujiki T, Nakashima T, Toyoda K, Hoshi H, Sakusabe T, Naito Y, Kawaguchi K, Watanabe H, Tani S, SS-MIX: a ministry project to promote standardized healthcare information exchange, Methods Inf Med 50 (2011), 131-139.
[3] HL7.org. FHIR v4.0.0. Available: http://hl7.org/implement/standards/fhir/ (2019-09-28).
[4] HAPI FHIR - The Open Source FHIR API for Java. Available: http://hapifhir.io/ (2019-09-28).