A Study on Transformation and Utilization of Common Data Model for Data Analysis of CDA-Based External Referral Document

Hyerim Ji,MS[1], Seok Kim,MPH[1], Soyoung Yi,RN[1], Sooyoung Yoo, PhD[1]
[1] Seoul National University Bundang Hospital, Gyeonggi-do, Seongnam, republic of korea

Is this the first time you have submitted your work to be displayed at any OHDSI Symposium?

Yes __V__ No___

Abstract

The paper converted CDA documents into OMOP CDM, a common data model, to analyze written requests accumulated in a hospital. The converted data were analyzed and the quantity and quality of exchanged information was measured to expand the scope of data analysis and measures for their use were prepared.

Introduction

In the present study, HL7 Clinical document architecture (CDA) was converted to a common data model OMOP CDM. There has been a study that converted Continuity of Care Documents (CCD) to CDM in the US, but there has been no case of converting CDA documents created in South Korea to CDM [1]. Seoul National University Bundang Hospital built a healthcare information exchange system in 2007 and collected 21,580 CDA written requests until April 2018. As the potential advantage of healthcare information exchange is proportional to the quantity of data, the analysis of exchange information is necessary to understand the usefulness of healthcare information exchange. If an extensive number of CDA documents can be constructed into a common data model, various big data analysis research using the healthcare information exchange CDA data may be possible.

Method

1. CDA document collection and parsing

A total of 21,580 CDA written requests accumulated in the healthcare information exchange system were collected as analysis. Version information of various CDA written requests was determined by identifying standard update details of each version of CDA and the application of standard codes of each item [2]. Using data parsing technique, each version of CDA documents were parsed and loaded into a database. If documents needed to be separated because English alphabets and numbers are mixed, they were classified using a regular expression and then parsed.

2. CDAtoCDM

A conversion rule was defined to review the quality of CDA document data and the standard codes in the document and then map them into CDM standard code table. If mechanical mapping with standard codes is difficult, CDA document codes were standardized by manually reviewing the documents, and the standardization results were collected to create a CDAtoCDM master table. By creating the details of ETL (Extract, Transform, and Load) to define the rule of using the parsed CDA documents and the master table, logic for converting CDA documents to CDM was materialized. The CDM conversion algorithm was performed by creating and executing CDM conversion code according to ETL definition. The results of OMOP CDM table matching of CDA data are presented in Table 1.

Table 1. CDM Condition Occurrence Table ETL , a part of ETL

OMOP CDM Table	HL7 CDA Section or xpath
Care site	ClinicalDocument/custodian
Provider	ClinicalDocument/Authenticator
Person	ClinicalDocument/recordTarget/patientRole/patient
Visit occurrence	Document Header
Condition occurrence	Problem section
Drug exposure	Medication section
Measurement	laboratory data section
Procedure occurrence	Procedure section

Result

The results of CDAtoCDM conversion showed that 97.7% of the entire CDA data were converted into CDM data. The results of each CDM table are presented in Figure 1. The concept of each table is the mapping rate of standard code, and CDAtoCDM is the results of CDA parsing data mapped into CDM table. Except for duplicate data in the parsing results of CDA and the data that cannot be mapped into standard codes, mapping of all data were completed. In the case of measurement and procedure occurrence, medical staff manually wrote in the CDA of written request in many cases, and because it is a value entered without a code, the concept mapping rate measured was low. In the case of measurements, because only the data that have test results from CDA Lab results table were defined to be converted, data without test results were not converted.

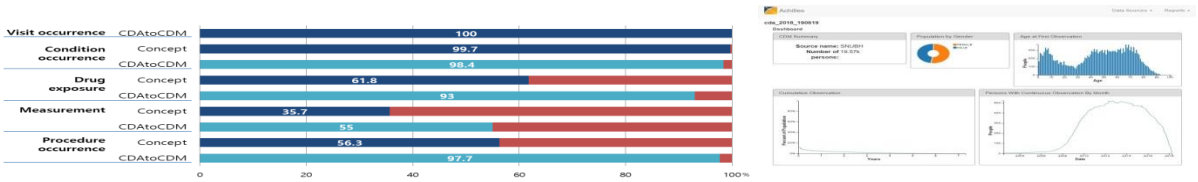

Figure 1. CDAtoCDM results and the results of Achilles application

In addition, the results of CDAtoCDM were applied to a visualization tool Achilles to ensure that the pattern, distribution, and trends can be easily identified. Figure 1 shows the Dashboard of CDAtoCDM data to which Achilles was applied and outputted.

Conclusion

The present study constructed a foundation for big data use of healthcare information exchange CDA document data. The results of the study that converted healthcare information exchange CDA into OMOP CDM can be used as basic data for researchers who want to conduct analysis using healthcare information exchange in future clinical observation studies.

This work was supported by grant no 18-2018-011 from the SNUBH Research Fund
No potential conflict of interest relevant to this article was reported

References

1. Hamed Abedtash, PharmD, Jon D. Duke, Assessing Mapping Performance from HL7 Continuity of Care Documents to the OMOP CDM. 2017 OHDSI Symposium
2. HL7 Clinical Document Architecture, Release 2, Journal of the American Medical Informatics Association 2006,13(1):30-9