MEDINFO 2019: Health and Wellbeing e-Networks for All
L. Ohno-Machado and B. Séroussi (Eds.)

doi:10.3233/SHTI190475

Expansion of EHR-Based Common Data Model (CDM)

Wona Choi[a], Soo Jeong Ko[a], Hyuck Jun Jung[a], Tong Min Kim[a], Inyoung Choi[a]

[a] *Department of Medical Informatics, College of Medicine, The Catholic University, Seoul, Republic of Korea,*

Abstract

We expanded and constructed a Common Data Model (CDM) based on hospital EHR to enable analysis and comparison of Adverse Drug Reactions(ADRs) integrated with external organizations with different data structures. This is significant in that it is possible to conduct joint research, analysis, and comparisons among institutions with the same type of CDM constructed, and provide the basis for conducting the same research simultaneously on various data sources.

Keywords:
Electronic Health Records, Observational Study, Database

Introduction

Traditionally, we have relied on the information obtained from clinical trials on drug side effects and post-marketing individual patient case reports to identify potential safety issues of marketed drugs [1; 2]. In fact, Korea has been surveilling drug side effects through the spontaneous ADRs reporting system.

However, in addition to the recognition that the spontaneous reporting system is not the only means that is effective in monitoring drug side effects [3; 4], and there is a growing interest in the monitoring of ADRs using hospital EHR.

Most hospitals in Korea use their own EHR system, and many Korean codes for diagnosis, medication, and treatment are not compatible with international coding systems [5]. Due to the differences in the data structure, format, and terminology used in these individual data sources, it takes a lot of time to analyze data and it is difficult to compare the results of studies using different databases [1].

The purpose of this study is to expand and build a common data model based on EHR to enable systematic analysis of different databases.

Methods

Data Source

The Catholic University of Korea Seoul St. Mary's Hospital is a tertiary university hospital with 1,356 beds and over 10,000 outpatient visitors per day. The hospital has a Clinical Data Warehouse (CDW) that extracted EHR data such as patient visits, examination, diagnosis, prescription and nursing needed for clinical research. We would like to convert the data stored in CDW of Seoul St. Mary's Hospital to CDM format.

K-CDM(Korean Common Data Model)

The Korean Common Data Model (K-CDM) has been defined in 2017 by the Ministry of Food and Drug Safety [6] through the optimization and localization of Observational Medical Outcomes Partnership Common Data Model (OMOP CDM) [7] and Sentinel Common Data Model (Sentinel CDM) [8]. K-CDM includes 8 tables (Person, Visit, Procedure, Condition, Observation, Drug, Measurement, Vital signs).

Code Mapping

The in-hospital codes in the EHR are mapped according to the international standard terminology system.

For the 584 drugs in the hospital, containing the designated 288 components, the component and in-hospital code were mapped at 1:N. The items including combination in the same component name were mapped as RxNorm and ATC by separating the single agent and the complexing agent. All in-hospital diagnosis codes for diagnosis and treatment and the in-hospital codes for 138 designated tests were mapped at 1:N to SNOMED-CT and LOINC, respectively.

Extraction, Transformation, Loading (ETL) Process

The ETL process involves the entire process of taking data from a database system and moving to another database system[5]. We created SQL statements for each table in the K-CDM and extracted, transformed, and loaded the patient, medication, diagnosis(condition), and test information(Measurement) according to the K-CDM format from the CDW(*figure 1*).

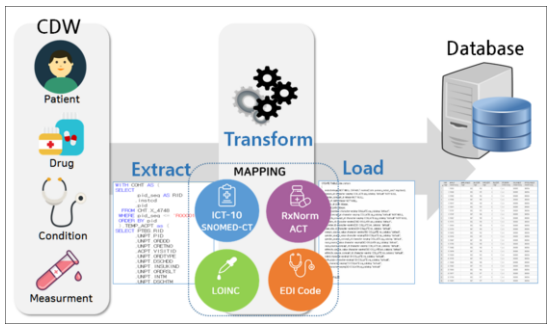

Figure 1 – ETL Process

Database

We used PostgreSQL, an open source object-relational database management system (ORDBMS), which is used as a medical information database to load data [9].

Data Analysis

Allopurinol is a kind of xanthine oxidase inhibitor that is commonly used as a preventive and therapeutic agent in

diseases that uric acid is excessively produced such as gout, uric acid nephrolithiasis, intracellular erythropoiesis, chemotherapy of lymphoma, Lesch-Nyhan Syndrome, chronic renal failure [10; 11].

Based on pervious study [12], data extraction and analysis were performed to confirm the association of allopurinol and TSH levels, which were not reflected in the Korean authorization.

Results

We extracted EHR data of 3,212,915 patients from Seoul St. Mary's Hospital from January 1, 1997 to December 31, 2017, transformed them into K-CDM format and completed the loading. The items used with in-hospital codes and domestic codes were mapped with the international standard terms, and data quality was verified using ATLAS and ACHILLES.

A Patients group (5,148 people) and a control group (20,581 people) were extracted from the established DB.

Table 1 – Number of Data

	Patient Group	Control Group
Number of group	5,148 (100%)	20,581 (100%)
Male	1,845 (35.8%)	7,377 (35.8%)
Female	3,303 (64.2%)	13,204 (64.2%)
Taking allopurinol		
Yes	172 (1.5%)	207 (1.0%)
No	4,976 (96.7%)	20,374 (99.0%)
Age-yr	56.06 ± 14.15	55.88 ± 13.93

Conditional logistic regression model was used to analyze allopurinol and the risk of elevated TSH levels (odds ratio, OR). As a result, the odds ratio was 3.46 times (95% CI: 2.91-4.26).

Discussion & Conclusion

We expanded and constructed the common data model based on hospital EHR to enable analysis and comparison of adverse drug reactions integrated with external organizations without releasing hospital EHR data to outside. This is significant in that it is possible to conduct joint research, analysis, and comparisons among institutions with the same type of CDM constructed, and perform the same research on various data sources simultaneously.

In fact, a multicenter study has been conducted on the subject of our study (risk of elevated thyroid stimulating hormone levels in allopurinol), and we hope to have meaningful results.

Our study has the following limitations. By applying a domestic common data model named K-CDM, part of OMOP CDM format was applied, which is widely used internationally. This can lead to limitations in international research, and supplementation is necessary in the future. In addition, governance for international standard medical terminology should be actively and specifically made at the national level, and many studies using medical data that are being accumulated infinitely in each hospital should be conducted in the future.

Acknowledgements

This research was supported by a grant from Korea Institute of Drug Safety and Risk Management in 2018.

References

[1] S.J. Reisinger, P.B. Ryan, D.J. O'hara, G.E. Powell, J.L. Painter, E.N. Pattishall, and J.A. Morris, Development and evaluation of a common data model enabling active drug safety surveillance using disparate healthcare databases, *Journal of the American Medical Informatics Association* **17** (2010), 652-662.

[2] Managing the risks from medical product use. Creating a risk management framework, in: *Report to the FDA Commissioner from the Task Force on Risk Management Rockville, MD*, 1999.

[3] R.S. Evans, S.L. Pestotnik, D.C. Classen, S. Bass, and J. Burke, Prevention of adverse drug events through computerized surveillance, in: *Proceedings of the Annual Symposium on Computer Application in Medical Care*, American Medical Informatics Association, 1992, p. 437.

[4] A.K. Jha, G.J. Kuperman, E. Rittenberg, J.M. Teich, and D.W. Bates, Identifying hospital admissions due to adverse drug events using a computer-based monitor, *Pharmacoepidemiology and Drug Safety* **10** (2001), 113-119.

[5] D. Yoon, E.K. Ahn, M.Y. Park, S.Y. Cho, P. Ryan, M.J. Schuemie, D. Shin, H. Park, and R.W.J.H.i.r. Park, Conversion and data quality assessment of electronic health record data at a Korean tertiary teaching hospital to a common data model for distributed network research, **22** (2016), 54-58.

[6] Development of a common data model for adverse drug reactions using hospital EMR, in, Ministry of Food and Drug Safety, Korea, 2017.

[7] https://www.ohdsi.org/data-standardization [accessed on March 22, 2019]

[8] https://www.sentinelinitiative.org/sentinel/data/distributed-database-common-data-model [accessed on March 22, 2019]

[9] https://www.postgresql.org/, postgreSQL [accessed on March 22, 2019]

[10] J.S. Cameron and H.A.J.B.m.j. Simmonds, Use and abuse of allopurinol, **294** (1987), 1504.

[11] S.H.S. Jung Hee Kim, Dong Won Lee, Soo Bong Lee, Ihm Soo Kwak, Case Reports : A Case of Allopurinol Hypersensitivity Syndrome with Esophageal Ulcer Bleeding, *Kidney Res Clin Pract* **25** (2006), 669-673.

[12] F. Perez-Ruiz, S.P. Chinchilla, J. Atxotegi, I. Urionagüena, A.M. Herrero-Beites, and M.A.J.R.i. Aniel-Quiroga, Increase in thyroid stimulating hormone levels in patients with gout treated with inhibitors of xanthine oxidoreductase, **35** (2015), 1857-1861.

Address for correspondence

Inyoung Choi, PhD, Assistant Professor

222, Banpo-daero, Seocho-gu, Seoul, 06591, Republic of Korea

Department of Medical Informatics, College of Medicine, The Catholic University, Seoul, Republic of Korea, iychoi@catholic.ac.kr