Name:	Seongwon Lee
Affiliation:	School of Medicine, Ajou University
Email:	seongwon.lee.16@gmail.com
Presentation type (select one):	Poster

Conversion of National Health Insurance Service (NHIS) Data of Korea to the Observational Medical Outcomes Partnership (OMOP) Common Data Model

Seongwon Lee, Ph.D[1], Soo-Yeon Cho, R.N., MPH.[1], Seng Chan You, M.D., MPH.[1], Hojun Park, BS.[1], Sungjae Jung, BS.[1], Rae Woong Park, M.D., Ph.D[1,3],Yunyoung Bae, MPH.[2], Hangil Lee, MPH.[2], Jahyun Cho, MPH.[2], Keunhui Park, MS.[2]

[1]Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, Korea; [2]Department of Insurance Benefits, National Health Insurance Service, Wonju, Korea; [3]BrainKorea21 (BK21) Plus, Korea

Abstract

Korea adopts the mandatory national health insurance system and National Health Insurance Service (NHIS) manage it. We are converting the nation-wide claim data of NHIS into Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) v.5 and we will apply international standard codes, perform ETL process in two phases, and validate converted data by replicating previously proven studies. We expect that CDM version of NHIS data will make various types of studies possible and enhance the usability and accessibility of data.

Introduction

The National Health Insurance Service (NHIS) holds the database containing demographic information, diagnosis, prescription from clinics and/or hospitals, ambulatory health care utilization, hospitalization, and health outcomes for nearly all Koreans. With the purpose of enhancing rapid research for potential adverse drug event and comparison of the results from various databases across the world, NHIS attempted to convert the Korean national claim data into the OMOP CDM v5. In this paper, we present the work process and validation methods for the conversion of national health insurance claim database into OMOP CDM.

Method

Data Source

The National Health Insurance Service (NHIS) was started in 1977 and achieved universal coverage by 1989 in Korea. Currently, the NHIS maintains and manages all databases for Korea's health services to utilize. In July 2014, NHIS started the National Health Insurance Sharing Service (NHISS) which is the fee-based claim data providing service. Five-year (from 2011 to 2015) claim data of NHIS will be converted into OMOP CDM v5. For further research of pharmacovigilance, we set the priorities of tables to convert as following: PERSON, OBSERVATION_PERIOD, OBSERVATION, DEATH, VISIT_OCCURRENCE, DRUG_EXPOSURE, CONDITION_OCCURRENCE, MEASUREMENT, CONDITION_ERA, DRUG_ERA, LOCATION, and CARE_SITE.

Application of Standard Code

The NHIS uses the 6th Korean Standard Classification of Diseases (KCD-6) code system for diagnosis and EDI (electronic document interchange) code system for the drug, procedure and measurement data. In converting process, the Korean standard terminologies will be mapped into OMOP vocabulary. We extracted the Korean code list which had been used for last 10 years (from 2006 to 2015) and plan to map the codes at 95% level in order of the amount of used. The overall process of conversion will be performed through two steps: Firstly, Standard Code Mapping Dictionary will be applied for NHIS database, which was developed by Ajou University [1][5]. At the second step, we will investigate the unmapped codes and map them on international standard codes. By doing so, we expect that we will be able to customize the NHIS code mapping and elevate the quality of our mapping dictionary.

ETL (Extraction, Transformation, and Loading) Process

Figure 1 depicts how ETL process will be performed through two phases and two versions of ETL documents and MS-SQL query scripts will be generated. At the first phase, we will make ETL document and scripts for the ETL from original NHISS to CDM. Original NHISS is the data-set to be provided to researchers, so NHISS is relatively refined, clean, and validated unproblematic in terms of personal information protection principles. This phase makes whole ETL process easy by using well-known data-set and will enhance the usability of NHISS data as well. And at the

second phase, we will create ETL documents and query scripts and perform actual ETL process from NHIS raw data to CDM.
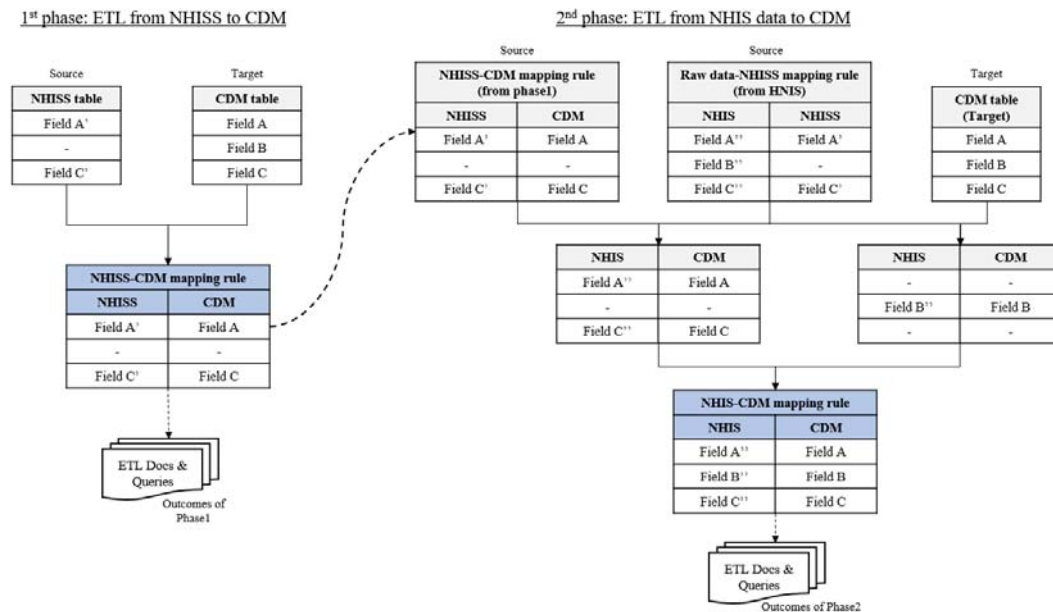

Figure 1. Whole ETL Process.

Data Validation

Replication of prior study is planned to confirm validity and feasibility of converted CDM structure of NHIS data [4]. Hence, the two same analyses will be performed, with original NHIS database and with converted CDM database, respectively. The result and consistency will be compared between these two analyses. Also, we will examine the length and complexity of analytic scripts of both to survey the easiness of research with two forms of data.

Conclusion

We are converting nation-wide claim data of Korean NHIS into OMOP CDM v5 with well-organized procedure. We expect that CDM version of NHIS data will make various types of studies (e.g., comparative study) possible and enhance the usability and accessibility of data across the world. In the further study, results of converting process and validation will be presented.

Funding Source

Acknowledgement

We have no conflict of interest.

References

1. Cho SY, Yoon D, Ahn EK, Shin D, Park RW. Mapping Korean national insurance billing code to OMOP code for drugs used in a Korean tertiary teaching hospital. OHDSI Symposium 2015.
2. National Health Insurance Service. http://www.nhis.or.kr. 2016.
3. National Health Insurance Sharing Service. https://nhiss.nhis.or.kr. 2016.
4. Rupa M, Ryan PB. Transforming the Premier Perspective® hospital database into the Observational Medical Outcomes Partnership (OMOP) Common Data Model. eGEMs 2014; 2(1): 1-10.
5. Yoon D, Ahn EK, Park MY, Cho SY, Ryan PB, Schuemie, MJ, Shin D, Park H, Park RW. 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. Healthcare Informatics Research 2016; 22(1): 54-58.