Name:	Clair Blacketer
Affiliation:	Janssen R&D
Email:	mblacke@its.jnj.com
Presentation type (s):	Poster

Creating a Framework for Evaluating Open Healthcare Claims Using the OMOP Common Data Model

Clair Blacketer, MPH[1,2], Erica A Voss, MPH[1,2]

[1]Janssen Research and Development, Titusville, NJ; [2]OHDSI collaborators, Observational Health Data Sciences and Informatics (OHDSI), New York, NY

Abstract

Clearinghouses are used by US healthcare providers for "claims scrubbing" between the point of care and the insurance company by inspecting the claims for errors before submission for payment. This poster will describe a methodology for evaluating open claims data by first obtaining the cross-section of patients from a claims clearinghouse and from a prescription benefit plan, defining metrics of interest, and then transforming the data to the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) (1) in such a way to quantify those metrics.

Introduction

In the US, claims clearinghouses provide transactional support between providers and payers to facilitate fast payment by minimizing errors during the claims submission process(2). The claims available from these institutions are considered "open" in that they contain the same information as data available from large payers except for final adjudication (considered as "closed"). While rich in information, this type of data has not been utilized for observational research because, unlike more traditional closed claims systems like such as prescription benefit or health insurance plans, members are enrolled at the provider level rather than at the patient level. To this point, there has not been a clear way to assess the completeness and validity of these claims. By first creating metrics to evaluate clearinghouse data, and then translating the data into the OMOP CDM to facilitate the calculation of those metrics, we created an open claims evaluation framework. While the CDM has been commonly used as an effective way to standardize observational health databases, it has yet to be employed in a way that enables assessment of the data itself for use in epidemiological research(3).

Methods

Medical and pharmacy data from a large claims clearinghouse was obtained and, to serve as a control, pharmacy data from a prescription benefits plan was linked at the patient level using protected health information (PHI). Non-linked commercial medical claims were also obtained from Truven Health MarketScan® Commercial Claims and Encounters Database for use as a high-level comparison.

Metrics

To understand the amount of information possibly missing from open claims and to understand how it compares to a traditional closed claims system the following metrics were chosen:

- Measure the overlap between the open claims and the closed claims for prescription drugs occurring per person per day during the same time period to quantify the missing data from the open claims that would be provided by the closed system and vice-versa.

- Compare the proportion of drug exposures at the ingredient level among persons in both the open and closed systems during the same time period to identify drugs that occur more often in one database versus the other.
- For a set of defined new user drug cohorts, compare the baseline prevalence of conditions between the open claims and Truven Health MarketScan® Commercial Claims and Encounters Database to assess if open claims capture the same proportion of conditions during the time prior to drug exposure.

CDM Conversion

To enable calculation of the metrics defined above, the intersection of patients between an open claims database and a closed claims database was de-identified and then transformed into the OMOP CDM. Unlike many extract, transform, load (ETL) processes that have been described previously(3), the raw dataset contained information from two separate sources which necessitated designing the tables such a way so that each source was still identifiable after transformation so that the metrics defined above could be quantified. This included:

- Using *_TYPE_CONCEPT_IDs as proxies for each database. For example, in the DRUG_EXPOSURE table the closed claims were given DRUG_TYPE_CONCEPT_ID = 38000177 and the open claims were given DRUG_TYPE_CONCEPT_ID = 38000175.
- Using the PAYER_PLAN_PERIOD and OBSERVATION_PERIOD tables to capture the amount of time contributed per person by each database. In this instance, the enrollment file from the prescription benefits closed claims system was used to create the PAYER_PLAN_PERIOD table, collapsing gaps of 32 days or fewer. The OBSERVATION_PERIOD table was created from the open claims by setting OBSERVATION_PERIOD_START_DATE as the earliest date available and OBSERVATION_PERIOD_END_DATE as the latest date available, for each patient. The two tables could then be used to find the overlapping time periods per person between the two source databases.

Conclusion

The use of open claims in observational health research is still a debated topic though little is understood about how well they represent the true picture of healthcare utilization for a given set of patients. By defining a set of metrics and designing an ETL in a way that allows for the assessment of those metrics, the OMOP CDM can be leveraged to create a framework that makes the evaluation of such data from an open system feasible.

References

1. OMOP Common Data Model [Webpage]. 2015 [cited 20 Jul 2015]. Available from: http://www.ohdsi.org/data-standardization/the-common-data-model/.
2. Sculley J. What is a clearinghouse for medical claims, and what do they do? 2014.
3. Voss E, Makadia R, Matcho A, Ma Q, Knoll C, Schuemie M, et al. Feasibility and utility of applications of the common data model to multiple, disparate observational health databases. Journal of the American Medical Informatics Association. 2015 2015 May;22(3):553-64. Epub 2015 Feb 10.