Contribution ID: **312** Type: **Poster**

Evaluating the impact of OMOP-CDM on data quality insight generation in respiratory disease management

Monday 13 October 2025 18:00 (1h 30m)

The increasing volume and heterogeneity of patient care data present significant challenges for comprehensive analysis and the generation of insights, particularly in specific disease areas such as respiratory diseases. Standardising diverse health data is crucial for enabling large-scale observational research and ensuring data readiness. The Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) provides a widely adopted standard for harmonising such data. However, evaluating the quality of data transformed into the OMOP CDM format is a critical step before its use in research or clinical decision support.

This study evaluates the impact of the OMOP CDM standardisation process on generating data quality insights for a respiratory disease dataset. The source dataset, initially in a paper-based format, was first converted to an electronic format. This historical dataset covers the years 2009 to 2024, containing 64 variables and 2,153 records.

The data underwent the standard Extract, Transform, and Load (ETL) process to convert it into the OMOP CDM format. Following this transformation, the quality of the resulting OMOP CDM instance was rigorously assessed. We utilised the Achilles tool, part of the OHDSI suite, designed for evaluating the quality of OMOP CDM databases. Achilles performs validation checks on the data based on key data quality dimensions, including conformance (adherence to standards), completeness (presence of values), and plausibility (believability of values).

The application of the OMOP CDM transformation and the subsequent quality assessment using Achilles successfully generated detailed insights into the dataset's quality. This systematic evaluation facilitated the identification of specific data quality issues across the conformance, completeness, and plausibility dimensions. Overall, the assessment conducted a total of 2,344 checks, of which 2,269 passed and 75 failed, resulting in a 97% overall pass rate for the Respiratory Diseases Inpatients data. It's also noted that 1439 out of 2269 passed checks were deemed 'Not Applicable' due to empty tables or fields, and 39 of the 75 failed checks were caused by SQL errors.

The standardisation of respiratory disease data using the OMOP CDM enabled a structured and transparent evaluation of data quality. Through the application of the Achilles tool, this study demonstrated the utility of OMOP CDM in generating meaningful data quality insights across multiple dimensions. These findings highlight the model's potential to enhance data readiness and support evidence-based decision-making in respiratory disease management.

Primary authors: KIRAGGA, Agnes (African Population And Health Research Center); MBATCHOU NGAHANE, Bertrand Hugo (Douala General Hospital, Data Science Without Borders project, Cameroon); MBOUAMBA YANKAM, Brenda (Douala General Hospital, Data Science Without Borders project, Cameroon); ANICET ONANA AKOA, François (Douala General Hospital, Data Science Without Borders project, Cameroon); BLAISE EBIMBE, Jean (Douala Gynaeco-Obstetric and Pediatric Hospital (DGOPEH), Cameroun); FANKOUA TCHAPTCHET, Luc Baudoin (DSWB); BARASA, Miranda (African Population and Health Research Center); ANDESO, Pauline (African Population And Health Research Center); IDDI, Samuel (African Population and Health Research Center)

Presenters: MBOUAMBA YANKAM, Brenda (Douala General Hospital, Data Science Without Borders project, Cameroon); ANICET ONANA AKOA, François (Douala General Hospital, Data Science Without Borders project, Cameroon); FANKOUA TCHAPTCHET, Luc Baudoin (DSWB); BARASA, Miranda (African Population and Health Research Center); IDDI, Samuel (African Population and Health Research Center)

Session Classification: Poster Session

Track Classification: SciDataCon Persistent Themes: Data and Research