**Reimagining Quality Reporting using the OMOP CDM NOTE\_NLP Table: A Pilot**

**Kristin Kostka, MPH1,2, Simon Beaulah, MSc1,4, Elizabeth Marshall, MD, MBA1,4;Benjamin N. Hamlin, MPH1,3**

**1Observational Health Data Sciences and Informatics, New York, NY, USA; 2IQVIA, Durham, NC, USA; 3National Committee for Quality Assurance, Washington, DC, USA; 4Linguamatics, an IQVIA Company, Marlborough, MA, USA**

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

 **Yes \_\_\_\_\_ No\_\_X\_\_\_\_**

**Abstract**

*HEDIS is the gold standard in healthcare performance measures. However, manual chart review is a substantial investment with an estimated $15.4 billion spent annually to report quality measures. We explored semi-automated methods to extract HEDIS and produce outputs in the OMOP CDM format. The NOTE\_NLP table can provide a home for HEDIS extractions that also could enable reuse of these data for other research purposes. Future work is needed to develop a fully validated pipeline in which full measure reporting can be constructed using the OMOP vocabularies and CDM. Ultimately any methods which utilize standards to improve the efficiency of HEDIS reporting could save US health plans millions of dollars in worker efficiency.*

**Background**

The Healthcare Effectiveness Data and Information Set (HEDIS®) is the gold standard in healthcare performance measurement, used by more than 90 percent of the nation's health plans and many leading employers and regulators (1). 190 million people are enrolled in plans that report HEDIS results (2). HEDIS includes more than 90 measures across 6 domains of care: Effectiveness of Care, Access/Availability of Care, Experience of Care, Utilization and Risk Adjusted Utilization, Health Plan Descriptive Information, and Measures Collected Using Electronic Clinical Data Systems. The National Committee for Quality Assurance (NCQA) collects HEDIS data from health plans, health care organizations and government agencies. HEDIS data help calculate national performance statistics and benchmarks, and set standards for measures in NCQA Accreditation. HEDIS measures are also included in the CMS Quality Rating System (QRS) which assist in determining appropriate payments made to Medicare Advantage plans (2).

**Methods: Assembling Measures**

HEDIS is a retrospective process. Each year health plans create targeted lists of patients for specific measures. The process regularly involves chart chasing – the process of hiring a medical professional to manually extract information from clinician notes in patient records to capture care delivered but not otherwise coded for. Manual chart review is a substantial time and monetary investment with more than 15.1 hours per week and an estimated $15.4 billion spent annually to report quality measures (4). As such, quality improvement teams are regularly investigating ways to utilize techniques such as Natural Language Processing to mine unstructured data and expedite the retrieval of non-coded information.

**Our Pilot: Breast Cancer and Colonoscopy Screening Data**

We evaluated a semi-automated approach to extract data to report out on two different HEDIS measures, Breast Cancer Screening and Colonscopy Screening, for a small set of selected patients. We used a proprietary NLP algorithm to extract data from a subset of selected patient charts. Each algorithm was validated in a prior exercise with a trained medical professional.



**Figure 1. Extract of NLP Data for Breast Cancer Screening Measure**

The resulting extracted data set was then mapped to the OMOP NOTE\_NLP table. As part of this pilot, we had other clinical data already loaded in the OMOP CDM for these patients. We performed an increment load to update the NOTE\_NLP table for these selected patients. The resulting OMOP data mart contains an amended view of the patient in which existing clinical records are stored simultaneously alongside NOTE\_NLP data. The intent of developing this ETL pipeline is to augment coded data with other unstructured data from the patient’s record in a systemized way eliminating the chart chase burden and standardizing the output.

**Conclusion**

We have only skimmed the surface on beginning to reimagine the pipeline in how HEDIS supplemental data can be extracted from clinical records with greater efficiency. Our project is a precursor to a hackathon session in which developers will brainstorm the utility of the ETL pipeline and discuss the next steps for developing a standard reporting pipeline. This process is an ongoing work and will be updated at the time of presentation to reflect further advancements made on the development of this use case. Our ultimate goal is to find ways to remove the burden of HEDIS measure reporting and enable streamlined, systemized methods to extract and build HEDIS measure reports. When stored efficiently alongside other clinical data, HEDIS measures represent highly valid cohorts which may be useful for other research purposes. More discussion is needed on the ways to modernize current reporting processes to take advantage of the OMOP vocabularies and CDM structure.

**References**

1. HEDIS. [Internet]. NCQA > HEDIS Publications: Overview - HEDIS & Quality Measurement. [cited 2019Jun24]. Available from: <http://store.ncqa.org/index.php/performance-measurement/hedis-publications-outline.html#vol1>
2. HEDIS [Internet]. NCQA. [cited 2019Jun24]. Available from: <https://www.ncqa.org/hedis/>
3. How to Maximize Incentives While Improving Outcome Measures [Internet]. SCPHCA Presentation. Absolute Total Care; 2012. Available from: http://www.scphca.org/media/34442/scphca\_presentation\_060912.pdf
4. Casalino LP, Gans D, Weber R, Cea M, Tuchovsky A, Bishop TF, et al. US Physician Practices Spend More Than $15.4 Billion Annually To Report Quality Measures. Health Affairs. 2016;35(3):401–6.