Database A
System A Data Warehouse
Only pulls data for structured variables available in CCD and not from medical transcriptions

Database B
HIE Data Warehouse
Uses NLP to pull data from medical transcriptions when CCD containing structured variables is not available

Patient aged 18-85 years with ≥1 visit at a System A healthcare facility in 2015

1) Healthcare facility that shares data only with System A data warehouse
2) Healthcare facility that shares data with both data warehouses
3) Healthcare facility that shares data with only the HIE data warehouse

A System A patient has an encounter at one of three categories of healthcare facilities

EHR encounter data sent as CCD and/or transcription to either one or both of the data warehouses

a All Healthcare System A facilities are under category 2, but each data warehouse receives data from System A patients from a variety of healthcare systems and healthcare facilities.

CCD: Continuity of care document, EHR: Electronic health record, HIE: Health information exchange, NLP: Natural language processing,