Simplifying Research that Involves Multiple Care Sites: insights and implementation at the US VA

Richard D. Boyce[1], Patrick R. Alba[1], Katherine R. Simon[1], Benjamin Viernes[1], William J. Obrien[1], Marc Suchard[1], Michael E. Matheny[1]

1. VA Informatics and Computing Infrastructure (VINCI), VA Salt Lake City Health Care System, Salt Lake City, UT, USA,

Background

The Veterans Health Administration (VHA) is the largest integrated health network in the United States, serving over 9 million patients at over 1,300 care facilities. To help manage information about VHA care sites, the health system maintains the detailed VA Site Tracking (VAST) database. Patient data going back to the 1990s and present in the Corporate Data Warehouse (CDW) is linked to information in the VAST database using station codes to identify geographic locations of care. In transforming and loading CDW data into the OMOP Common Data Model (CDM), the same data are represented by facts relating one care site to another. While this approach enables researchers to analyze VHA data in the CDM using various care site facets, such as site location, researchers can only do so using SQL queries over the CDM rather than OHDSI tools such as Atlas and HADES. Moreover, the sole use of fact relationships obscures the hierarchical organization of VHA care sites, making it more difficult than necessary for researchers to conduct research focused on care sites within a given region or serviced by a specific health care center or integrated service network. This project advances the representation of care sites in the CDM to simplify research that involves multiple care sites at the VA and delivers more general applicability for geo-coded research across the OHDSI community.

Methods

We collected use cases from VHA researchers that require the use of VA Care Sites. We then studied the VAST and developed both vocabulary and CDM extensions to enable fulfillment of the use cases, seeking to reduce cognitive burden and enable the use of OHDSI tools such as Atlas and HADES.

Results

Work is ongoing. Currently, a new vocabulary extension explicitly captures the hierarchical organization of VHA care sites: Level 1: national or one of 4 geographic regions and National; Level 2: one of eighteen veteran's integrated service networks (VISN); Level 3: regional benefit offices, VHA medical centers, or health care centers; Level 4: individual care institutions. Since Station Number Identifiers (sta3n) are used by VHA directive as the primary identification for care sites at Level 3 (with suffixes added for sites at Level 4), we concatenated sta3n codes with the site names to enable their use in concept set creation. We also represented the various site levels using the concept_class_id field and built records for the CONCEPT_RELATIONSHIP and CONCEPT_ANCESTOR table to simplify the grouping and filtering of VHA care sites during research cohort creation. To enable use in Atlas and HADES, we also modified the two tables in the CDM v5.4, LOCATION and

LOCATION_HISTORY, to add fields that enable historical locations records for both persons and care sites in VHA. We developed a preliminary implementation within the Circe back-end that processes cohort query criteria against the CDM extension using our vocabulary and modified Atlas for direct researcher engagement with the implementation.