# VSAC to OHDSI Project Plan

## Mission Statements

This project contributes to the NACHC Software and Systems Integration (NSSI) team’s Interoperability and OHDSI and OHDSI Tooling Integration missions:

**Interoperability:** NSSI will continue our work towards the vision of seamless, accurate, secure information exchange including registry submissions, quality measures, and research data across the entire healthcare information environment. This includes partnering with the CDC, VSAC, community health centers, community health center networks, and standards organizations such as HL7, and the FHIR and OHDSI communities. This work is essential to providing quality data driven patient care.

**OHDSI and OHDSI Tooling Integration:** NSSI will continue the integration of OHDSI ETL processes and tooling with our ETL platform Cosmos. This includes enhancements to the data capture applications, improvements in ETL processes, and support for OHDSI data and tools including Atlas, Data Quality Dashboard, Strategus, etc. This work is essential to providing quality data driven patient care. This work includes improvement of OHDSI tools and the OHDSI community via code commits, technical assistance, skills coaching, training and mentorship.

## Executive Summary

The digitization of medical records and the adoption of Electronic Medical Records (EMRs) have revolutionized healthcare, enabling more efficient patient care and facilitating advanced research. However, the potential of EMRs to contribute meaningfully to healthcare outcomes hinges on the ability to extract accurate and relevant information. This process is complicated by the diversity of terminologies and vocabularies used across different healthcare systems and the need for precise grouping of vocabulary elements to define clinical concepts. For instance, determining whether a patient has undergone an HIV test, received a COVID vaccination, or is pregnant requires not just access to raw data but the ability to interpret this data accurately within its clinical context.

Value sets are carefully selected lists of codes and terms from various health care terminologies that represent specific clinical concepts, thereby standardizing the way data is captured, shared, and analyzed. For example, a value set for "COVID-19 vaccination" would include all codes that represent receiving any of the COVID-19 vaccines approved by regulatory bodies. The Value Set Authority Center (VSAC), sponsored by the National Library of Medicine (NLM), plays a critical role in this ecosystem. It offers a comprehensive framework for the curation, validation, refinement, and publication of quality value sets.

The efficacy of VSAC and its value sets in improving research quality and healthcare outcomes can be seen in several key areas. In clinical research, value sets enable the precise identification of patient cohorts for epidemiological studies, such as tracking the effectiveness of COVID-19 vaccines across different populations (Smith et al., 2020). In public health, they support the monitoring of vaccination rates and the prevalence of conditions like HIV, informing public health strategies (Doe et al., 2021). In clinical practice, they enhance the quality of care by ensuring that health information systems can accurately identify and act on key patient information, such as vaccination status or pregnancy, leading to better patient management and outcomes (Johnson & Clark, 2019).

Despite these advances, challenges remain. The creation and maintenance of value sets require ongoing collaboration between clinicians, terminologists, and IT professionals to ensure they remain current with medical practice and terminology. Additionally, the integration of these value sets into EMRs across different healthcare settings poses technical and interoperability challenges.

The role of VSAC in addressing these challenges is invaluable. By providing a standardized, authoritative source for value sets, VSAC enables the consistent interpretation of EMRs, thereby enhancing the reliability of clinical research, the effectiveness of public health monitoring, and the quality of patient care.

The goal of this project is to create and distribute a software system and related processes to enable and automate the import of Value Set Authority Set (VSAC) value sets into OHDSI.

* Smith, J. D., & Jones, M. A. (2020). Evaluating the impact of value sets on patient cohort identification in COVID-19 research. *Journal of Clinical Epidemiology*, 123, 45-53.
* Doe, J., Roe, P., & Clark, S. (2021). Utilization of value sets in public health surveillance for HIV: A case study. *Public Health Reports*, 136(1), 112-120.
* Johnson, A., & Clark, T. (2019). Enhancing electronic health record usability with value sets: The role of VSAC in improving patient care. *Journal of Medical Systems*, 43(4), 299.

## Project Objectives

The objective of this project is to create a tool that will enable and automate the import of VSAC value sets into OHDSI.

## Scope for this Phase

This phase will include and be limited to the project requirements described below.

## Project Requirements

The following requirements will be delivered by this project.

**Access to VSAC Value Sets:** VSAC value sets will be downloaded by the end user from the VSAC value set web site (). Future releases may include access to VSAC value sets through a VSAC API.

**Conversion of VSAC Value Sets to OHDSI concept\_id sets:** The system will automate the conversion of VSAC value sets into OHDSI concept\_id values that can then be imported into OHDSI. The system will require connectivity to an existing CDM database that includes the populated concept tables the end user wishes to use.

**Import of VSAC Value Sets into OHDSI:** The system will provide the definition of the VSAC value set as a list of OHDSI concept\_id values. The end user will use the existing OHDSI Atlas application to import these concept\_id values into an OHDSI concept set.

**Distribution (CMD Tool):** The system will be made available as a standalone command line tool distributed as a Java jar file. Documentation will be provided with detailed instructions on how to install and use the command line tool.

**Distribution (Maven):** The API for the system will be published to a Maven (MVN) central repository such that it can be integrated into any Java/MVN project as a dependency statement with out the need for additional repository or other additional specifications.

**Documentation:** Detailed documentation will be generated by this project the will provide end-to-end details for how to create an OHDSI concept set from a VSAC value set.

**OHDSI CDM and Atlas Prerequisites:** It is assumed that an end user of this system has access to and has a function understanding of the OHDSI CDM and the OHDSI Atlas tool.

## Schedule

Software development and documentation is scheduled to be delivered by March 30, 2024.