

VSAC to OHDSI: Automation of the management of OHDSI concept sets using the Value Set Authority Center (VSAC)

**John E. Gresh, Raymonde Y. Uy, Julia L. Skapik
Curlew Consulting, National Association of Community Health Centers (NACHC), NACHC**

Background

The digitization of medical records and the adoption of Electronic Medical Records (EMR) systems 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-19 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 curated lists of codes and terms from various healthcare 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 codes that represent receiving any of the COVID-19 vaccines. The Value Set Authority Center (VSAC), sponsored by the National Library of Medicine (NLM), offers a comprehensive framework for the curation, validation, refinement, and publication of high-quality value sets, including all official versions of vocabulary value sets contained in the Clinical Quality Measures (CQMs) used in the certification criteria for electronic health record systems¹.

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. In public health, they support the monitoring of vaccination rates and the prevalence of conditions like HIV, informing public health strategies. 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.

Despite these advances, challenges remain. The creation and maintenance of value sets requires ongoing collaboration between clinicians, terminologists, and health informatics professionals to ensure they remain current with medical practice and terminology updates.

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 the concepts represented by these value sets, thereby enhancing the reliability of clinical research, the effectiveness of public health monitoring, and the quality of patient care.

Methods

Publicly available open source (Apache 2 license) software has been developed that enables the user to upload value sets downloaded from VSAC into the OHDSI common data model (CDM). This software is available as a standalone application or as an API that is available through Maven (MVN) central.

The standalone application is provided as an executable jar file that can be downloaded directly from the vsac-to-ohdsi github site (<https://github.com/NACHC-CAD/vsac-to-ohdsi>). Complete detailed instructions for downloading, configuring, and running the standalone application are available at https://nachc-cad.github.io/vsac-to-ohdsi/getting-started-vsac-to-ohdsi.html#Run_VSAC-TO-OHDSI.

The API can be integrated into any Java application through the simple inclusion of a dependency in the pom.xml file for that application as shown below.

```
<dependency>
  <groupId>org.nachc.cad.tools</groupId>
  <artifactId>vsactooohdsi</artifactId>
  <version>1.1.025</version>
</dependency>
```

The VSAC to OHDSI project also provides detailed information regarding using and exporting VSAC value sets as well as detailed instructions on using the software and API.

Results

The VSAC to OHDSI standalone application and API provide an automated way to import VSAC value sets into OHDSI as concept sets. These concept sets can then be accessed and used like any other concept set in OHDSI using Atlas and other OHDSI tools.

Figure 1. The COVID19 Vaccine_CVX (2.16.840.1.113762.1.4.1235.349) imported into the OHDSI CDM using the VSAC to OHDSI tool as viewed through Atlas.

Id	Code	Name	Class	RC	DRG	Domain	Vocabulary	Ancestors
722118	230	SARS-COV-2 (COVID-19) vaccine, mRNA, spike protein, LNP, bivalent, preservative free, 10 mcg/0.2 mL dose	CVX	0	0	Drug	CVX	0
722117	301	SARS-COV-2 (COVID-19) vaccine, mRNA, spike protein, LNP, bivalent, preservative free, 10 mcg/0.2 mL dose, tris-sucrose formulation	CVX	0	0	Drug	CVX	0
722119	302	SARS-COV-2 (COVID-19) vaccine, mRNA, spike protein, LNP, bivalent, preservative free, 3 mcg/0.2 mL dose, tris-sucrose formulation	CVX	0	0	Drug	CVX	0
722120	300	SARS-COV-2 (COVID-19) vaccine, mRNA, spike protein, LNP, bivalent, preservative free, 30 mcg/0.3 mL dose, tris-sucrose formulation	CVX	0	0	Drug	CVX	0
778267	229	SARS-COV-2 (COVID-19) vaccine, mRNA, spike protein, LNP, bivalent, preservative free, 50 mcg/0.5 mL or 25 mcg/0.25 mL dose	CVX	0	0	Drug	CVX	0
702678	218	SARS-COV-2 (COVID-19) vaccine, mRNA, spike protein, LNP, preservative free, 10 mcg/0.2mL dose, tris-sucrose formulation	CVX	0	0	Drug	CVX	0
724906	207	SARS-COV-2 (COVID-19) vaccine, mRNA, spike protein, LNP, preservative free, 100 mcg/0.5mL dose or 50 mcg/0.25mL dose	CVX	0	0	Drug	CVX	0
702676	219	SARS-COV-2 (COVID-19) vaccine, mRNA, spike protein, LNP, preservative free, 3 mcg/0.2mL dose, tris-sucrose formulation	CVX	0	0	Drug	CVX	0
724907	208	SARS-COV-2 (COVID-19) vaccine, mRNA, spike protein, LNP, preservative	CVX	0	0	Drug	CVX	0

Conclusion

The use of VSAC in the OHDSI community, including the development of open-source software for the automation of value set management and integration into the OHDSI framework, represents a significant advancement in the curation and management of OHDSI concept sets. This initiative streamlines the process of managing and utilizing value sets. The use of VSAC to manage and curate OHDSI concept sets also improves the consistency and accuracy of data across different healthcare systems. By automating the import of VSAC value sets into the OHDSI common data model, this project reduces the technical barriers that have historically hindered the use of VSAC in the OHDSI community. The success of this integration can potentially lead to better patient outcomes by facilitating more precise and timely health interventions.

References

- [1] Bodenreider, Olivier et al. "The NLM Value Set Authority Center." MEDINFO 2013: PROCEEDINGS OF THE 14TH WORLD CONGRESS ON MEDICAL AND HEALTH INFORMATICS, PTS 1 AND 2. Vol. 192. AMSTERDAM: IOS Press, 2013. 1224–1224. Abstract available at <https://pubmed.ncbi.nlm.nih.gov/23920998/>.