



 <https://www.linkedin.com/in/akamauu/>

## Toward real-world reproducibility: Verifying value sets for clinical research

---

Aaron Kamauu MD MS MPH

CEO & Managing Director  
*Navidence*





## Authors

Scott L. DuVall<sup>a</sup>, Craig G. Parker<sup>b</sup>, Amanda R. Shields<sup>b</sup>, Patrick R. Alba<sup>a</sup>, Julie A. Lynch<sup>a</sup>,  
Michael E. Matheny<sup>c</sup>, and Aaron W. C. Kamauu<sup>b</sup>

<sup>a</sup> VA Informatics and Computing Infrastructure, VA Salt Lake City Health Care System and Department of Internal Medicine Division of Epidemiology, University of Utah School of Medicine, Salt Lake City, UT, USA

<sup>b</sup> Navidence LLC, Aurora, CO, USA

<sup>c</sup> VA Informatics and Computing Infrastructure, Tennessee Valley Healthcare System Veterans Affairs Medical Center and Department of Biomedical Informatics, Department of Medicine, Department of Biostatistics, Vanderbilt University Medical Center, Nashville, TN, USA

This project was supported by AstraZeneca



## Background / Introduction

- The adoption of robust, reproducible methods for generating RWE from RWD is critical.
- Standardized operational definitions of clinical concepts (including value sets) are a core component of a reproducible approach.
- This method was used in support of a global RWE program to evaluate the effectiveness of AZD7442 (tixagevimab/cilgavimab) across many health systems covering several countries.
- A technology-enabled approach for handling multiple, complex value sets was piloted within the US VA EHR data using VINCI, which is mapped to the OHDSI OMOP common data model.



## Methods

### Clinical Concepts

### Verified Value Sets

#### Operational Definitions

Algorithms were created to define clinical concept(s) (e.g. 1 diagnosis code for a relevant condition in an inpatient setting OR 2 diagnosis codes in an any setting)

#### Associated Value Sets

The list of codes for each operational definition were defined as intensional algorithms or extensional (explicit) lists. Value sets encompassed multiple coding schemas (e.g. ICD-10-CM, CPT®, HCPCS, ICD-10-PCS, RxNorm, ATC).

#### Terminology Crosswalk

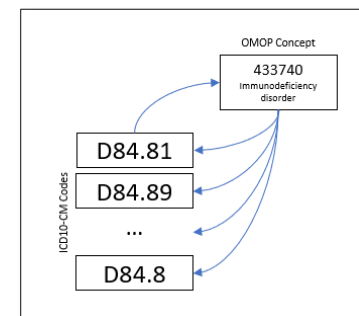
All concepts were processed against their respective hierarchies to identify additional granularities that were not included.

#### (1) EHR Usage Check

Broad queries were run against the source EHR data to identify any codes that did not map or those that resulted in zero instances.

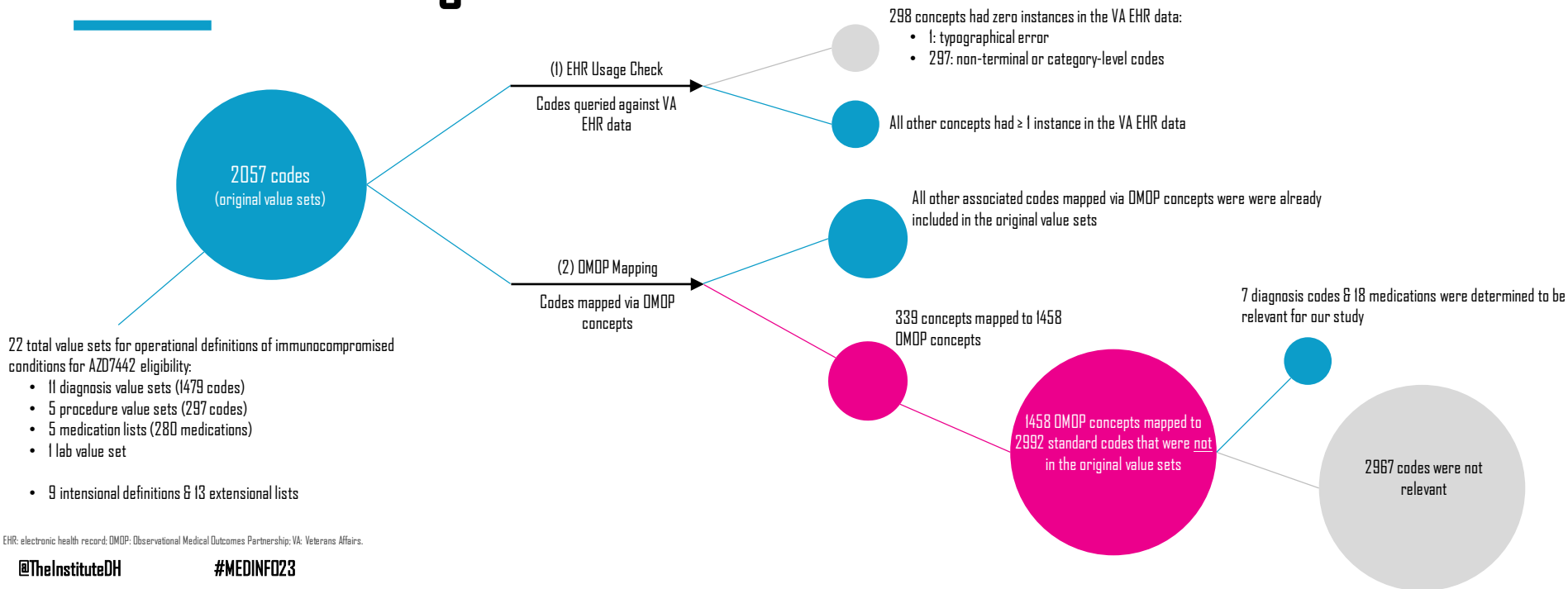
#### (2) OMOP Mapping

Standard codes were mapped with corresponding concepts in OMOP, and then mapped back to all codes in the original terminology that mapped to the same OMOP concept. The resulting codes were compared to the original value sets and evaluated to identify potentially relevant additional concepts.





## Results / Findings





## Discussion

### (1) EHR Usage Check

- Identified a typo that was fixed
- Non-terminal codes
  - not expected in patient data
  - we decided to include them in our value sets for completeness and hierarchical reference

### (2) OMOP Mapping

#### 7 Relevant Diagnosis Codes

New 'children' ICD-10-CM codes were released in Oct 2022

Added to our value sets

#### 18 Relevant Medications

Drug names did not match in the VA EHR

An RxNorm ingredient code was determined from OHDSI Athena tool

Added to our value sets

#### 2967 codes were not relevant

- OMOP concepts were not relevant  
ICD-10-CM code O98.711 <HIV disease complicating pregnancy, first trimester>  
→ OMOP concept "First trimester pregnancy"
- Individual codes were not relevant  
ICD-10-CM code C7A.012 <Malignant carcinoid tumor of the ileum> → OMOP concept → ICD-10-CM code D3A.012 <Benign carcinoid tumor of the ileum>



## Conclusion

- This pilot project demonstrated the use of
  - a robust and reproducible method for validating, enhancing and updating value sets used in operational definitions for clinical research
  - leveraged standard terminologies, dictionaries and data models to provide a comprehensive set of relevant concepts
- This provided confidence in operational definitions for our RWE research program across several countries
- Further application to additional studies, possibly via the OHDSI collaboration, can help socialize the value of this method to the broader clinical research community



## References

1. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, Saunders LD, Beck CA, Feasby TE, Ghali WA. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Medical care*. 2005 Nov 1;1130-9.
2. Quan H, Li B, Couris CM, Fushimi K, Graham P, Hider P, Januel JM, Sundararajan V. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. *American journal of epidemiology*. 2011 Mar 15;173(6):676-82.
3. Beyrer J, Manjelienskaia J, Bonafede M, Lenhart G, Nolot S, Haldane D, Johnston J. Validation of an International Classification of Disease, 10th revision coding adaptation for the Charlson Comorbidity Index in United States healthcare claims data. *Pharmacoepidemiology and drug safety*. 2021 May;30(5):582-93.
4. de novo code lookup informed by Informed by National Library of Medicine Value Set Authority Center. NLM Value Set Repository. [Accessed 10 April 2022] Available from: <https://vsac.nlm.nih.gov/welcome>
5. Carnahan R. RxNorm RxCUI codes for Cancer Therapies. PCORnet Distributed Research Network Operations Center. University of Iowa. PheKB; 2021 Available from: <https://phekb.org/phenotype/I578>
6. Haendel, M. A., Chute, C. G., Bennett, T. D., Eichmann, D. A., Guinney, J., Kibbe, W. A., ... & Gersing, K. R. (2021). The National COVID Cohort Collaborative (N3C): rationale, design, infrastructure, and deployment. *Journal of the American Medical Informatics Association*, 28(3), 427-443.
7. Master Protocol Development: COVID-19 Natural History. Sentinel Initiative. 2020 [Accessed 5 April 2022]. Available from: <https://www.sentinelinitiative.org/methods-data-tools/methods/master-protocol-development-covid-19-natural-history>
8. Lynch KE, Deppen SA, DuVall SL, Viernes B, Cao A, Park D, Hanchrow E, Hewa K, Greaves P, Matheny ME. Incrementally Transforming Electronic Medical Records into the Observational Medical Outcomes Partnership Common Data Model: A Multidimensional Quality Assurance Approach. *Appl Clin Inform*. 2019 Oct;10(5):794-803.





 <https://www.linkedin.com/in/akamauu/>

## Toward real-world reproducibility: Verifying value sets for clinical research

---

Aaron Kamauu MD MS MPH

CEO & Managing Director  
*Navidence*

