Managing Interoperability of Healthcare Data for Decision Analytics and Research

JR Campbell,JC McClay,CR Geary, WS Campbell, & AJ Anzalone

1 Introduction

This four-hour workshop for the advanced participant will promote interoperability and analytics in healthcare datamarts being developed and implemented for decision support, research and public health. It will guide the participant in implementing the guidelines of the ONC Interoperability Standards Advisory(ISA). The workshop team will employ their research and experience with HL7 and SNOMED CT to demonstrate the integration of SNOMED CT, LOINC and RxNORM as domain ontologies integrated with FHIR resources. Using a series of use cases taken from implementation of research network datamarts, they will present and discuss the relevant ISA specifications, and the associated ontology standards, some being developed at Nebraska. For each use case they will highlight any ambiguities in the ONC specifications and discuss the decisions to be made relative to deploying the semantics of the reference domain standards. They will present resources for managing and implementing relevant ONC sanctioned ontologies within the ISA model, including an integrated ontology of SNOMED-LOINC-RXNORM developed and published as terminology resources by Nebraska. Special attention will be focused on addressing problem areas of translational research that require integration of reference domain standards to bridge laboratory medicine, pathology, microbiology, genomics and patient phenomics. Each use case will be accompanied by demonstration of the data analytic utility of ontologies for extraction of meaningful knowledge from the EHR. Participants will conclude the workshop with an understanding of and access to a set of open source resources and tools for binding, integrating and maintaining the ontologies at their own institutions.

1.1 Learning Objectives

After participating in this session the learner should be better able to:

- Describe and discuss features of an interoperable data model compliant

with the 2021 ONC Interoperability Standards Advisory for EHR core data domains

- Understand the semantic scope, strengths and weaknesses and tooling for terminology standards: SNOMED CT, LOINC, RxNorm, NDC, (CPT-4, ICD-10-CM/PCS, CVX)
- Identify and manage knowledge gaps present in domain ontologies in support logical integration of ontologies
- List resources and tools of use in deploying, maintaining and updating ONC interoperability standards

1.2 How to Use this Pamphlet

This accompanying pamphlet contains references and resources that accompany the sections of this workshop. We have compiled a list of important resources on the management and maintenance of medical terminology and other measures that are available for use during and following the workshop in the handouts folders.

1.2.1 Handouts Inventory

- RUCA2010ZIPCODE.csv
  - 2010 Rural-Urban Commuting Area (RUCA) Codes for the US organized by 5 digit zip code.

- RXNORM_NDC_RELATIONSHIPS_20220501.zip
  - Comprehensive and historically complete file of useful relationships which have been published by NLM and the FDA. This includes relationships for RXCUIs, NDCs and VA Drug Classes that might now be listed as inactive in current UMLS releases. The Relationship types included are:
    1) VA Class to Ingredient,
    2) Active Ingredient to Clinical Drug,
    3) Active Ingredient to Manufactured Product Package code, and
    4) Clinical Drug to Manufactured Product package Code.
  - Active ingredients are RXCUIs from TTY classes PIN, MIN and IN. Clinical Drugs are RXCUIs from TTYs of SCDC,SCD,GPCK, SBD and BPCK. Manufactured Product Package codes are exclusively NDC.

- SDH_COLUMNS.csv
  - Data dictionary including for the current release of community-level Social Determinants of Health Observation concepts in Nebraska Lexicon©

- SDH_Stats.csv
  - Data inventory and Nebraska Lexicon© SNOMED CT extension codes for 63 Community-level data items computed from American Community Survey 2019 Census data including Bird Index, Wisconsin Area Deprivation Index, American Community Survey detail and the Social Vulnerability Index.

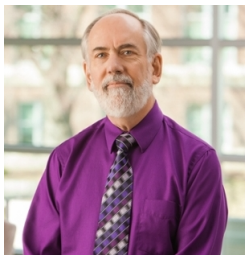
- SNOMEDCT_USExtension_NebLexicon_LOINC_RXNORM.zip
  - OWL file with integrated ontology of SNOMED CT© International Edition 20220130 + US Extension© 20220330 from NLM + Nebraska Lexicon© SNOMED Extension + 23K Laboratory LOINC© concepts imported as axioms + RxNorm© generic clinical drugs imported as axioms

- SVI2018DocumentatioN-h.pdf
  - CDC/ATSDR Social Vulnerability Index(SVI) documentation for Social Determinants of Health Data data items in SDH_Stats table related to 4 SVI themes and summary score.

- Doc_ExpressionConstraintLanguage_v2.0-en-US_INT_20220404.pdf
  - SNOMED CT Expression Constraint Language Specification and Guide Version 2.0 This includes the technical specifications and ABNF syntax for current version of Expression Constraint Language.

- SNOMED CT Editorial Guide-2022-01-31.pdf
  - Current version of the authoritative technical reference guide for the structure and function of SNOMED CT© and the specifications of its concept model (ontology conceptualization)

1.3 About the Presenters



James R. Campbell, MD

Jim Campbell MD, FACP, FACMI mustered out of the Air Force in 1972 and earned his MD from the University of Nebraska in 1976. He is boarded in Internal Medicine and has served as a primary care physician for 45 years. He became interested in informatics through involvement in the newly formed Society for Computer Applications in Medical Care(SCAMC) and developed his first electronic health record(EHR) for installation in the outpatient medicine clinic in 1982. He was convinced at then that standardization of healthcare terminologies were core requirements for an EHR that could support reliable and consistent outcomes for his patients and joined the SNOMED Editorial Board in 1990, He has worked actively with Standards Development Organizations including HL7 Decision Support and Guidelines projects, LOINC and SNOMED International and ICD-11 technical advisory team through the intervening years.He is a certified Consultant Terminologist with SNOMED International and has certified as an Epic® Ambulatory Analyst. His ongoing research interest is the standardization and interoperation of healthcare terminologies in service of research analytics and clinical decision support.

James C. McClay, MD, MS

James McClay, MD, MS, has over thirty years of clinical, informatics, and research experience as an Emergency Physician. He recently became the Chief Research Informatics Officer at the University of Missouri. Prior to this role, he was principal investigator for the University of Nebraska Clinical Research Analytics Environment (CRANE), director of the Great Plains IDeA Clinical and Translational Research Center Biomedical Informatics Core, and chaired the University of Nebraska Medical Center (UNMC) Biomedical Informatics Graduate Degree Program. His work focused on developing and operationalizing evidence generation pipelines for learning health systems. Dr. McClay is a long-standing Health Level Seven (HL7) participant as a workgroup chair and a clinical leader in the development of FHIR specifications.

Carol R. Geary, PhD, RN, MBA

Carol Geary, PhD, RN, MBA, is a nurse researcher who began her focus on use of structured electronic health record research during postdoctoral study with Dr. McClay. She spends much of her time focused on developing and instituting infrastructure necessary to support such usage. She has recently become the UNMC PCORnet Principal Investigator, building upon previous experience as Patient Engagement Officer. In this context, she aims to enhance availability and usability of structured electronic health data in nursing research.

W. Scott Campbell, PhD

Scott Campbell, PhD, MBA is a tenured, associate professor in the Department of Pathology and Microbiology at the University of Nebraska Medical Center (UNMC) in Omaha, Nebraska, USA. He is also the Director of Informatics for the Nebraska Public Health Laboratory and the Sr. Director of Research Information Technologies for UNMC. Dr. Campbell specializes in clinical informatics with special emphasis in cancer pathology and microbiology.

The development of SNOMED CT concept expressions for use in cancer synoptic reporting which includes biomarker and next generation sequencing data comprises significant aspects of Dr. Campbell's research efforts. He currently serves as the Chair of the Cancer Synoptic Reporting Working Group for SNOMED International which is creating the necessary knowledge representation and concepts for the SNOMED CT international release for use in structured cancer reporting.

A. Jerrod Anzalone, MS

Jerrod Anzalone, MS, is a PhD candidate in biomedical informatics and works as a clinical research informatics specialist for the University of Nebraska Medical Center (UNMC). His research interests are in applied clinical informatics, the secondary analysis of healthcare data for research, and the maintenance of medical terminology.. Over the past two years, he has worked heavily in the National COVID Cohort Collaborative (N3C), focusing on rural populations, immunosuppresed/compromised persons, and data quality issues in multi-center studies.

1.4 Learn More

To learn more about the work done by this research team on terminology development, please visit:

- Nebraska Lexicon
- University of Nebraska Medical Center's Clinical Research Analytics Environment GitHub

2 Introduction to semantic interoperability; Top-level ontologies for the US domain

This chapter provides references and resources for learning more about the state of interoperability in the United States. It is not meant to be comprehensive, but should provide the workshop attendee with the requisite background to find out what is being done, by whom, and where to go to learn more.

2.1 Learning Healthcare System

- Institute of Medicine (US) Roundtable on Evidence-Based Medicine, Olsen L, Aisner D, McGinnis JM, eds. The Learning Healthcare System: Workshop Summary. Washington (DC): National Academies Press (US); 2007.

2.2 US Core Data for Interoperability

- US Core Data for Interoperability v2

2.3 Fast Healthcare Interoperability Resources (FHIR)

- Lehmann HP, Downs SM. Desiderata for sharable computable biomedical knowledge for learning health systems. Learn Health Syst. 2018;2(4):e10065. https://onlinelibrary.wiley.com/doi/full/10.1002/lrh2.10065.
- USCDI FHIR Implementation Guide
  - Condition Profile
  - Laboratory Results Reporting
  - Vital Sign Profile
  - Medication Profile
  - MedicationRequest Profile

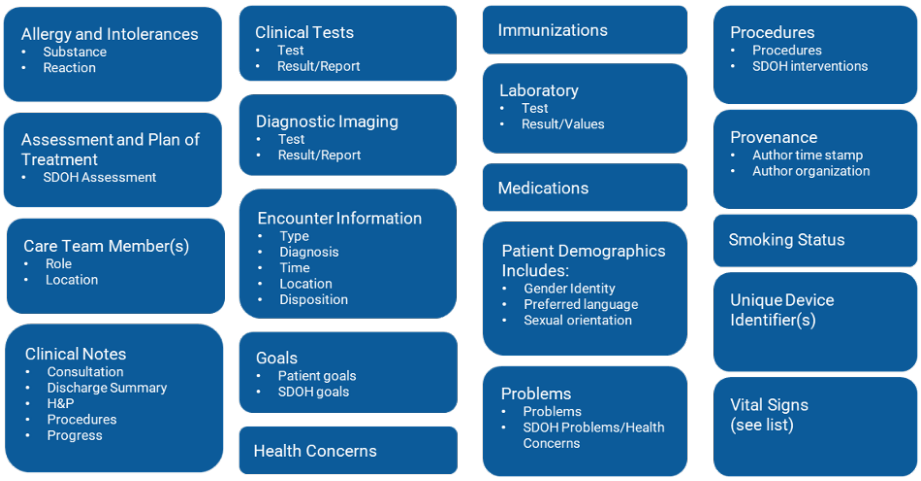


Fig. 1: USCDI v2

2.4 Common Data Models

- The National Patient-Centered Clinical Research Network (PCORnet)
  - PCORnet Common Data Model v6.0
- Observational Health Data Sciences and Informatics (OHDSI)
  - The Book of OHDSI
  - OMOP Common Data Model v5.3

3 Interoperable Use Cases for CONDITIONS and DIAGNOSES

This chapter provides references and resources for learning more about patient condition diagnoses.

3.1 Computable Phenotyping

- Electronic Health Records-Based Phenotyping
- Richesson LR, Sun J, Pathak J, Kho AN, Denny JC. Clinical Phenotyping in selected national networsk: demonstrating the need for high-throughput, portable and computational methods. Artif Intell Med. 2016 July; 71:57-61.

3.2 Terminologies

- ICD-10 CM
  - Tabular browser view download
  - Alpha browsing tool
  - Available from CMS or CDC
- SNOMED CT
  - SNOMED CT Browser
  - Downloads of SNOMED International + US Extension available twice yearly (March and September) from NLM terminology services

4 Interoperable Use Cases for LAB RESULTS including COVID-19 and ANATOMIC PATHOLOGY

This chapter provides references and resources for learning more about the interoperation of laboratory terminology and anaomtic pathology.

4.1 LOINC

- LOINC
  - LOINC publishes release twice annually and supports downloads and browsing tools
  - Browsing and term lookup available in RELMA tool and from the LOINC website
  - Frequency analysis of US lab results from our site included in handout material: NEBRASKA_SEMANTIC_CODE_FREQUENCY.xlsx
  - Structured Terminology Model of Meaning: LOINCUsersGuide.pdf

4.2 Synoptic Pathology

- College of American Pathologists Cancer Protocol Templates
- Nebraska Lexicon

5 Interoperable Use Cases for MEDICATION ORDERS, MEDICATION ADMINISTRATION and DRUG ALLERGIES

This chapter provides references and resources for learning more about the interoperation of medication orders, medication administration, and drug allergies.

5.1 RxNorm

- First published by NLM in 2002 as UMLS project for standardized nomenclature for US clinical drugs
- Downloads available at UMLS Terminology Services
- Browsing supported by RxNav for terms and semantic network
- RxNav-in-a-Box supports local-installable Docker version of all APIs and allows updating of all RXNorm data files
- RxClass API supports maps to drug classification schemes
- Handouts: RXNORM_NDC_RELATIONSHIPS_20220501.zip

5.2 National Drug Codes (NDC)

- Database maintained by the FDA and reported in RxNav which requires manufacturers to report all drugs manufactured, prepared, propagated, compounded or processed for sale in the US
- Handouts: RXNORM_NDC_RELATIONSHIPS_20220501.zip

6 Interoperable Use Cases for VITALS, CLINICAL ASSESSMENTS, NURSING ASSESSMENTS and SOCIAL DETERMINANTS OF HEALTH

This chapter provides references and resources for learning more about the interoperation of vitals, assessments, and social determinants of health (SDOH).

7 Assessments

- LOINC
- SNOMED CT # Social Determinants of Health (SDOH)
- Healthy People 2030
- Efforts made to standardize SDOH data elements:
  - World Health Organization
  - The Gravity Project
  - USCDI
  - Research Networks
    - OHDSI Health Equity Workgroup
    - N3C SDOH Domain Team
  - Regional and National Measures and Indices
    - US Census American Community Survey

* Wisconsin Area Deprivation Index
* CDC Social Vulnerability Index
* BIRD Index
* USDA ERS Rural Urban Commuting Area (RUCA) Codes
* USDA Food Environment Atlas

8 References

Selected references from this workhop include:

1. The Learning Healthcare System. Ed: LeeAnne Olsen, Dara Aisner, Michael McGinnis. Institute of Medicine, National Academy of Sciences 2007
2. Lehmann HP, Downs SM. Desiderata for sharable computable biomedical knowledge for learning health systems. Learn Health Syst. 2018;2(4):e10065. https://onlinelibrary.wiley.com/doi/full/10.1002/lrh2.10065.
3. OHDSI. The Book of OHDSI: Observational Health Data Sciences and Informatics. OHDSI; 2021.
4. Yu Y, Zong N, Wen A, et al. Developing an ETL tool for converting the PCORnet CDM into the OMOP CDM to facilitate the COVID-19 data integration. J Biomed Inform. 2022;127:104002. doi:10.1016/j.jbi.2022.104002
5. Richesson LR, Sun J, Pathak J, Kho AN, Denny JC. Clinical Phenotyping in selected national networsk: demonstrating the need for high-throughput, portable and computational methods. Artif Intell Med. 2016 July; 71:57-61.
6. Cimino-J. Desiderata for Controlled Medical Vocabularies in the Twenty-First Century. Methods Information in Medicine. 1998 Nov; 37(4-5):394-403.
7. Bodenreider O, Cornet R, Vreeman DJ. Recent developments in clinical terminologies SNOMED CT, LOINC and RxNorm. Yearbook Med Inform 2018;27(1):129-139.
8. Teasdale G, Jennett B. Assessment of coma and impaired consciousness. Lancet 1974; 81-84.
9. Teasdale G, Jennett B. Assessment and prognosis of coma after head injury. Acta Neurochir 1976; 34:45-55.
10. Westra, Bonnie L., et al. "Validation and refinement of a pain information model from EHR flowsheet data." Applied clinical informatics. 9.01 (2018): 185-198.
11. Healthy People 2030, U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion. Retrieved [08 May 2022], from https://health.gov/healthypeople/objectives-and-data/social-determinants-health

12. McClure et al. Gender harmony: improved standards to support affirmative care of gender-marginalized people through inclusive gender and sex representation [published correction appears in J Am Med Inform Assoc. 2021 Nov 25;:]. J Am Med Inform Assoc. 2022;29(2):354-363. doi:10.1093/jamia/ocab196