Mapping Dental Use Cases to the OMOP-CDM: Vocabulary and Common Data Model Evaluation

Robert Koski[1], Gopikrishnan Chandrasekharan[2], William D. Duncan[3]
[1] Johns Hopkins University, [2, 3] University of Florida

Background

Medical specialties that wish to transform their data to the Observational Medical Outcomes Partnership common data model (OMOP-CDM) follow a specific process that begins with use case generation. Use case generation is important because it helps define what data is important to the transformation and what terminologies will be needed to conduct analysis (1). Use cases also restrict the scope of transformation so that time and resources are not wasted with data that may not be needed for analysis. Dentistry is in the earliest stages of adopting the OMOP-CDM, so developing use cases and attempting to map them to the OMOP-CDM is a logical step toward using dental data for observational research. This project uses one hypothetical dental use case to simulate how dental data would be mapped to the OMOP-CDM and to describe the types of analyses that could be conducted from such a mapping. More importantly, this mapping shows where dentistry should focus efforts to improve its capability to conduct high-quality observational research.

Methods

Hypothetical use cases were developed based on the three types of questions that observational research can answer: characterization, population health estimation, and patient level prediction (2). One use case was selected: Amongst patients that received a posterior composite restoration, how many patients experienced restoration failure within five years? The use case was selected because of the variety of general dentistry related concepts in the scenario.

Use case data was generated by a general dentist advanced training in general dentistry and advanced training in health informatics. The use case was mapped by three individuals with advanced health informatics training and familiarity with dental concepts and terminology.

The hypothetical use cases were presented in Miro, a virtual board software program. This program was chosen because of its ability to visualize all aspects of the case at all levels of granularity (from high level overview of the project to viewing individual data elements in specific tables). The software also has several features that allow easy navigation between boards.

Synthea Generic Module Builder is a tool that allows users to create modules. These modules represent a model of a clinical case in a domain. For example, Synthea has pre-built modules for conditions like Asthma and Epilepsy. The module can model the different states a patient can encounter in the process of having a medical condition. This includes the onset of clinical conditions, medical encounters, medications prescribed, procedures done, and end of clinical conditions. Public health data can be used to model the probabilities of each state.

The Synthea Generic Module Builder was used to model a dental patient who has a caries condition. We were able to model the onset of caries, a visit to the dental clinic for a periodic dental exam, obtaining radiographs, administering of local anesthetic, and a procedure for a dental composite restoration. SNOMED, RxNorm, and LOINC codes were used to annotate the dental terms for each of these states.

We used an ETL tool called ETL-Synthea (https://github.com/OHDSI/ETL-Synthea) to create relational

tables in a Postgres database. 2 separate schemas were created by the tool, one for OMOP CDM and one for the Synthea FHIR data. The tool also loaded vocabulary downloaded from the Athena website. Once the tables were generated and the vocabulary loaded onto the respective tables in the database the ETL tool loaded the Synthea FHIR data onto the Synthea tables. Then the ETL tool would extract data from the Synthea tables, transform the data to OMOP representation, and then load the data to OMOP tables in the database.

The Observational Health Data Sciences and Informatics (OHDSI) community has comprehensive documentation for the open-source OMOP-CDM. The latest release of the OMOP-CDM, v5.4 (3), was used to map the hypothetical use case. Dental concepts were selected from the use case by consensus of the three mappers and mapped to the OMOP-CDM. The concepts were evaluated by two metrics.

First, the concept was judged based on terminology available in standardized vocabularies to adequately describe the concept. Second, the concept was judged based on its ability to map distinctly to a parameter in the OMOP-CDM. The criteria for evaluating terminology and schema fit are underway in collaboration with the OHDSI Vocabulary Subgroup.

The proposed SITE and SITE_EVENT tables for the tooth-centric mapping demonstrates that it is possible to use the OMOP-CDM to represent data to the granularity of teeth; however, this is a complex mapping schema that may be unsuitable for many use cases and analyses. Further discussion and consensus is needed prior to adoption.

Mapping the concepts is ongoing and additional results are expected in late 2023.

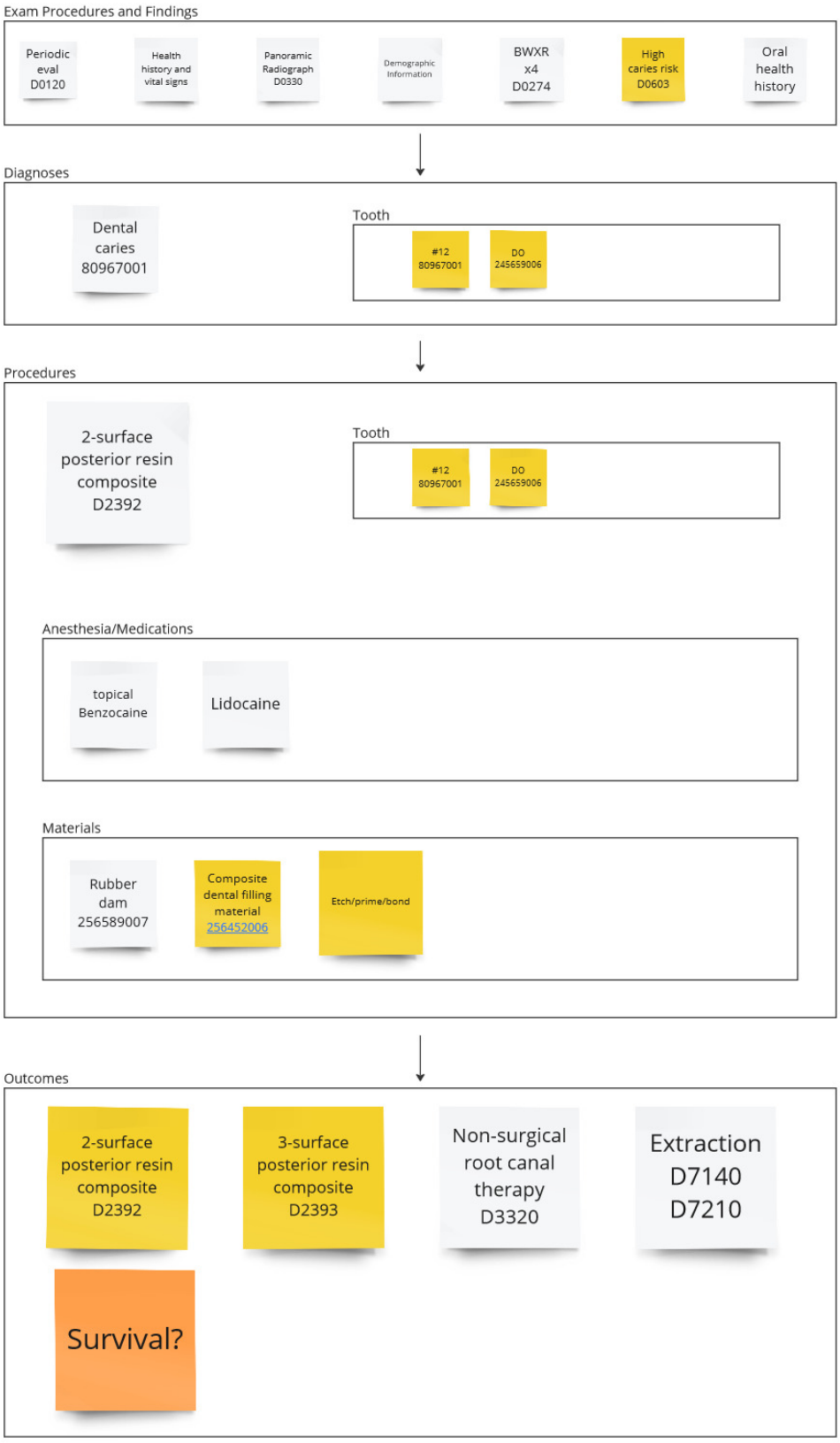
Results

Synthea allowed us to create realistic synthetic patient cases with multiple comorbidities. It also automatically created claims, allergies, organizations, providers, and medication data based on the built-in modules. However, the module builder is limited by its ability to model body sites for conditions and procedures. Even though we were able to model conditions like dental caries we were unable to model the tooth, the location of the tooth, and the extent of the spread of dental caries on the tooth even though terms existed for these concepts. We were unable to model dental caries on multiple teeth, even though this is a common occurrence among patients.

In total, 18 concepts are being considered for this mapping. The diagram on the following page shows the use case concepts being considered for mapping. Concepts in gray are projected to have both adequate terminology and adequate schema representation. Concepts in dark yellow have either inadequate schema or inadequate terminology representation. Representation of the survival of the tooth, indicated in orange, is being discussed for how this might be represented in dental data.

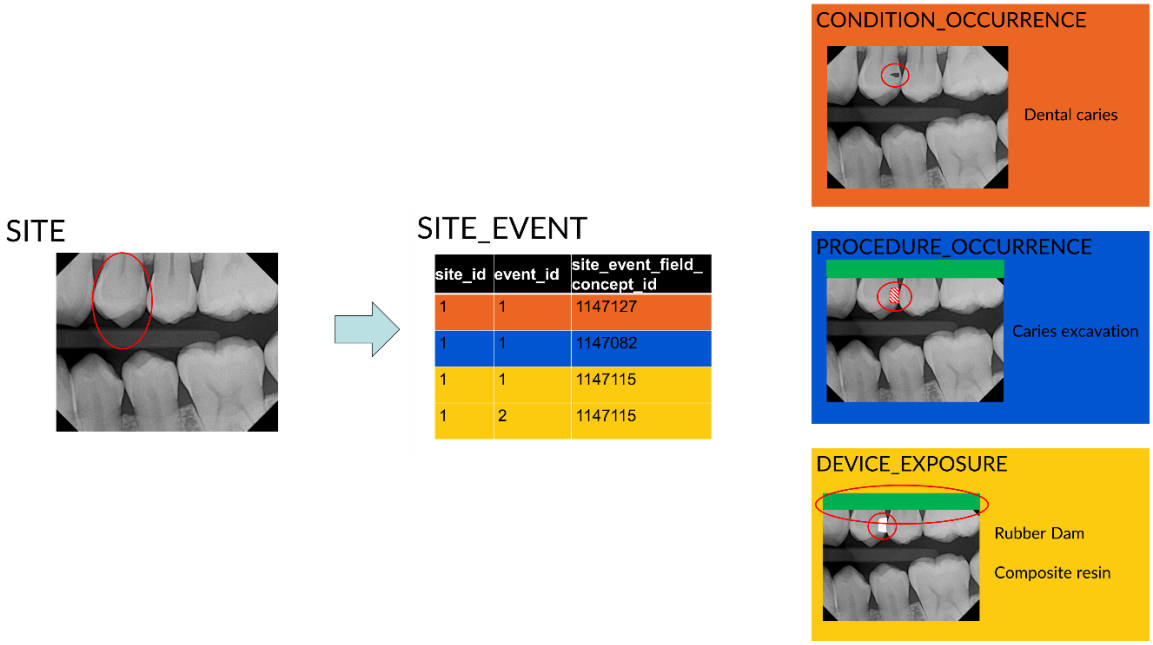
Total projected mapping and terminology representation from current concepts.

Projected Schema and Mapping Representation	Count
Adequate Schema and Terminology	11
Inadequate Schema or Terminology	7


Exam Procedures and Findings
Periodic eval D0120
Health history and vital signs
Panoramic Radiograph D0330
Demographic Information
BWXR x4 D0274
High caries risk D0603
Oral health history
Diagnoses
Dental caries 80967001
Tooth
#12 80967001
DO 245659006
Procedures
2-surface posterior resin composite D2392
Tooth
#12 80967001
DO 245659006
Anesthesia/Medications
topical Benzocaine
Lidocaine
Materials
Rubber dam 256589007
Composite dental filling material 256452006
Etch/prime/bond
Outcomes
2-surface posterior resin composite D2392
3-surface posterior resin composite D2393
Non-surgical root canal therapy D3320
Extraction D7140 D7210
Survival?

One of the key findings from the use case is that the person-centric nature of the OMOP-CDM does not allow adequate representation of dental specific use cases. Dental use cases, such as characterization of posterior composite resin restorations, often require longitudinal tracking of individual teeth. The OMOP-CDM does not currently able track anatomic site data and events at the site level. The proposed SITE and SITE_EVENT tables are a first iteration to expanding the current schema to allow anatomic sites to be adequately represented and tracked.

In the proposed tables represented below, the anatomic site would be represented in the SITE table, similar to episodes in the EPISODE table in OMOP-CDM v5.4. The sites would be related to the events and exposures that occur to them over their life through the SITE_EVENT table, which connect the site to the unique identifier in each table that the event exists in. Further development of this schema is ongoing.



Conclusion

Mapping the dental use cases highlights the need for additional discussion within the dental professions about common data model use. Schema and vocabulary concerns must be addressed to maximize the benefits of applying observational research to dental use cases. The Dentistry WG hopes to highlight the challenges of observational research that are unique to dentistry as well as generate interest for dental observational research with this project. Additional concepts will be considered in the use case and mapping is ongoing.

References

1. Lamer A, Abou-Arab O, Bourgeois A, Parrot A, Popoff B, Beuscart JB, et al. Transforming Anesthesia Data Into the Observational Medical Outcomes Partnership Common Data Model: Development and Usability Study. J Med Internet Res. 2021;23(10):e29259.
2. The Book of OHDSI: Observational Health Data Sciences and Informatics; 2021.
3. OHDSI. OMOP Common Data Model: OHDSI; 2022 [Available from: https://ohdsi.github.io/CommonDataModel/.