Meeting with Alex Wagner 5/8/2020

Next Steps:

1. Since the goal of the OHDSI Genomic efforts is to have a list of concepts, it was decided that metaKB is a good place to start. OHDSI Genomic WG to figure out what attributes are needed for our conceptual starting point. Based on the attributes, Alex will pull out the relevant data from metaKB, standardize it to the specification that is required by OHDSI and create a recurring flat file that we can grab. Coordinate on a repo to place this information.
2. If OHDSI gets new concepts, assuming they are in the cancer domain, OHDSI can provide these to CIVIC that has a curation interface. metaKB updates from CIVIC so it goes through the natural flow without manual intervention.

Key points discussed:

1. Goal of the OHDSI Genomic efforts are to create a min viable product with a limited list of concepts. For this at a min we need a unique ID, human readable name and a bunch of links to the original sources of information. As a starting point, we do not need to have all facets.
2. VICC is a driver project of GA4GH, metakb is a VICC product where all knowledgebases work together to contribute to the interoperability of the data elements and how to bring them together.
3. CinGen, VRS and metaKB were evaluated and discussed to access the right fit as a starting point for OHDSI
4. ClinGen Allele Registry (CAR) is comprehensive and has ClinVar, separate concepts for protein changes. For future, CNV is in development. ClinGen does have concepts for structural variation or level of granularity required by OHDSI. One recommendation discussed was to point to ClinGen Allele Reg (CAR), get normalized concepts, takes those concepts and compute VRS ID on top of them. Concerns with ClinGen are that it cannot be downloaded as its not open.
5. VRS (VR Specification) provides a flexible framework to precisely and consistently exchange genetic variation data across communities. VRS aims to improve the reliability and utility of the clinical annotations that are central to personalized medicine. It’s not designed to be central registry. A centralized registry could be created using VRS to build upon the identifiers. VRS continues to evolve as it has major stakeholder. VRS has variation set composed of individual variations.
6. metaKB is a flattened version of the concepts. It is not recommended as a standard but rather a baseline and not an annotation tool. It’s designed to help identify disparities and similarity between standard to develop good standards. metaKB harmonizes clinical interpretation evidence which includes variation concepts and other things. metaKB is not going to be expanding under the VICC project efforts anymore due to focus on variant lexicon. VRS IDs exist in metaKB.
7. To maximize searchability, MetaKB captures all corresponding variants that link to the protein change from ClinVar, captures the HGVS strings and normalizes HGVS strings using VRS to obtain a unique VRS ID across those concepts. Tooling exists to track back the HGVS ID and VRS ID which allows for searching between systems.