


Patient / Research Subject / Sample Level Common Metadata (Draft v0.1)

Patient / Research Subject Public Metadata (Common, Minimal Metadata)

A. From each data platform in the hub

- Age (90 or under)
- Anchor Event (for durations)
- Biological Sex
- Race/Ethnicity (there may be more than one, it will be from a controlled vocabulary)
- Project-specific ID or GUID
- Prometheus Global ID
- Apollo ID (if relevant)
- Primary Disease
- Cancer Type
- Metastasis (Y/N)
- Cancer Grade
- Cancer Stage
- Type of Exposure ([list of exposure types from](#) a controlled vocabulary)
- Chemical/agents exposed to (list of exposure types from controlled vocabulary, check if we can use SNOMED codes) (Information shared by Danne:
 SNOMED_Interoperability_Update_20240708.pptx , To check any substances in SNOMED: [SNOMED CT - Home \(ihtsdotools.org\)](https://ihtsdotools.org/))

B. Computed by Hub

- Boolean variables indicate (Y/N) (node/commons level data)
 - Is there Metabolomic data associated with subject
 - Is there exposure data associated with subject
 - Is there liquid biopsy data associated with subject
 - Is there VPODC data associated with subject
 - Is there Imaging Data (TCIA) associated with subject

Possible attributes to include

- Gender Identity
- Assay method/lab tests for determining exposures (controlled vocabulary or string)

More generally, what principles do we use for pulling data from particular studies in particular (Prometheus commons/nodes) into the public metadata?

Study Level Common Metadata (Draft v0.1)

Study Level Common, Minimal Metadata

- Study/Project Name
- Prometheus Study ID
- Project Number (may be in addition to Prometheus Study ID)
- Institution or Organization
- PI
- Investigators
- Study Description Summary
- Research Focus Area
- Type of study (preanalytic, clinical utility, clinical efficacy, clinical trial, clinical trial stage, questionnaires, etc., from a controlled vocabulary)
- Types of subjects (human, animal, from a controlled vocabulary)
- Current number of subjects
- Version Number
- Date of last updated
- Controlled Access, open access, both open and controlled access (controlled vocabulary)
- Accrual status (not started accruing, ongoing accrual, closed, etc., from a controlled vocabulary, relevant for some studies, so include NA)
- Dataset description
- study population (controlled vocabulary: active duty, veterans, civilian, mixed)
- Commons/Nodes hosting data
- Tags to support search

Possible attributes to include

- Geographical regions associated with the study
- Research Program (in addition to Research Focus Area)
- List of data types included (from a controlled vocabulary)