-- Overall count of pre-coordinated ICD-O conditions: 62,535

-- Count of ICD-O topography concepts: 400

-- Count of pre-coordinated ICD-O conditions with topography only: 400

-- Count of ICD-O histology concepts: 1,536

-- Count of pre-coordinated ICD-O conditions with histology only: 1, 476 (less than histology concepts)

--Counts of valid and invalid concepts

|  |  |  |
| --- | --- | --- |
| **invalid\_reason** | **count(\*)** | **comments** |
| NULL | 59877 |  |
| D | 830 | The reason for invalidation should be explained. It looks like majority of these concepts should not have been created at all because they are based on non-existing topography concepts. |
| U | 1828 | I see patients with some of these invalidated concepts in our tumor registry (e.g. 8460/0-C56.9). If these conditions do not exist, the respective concepts should be removed from the vocabulary. For the existing ones, they should be re-instated. A lot of them do not have established relationships. |

-- Counts of ICD-O conditions mappings

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **standard\_concept** | **maps to** | **vocabulary\_id** | **count(\*)** | **Comments** |
| N | Maps to | ICDO3 | 1792 |  |
| N | Maps to | SNOMED | 2713 |  |
| N | Not mapped | NULL | 796 | Looks like most of them are invalid (‘D’) and based on non-existing topography concepts, but there are a few of valid concepts with incorrect concept codes. All of them should be removed. |
| S | Maps to | ICDO3 | 56858 |  |
| N | Maps to | Cancer Modifier | 377 |  |

-- Completeness and concordance between ICD-O and SNOMED topography relationships for Standard ICD-O condition concepts only: see Topography Relationship sheet in the attached Excel file.

-- Completeness and concordance between ICD-O and SNOMED histology relationships for Standard ICD-O condition concepts only: see Histology Relationship sheet in the attached Excel file.

There are a few ICD-O conditions that do not have SNOMED relationships (e.g. cancers of Blood do not have SNOMED Finding Site). This may results in omitting them when creating SNOMED-based phenotypes. This should be looked into.