Abstract: last

Background

The Federal Health Information Model is an effort supported by the United States Office of the National Coordinator for Health Information Technology, and it represents health information that might be exchanged among US federal agencies and their partners. This representation serves two purposes. First, it provides a venue for conceptual harmonization across organizations that may otherwise define data elements for similar purposes in ways different enough to be incommensurable. Second, it does so in a toolset, the Model Driven Health Tools, that can use these elements to generate both implementable specifications (e.g., Clinical Document Architecture, NIEM, or FHIR implementation guides) and executable application code to generate, validate, and consume instances of these specifications.

One of the tasks of element harmonization is to specify the semantics of coded elements by defining value sets and publishing them in a commonly accessible repository (the CDC’s Public Health Information Network Vocabulary Access and Distribution System, or PHIN VADS, and the NLM’s Value Set Authority Center, or VSAC).

We report our experience defining the value set for the allergy & intolerance reaction data element.

Approach

We proposed to use existing reaction lists from three healthcare systems to create a common list. This list would support interoperability among healthcare systems and provide a common practice for the assignment of reaction values in clinical practice.

We anticipate that all clinical practices will need the ability to capture reactions not included in our list; the universe of reaction types has, in statistical parlance, a very long tail. This would mean that the element definition should permit the use of virtually any disorder concept, so there was some question regarding the utility of a smaller value set. The utility envisioned was to provide a “short list” that might serve both technical and clinical needs. Technically, the list could be used to populate user interface selection tools and provide a more expedient validation routine than the entire set of disorders. Clinically, it could provide a forum for convergence of practice.

It was immediately clear that different institutions used different lists. While there was some overlap in their concept lists, there was appreciable dispersion as well. One practice used a long list of fairly detailed concepts, one used a shorter list of concepts but selected, in some cases, even more specific concepts, and one used a very short list, recording two thirds of their instances with only two codes.

[measured kurtosis]

Candidate systems

System 1 includes 146 hospitals serving X patients. The 718,000 instances of sample data represent the records taken in 2013. This system encoded its values in SNOMED CT in 2006, and currently lists 437 values, of which 12 are not encoded in SCT.

System 2 includes X hospitals serving X patients. The 3.8 million instances of sample data represent the records taken in 20xx. This system currently lists 63 values, all but two of which were encoded in SNOMED CT for this project.

System 3 includes X hospitals serving X patients. The 255,000 instances of sample data represent the records taken in 2013. This system encoded its values in SNOMED CT in xxxx, and currently lists 227 values, of which 36 are not encoded in SCT.

Because the lists are based on interface values, and the standard SNOMED CT values are used only for transmission, there are cases where two interface values may map to a single SNOMED CT code. And even those systems that had been using SCT for years had a small number of mapping issues. The numbers of values

|  |  |  |  |
| --- | --- | --- | --- |
|  | **1** | **2** | **3** |
| total items | 437 | 63 | 227 |
| duplicates | 8 | 5 | 6 |
| problematic | 5 | 0 | 1 |
| unencoded | 12 | 2 | 36 |
| unique concepts | 412 | 56 | 184 |

Most of the concepts in system two were present in system one



Clearly, concepts that are included in set A but not others may simply be the result of greater coverage, whether due to coverage of less common conditions or of more specific breakouts of conditions included elsewhere.

Most of the concepts included exclusively in B, the smallest set, result from that organization’s use of more general classifications, for instance “Adverse reaction to drug.”

Asthma

Acute labyrinthitis

Cough, chronic; Cough after eating; cough at rest

Partly due to granularity; esp kp

Also to ??

Practice: oral edema vs edema of tongue; lip swelling & angioedema

Found & disqual: inferred causes, grouping concepts; situation where disorder exists

Do some tests: can all items in list A be captured in list B somehow?