

ATTACHMENT A

Vocabulary Subsets and Value Sets (Version3 - 2/6/2010)

I.A. General Definitions

- 1) Subset - a set whose members are members of another set; a set contained within another set
- 2) Value set – the set of all possible values for a given purpose.

I.B. HIT Vocabulary Standards Context definitions

- 1) Vocabulary Subset – a terminology set that is a subset of a larger standard vocabulary. The subset can be defined for one or more of the following purposes:
 - a. To produce or contribute to a value set
 - b. To define the universe of relevant values from a standard vocabulary for a particular segment of an EHR or some other purpose, WITHOUT necessarily constraining all possible valid values for that segment or purpose (for example, if a patient’s specific problem cannot be represented using the universe of all disease, injury, symptom, etc. terminology in SNOMED CT, it would still be entered in the patient’s record.)
 - c. Convenience: To ease implementation of standard vocabularies by providing a relatively compact set to facilitate data entry, prioritize one-time mappings from local or legacy terminologies, etc. Two common types of convenience subsets are:
 - i. Frequency-based subsets – based on the level of actual occurrence in large bodies of patient data, preferably from multiple institutions. Examples include the: CORE Problem List Subset of SNOMED CT (http://www.nlm.nih.gov/research/umls/Snomed/core_subset.html) and Frequency-based LOINC subsets for test orders and results (<http://loinc.org/usage>)
 - ii. Specialty-based subsets
- 2) Value set - All valid vocabulary values for a defined purpose. These may or may not be – or contain – subset(s) of one or (possibly) more standard vocabulary (ies). Value sets can be described by reference to part(s) of the structure of a hierarchically organized vocabulary, by an enumerated list, or by a combination of the two, e.g., to describe the specific exclusions. A value set can define:
 - a. All valid values for specific data elements or segments of standard messages
 - b. a parameter needed to identify a specific population within a group of health records. One important purpose for value sets for such parameters is to define the denominator and the numerator for a quality measure.

I.C. Vocabulary TF Subset and Value Set Priorities

- 1) Convenience subsets that can aid in the achievement of “meaningful use” –
A variety of subsets of standard vocabularies may be immediately useful to different providers or hospitals attempting to achieve “meaningful use”. It is appropriate and desirable for the Vocabulary TF to promote the development of an effective “marketplace” of easily identified and obtained subsets of standard vocabularies that are described in a uniform way - so that interested parties can easily locate those that can be useful to them in facilitating “meaningful use” of EHRs. It is NOT appropriate for the Vocabulary TF to develop or anoint particular subsets.
- 2) Value sets defined for “meaningful use” quality measures
- 3) Value sets for parts of message standards that are required for “meaningful use” IF the value set draws some or all of its valid values from standard terminologies
- 4) Agreement on detailed clinical models for some high priority uses

Re: 3 and 4: The US lacks - but needs – an effective mechanism for defining, anointing, and enforcing adoption of uniform value sets for US-wide use in message standards (e.g., HL7, NCPDP, CDISC, etc.) – perhaps especially when the same data are exchanged using different message standards – and for reaching agreement on detailed clinical models.