

Responses for HIT Standards Committee Feb. 23rd Meeting

Brian Levy MD
Chief Medical Officer and Senior Vice President
Health Language

Brian.levy@healthlanguage.com

4600 South Syracuse Street

Suite 1200

Denver, CO 80237

720-940-2940

Cell: 720-308-9363

General comments:

In general, I think of dividing the subset/value set into those that are needed to support HIT infrastructure such as messaging and reporting and those that are useful for facilitating data entry from EMRs. The use of both subsets and value sets enhances interoperability by constraining the terminology concepts that should be used for specific use cases.

I think that subsets will be especially useful for facilitating data entry into EMRs. An example of such subset is the Core SNOMED CT problem list subset. However, it may not contain every finding or procedure concept that certain users and vendors want.

One pitfall of using a standardized terminology such as SNOMED CT is that even though we might all be using a set standard to record problems, we still might not be able to achieve semantic interoperability if we are using the ‘wrong concepts’ or different concepts to mean the same thing. For example, using a SNOMED morphology concept instead of a concept from the finding hierarchy is a common mistake. In another example, if a cardiologist chooses a specific type of myocardial infarction – say “posterior inferior MI” with SNOMED but a GP just chooses “MI”, the SNOMED model will allow us to say that one is a type of another. These multiple levels of granularity are important for data capture, but they must be then ‘rolled up’ to support reporting, data analysis, and decision support. Thus even with the use of subsets, there will likely still be some ambiguity of concept meaning. Terminology services will assist in performing additional translations and terminology navigations to allow understanding of similar concepts.

One final general point is that many of these questions and issues will also pertain to the creation, distribution, and updating of necessary mappings among terminologies as well.

Responses to questions:

1) Who should determine subsets and/or value sets that are needed?

For subsets used in data or order entry within EMRs, standards groups such as the HIT Standards Committee, HITSP, and HL7 will need to identify areas such as problem and procedure lists that would benefit from the use of subsets. In addition, vendors and even end user hospitals and clinicians will need to identify their requirements for localized subsets.

Messaging organizations such as HL7 could determine the value sets that are needed. I think that HITSP also has served an important role for determining where value sets are needed for the CCD. Organizations that require reporting such as CMS can also determine certain value sets that are needed. Since it is unrealistic to assume that there will be one organization to determine when value sets are needed, there will likely be the need for these organizations to collaborate to minimize duplicative and contradictory value sets.

2) Who should produce subsets and/or value sets?

I like the idea of the NLM managing and distributing the SNOMED Core, and receiving feedback from users. I also see medical societies as playing a role in helping to define useful subsets for their specialties. Over time, there could be many dozens of subsets created by many groups. This fact may lead to too many subsets, but it is probably inevitable. Although subsets are currently available from differing sources, the NLM would be a good central repository for finding subsets and value sets produced by these many organizations whenever possible.

National ownership may be challenging in the US – but in the UK there are national efforts to create and manage subsets using medical societies and other interested parties as domain experts. There needs to be clear education – maybe from the NLM and the IHTSDO on how organizations can create subsets ‘correctly’ and to determine best practices. Even with the use of these organizations, there will need to be mechanisms to easily create, collaborate, share, and distribute subsets. Terminology services vendors will play a role here.

Value sets – particularly those required for messaging such as HL7 or the CCD/CDA/CCR, could be determined by the respective messaging organizations. Alternatively or in addition, the standards bodies themselves such as the IHTSDO or Regenstrief could define value sets. But these value sets need to be readily shared – we would not have semantic interoperability if HL7 derived one value set for Race codes that are different than the value set required for the CCD. Lastly, these value sets must be created in a timely manner.

Efforts by individual organizations such as HITSP defining value sets, as well as encouraging the standards bodies to develop their own would facilitate adoption.

3) Who should review and approve subsets and/or value sets?

For subsets, I think it is useful to have a core set, such as the most common findings and procedures, 'validated' by an organization such as the NLM , IHTSDO, or Regenstrief for LOINC. Thus vendors can have some assurance that when they use SNOMED CT or LOINC concepts for example, they can start with a validated set. Of course, there is no definition of what this 'validation' actually means and there will likely be disagreement. Over time, as discussed above, I think there will be many dozens of subsets with a considerable local influence. In addition, these users will need the ability to manage these local lists yet still receive updates from the standards bodies and reconcile them. At a local hospital level, for example, there will likely be governance policies established to manage these subsets. In many cases, these local subsets will be derived from validated larger subsets. Validation from specialty organizations and medical societies will also be beneficial.

4) *How should subsets and/or value sets be described, i.e., what is the minimum set of metadata needed?*

Basic good database practices will be needed. But in addition, the dependent terminologies and versions of the terminologies upon which the subsets are based are critical. Metadata regarding specific change types and history will also be beneficial.

5) *In what format(s) and via what mechanisms should subsets and/or value sets be distributed?*

A common distribution format such as the SNOMED format or Refset format would be useful. A standard machine readable format would make it easier to have vendors automatically consume subsets and value sets created by disparate organizations.

6) *How and how frequently should subsets and/or value sets be updated, and how should updates be coordinated?*

Updating the subsets and value sets are certainly critical to maintain semantic interoperability over time. The harder problem will be ensuring that subsets and pick-lists created locally are updated as the standards change. Local modifications will need to be reconciled with those from the standards bodies to preserve these local changes. Terminology tools and services will be required to support these kinds of local updates.

The frequency of updates will be dependent on the changes in the standards bodies and requirements for their use. Coordination of subset release would certainly be beneficial, but may be unrealistic. Organizations such as the NLM could help to some extent by providing a source for these updates along with the appropriate metadata regarding the update. Subsets will likely be available from many sources and have different update schedules.

7) *What support services would promote and facilitate their use?*

With dozens of potential subsets and value sets in use, one issue will be for vendors, hospitals and users to know when they are available and where to obtain them. Central coordination from the NLM would be useful, but even then users would likely need the support of vendors and terminology service providers to keep track of the terminology subset and value set landscape – particularly for updates.

Also promoting the value of using subsets will help in their adoption. For example, links to decision support from subset concepts, better reporting, improved interoperability, and certification compliance are value points.

8) *What best practices/lessons learned have you learned, or what problems have you learned to avoid, regarding vocabulary subset and value set creation, maintenance, dissemination, and support services?*

Subset creation is often more complicated than initial appearances. If we become bogged down in the semantic issues of creating standard subsets and value sets, the market may not wait for their creation and validation. So we need to be practical and start using subsets and improve them over time with feedback. Expect localizations and build the systems to handle them and the updates to the underlying terminology standards over time.

9) *Do you have other advice or comments on convenience subsets and/or value sets and their relationship to meaningful use?*

Specifications of subsets and value sets are important for testing – particularly CCHIT certifications.

10) *What must the federal government do or not do with regard to the above, and/or what role should the federal government play?*

The government agencies and committees can continue to stimulate, encourage, and drive the standardizations around the use of subsets and value sets. In addition, providing education to vendors and users of the terminologies and subsets will improve their adoption.