

**Written Comments Submitted by
Floyd Eisenberg**

**ONC HIT Standards Committee
Clinical Operations Workgroup - Vocabulary Task Force
Part 1: Measure Developers – Value Set Creators**

September 1, 2010

Good Morning Chairs and Members of the Taskforce. Thank you for providing the National Quality Forum with the opportunity to provide comments on the requirements for the production, maintenance, and oversight of value sets for use in quality measurement and clinical-decision support.

NQF is a public and private partnership with more than 400 members representing virtually every sector of the health care system. NQF operates under a three-part mission to improve the quality of American health care by:

- setting national priorities and goals for performance improvement;
- endorsing national consensus standards for measuring and publicly reporting on performance; and
- promoting the attainment of national goals through education and outreach programs.

NQF is recognized as a private sector standard-setting body under the National Technology Transfer and Advancement Act.

NQF endorsement, which involves rigorous, evidence-based review and a formal Consensus Development Process, has become the “gold standard” for health care performance measures. Major health care purchasers, including the Centers for Medicare & Medicaid Services, rely on NQF-endorsed measures to ensure that the measures are scientifically sound, relevant and help standardize and raise the bar for performance across the industry. To date, NQF has endorsed more than 500 measures.

NQF drives improvements in care by endorsing evidence-based measures of performance-- focusing on measurement for accountability and quality improvement. Measurement has the greatest impact on quality when it supports transparency and public reporting, but also provides actionable information to clinicians to make improvements in care delivery. To date, quality measurement and public reporting has been thought of as a secondary data use versus a driver of care. However, by setting standardized performance measures and properly designing and building HIT, it will now be possible to capture performance data as part of the care process and to provide immediate information feedback and clinical decision-support to clinicians to drive improvement.

Designing and building HIT to support performance improvement requires close collaboration between the “quality community” and the “HIT community.” The “quality community” includes organizations that set practice guidelines, develop performance measures, and set standards for measurement. NQF plays a key role in the “quality community” as the national standard-endorsing body for performance measures and as a neutral convener of multiple stakeholders to recommend National Priorities for

Improvement and advance the quality agenda. The “HIT community” includes HIT suppliers, standard-setting organizations, and users. The DHHS Office of the National Coordinator for HIT is leading the effort and working to establish a Nationwide Health Information Network.

Recognizing the importance of a close collaborative relationship between the “quality community” and the “HIT community,” in 2008, Congress directed HHS to contract with a consensus based organization, such as NQF, to endorse standardized performance measures and to “promote the development and use of EHRs that contain the functionality for automated collection, aggregation, and transmission of performance measure information.” NQF was awarded this contract in 2009 which provides for annual renewal for up to four years.

Recently, NQF has had the opportunity to work directly with measure stewards to retool a set of existing measures to address data obtained through meaningful use of electronic health records. Each of these measures has been endorsed through NQF’s Consensus Development Process. The scope of the project involves preserving the intent of each of 110 measures while providing sufficient detailed electronic specifications to allow implementation in EHRs. To allow flexibility in those implementations and provide guidance for the future, concepts are provided in terminologies specified in EHR certification rules where they are so indicated and also in terminologies recommended by the HIT Standards Committee for future adoption. The first 57 measures retooled as of today include 802 value sets created or updated as part of the effort; 53 measures remain to be completed by the end of the calendar year. Based on NQF’s experience with this project, my comments this morning will address requirements for value set development and governance.

1. What are the requirements for a centralized infrastructure to implement “one-stop shopping” for obtaining value sets, subsets, and vocabularies for meaningful use?

- A. The most urgent requirement is definition of governance and the roles of groups and stakeholders in developing and using value sets. The governance must include:
 - a. Stewardship by content experts in each domain (Quality, Research, Public Health) while providing guidance and infrastructure for appropriate harmonization.
 - Allows for rapid development in content areas by domain experts
 - Avoid stifling domain needs
 - Provide for appropriate harmonization
- Domain content areas include:
- i. Quality – measures, practice guidelines, clinical decision support rules
 - ii. Research – clinical effectiveness research, pharmaceutical research, etc.
 - iii. Public Health

The process should move at the appropriate pace for each domain, balancing speed against harmonization. A carefully designed process can allow emergence of best practices and encourage adoption in each domain area. A decentralized approach within domain areas will be key to initial adoption because domain areas have different needs and are at different stages of readiness to adopt. Over time, greater emphasis can be placed on harmonization across domain areas.

- b. Collaboration with experts in each respective taxonomy to encourage selection and use of appropriate concepts for value sets.
- B. Tools and infrastructure to create and manage value sets:
The Vocabulary Task Force (2/6/2010) defined subsets and value sets as:
- 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.” [Vocabulary Taskforce meeting material, February 23, 2010, Attachment A (available at: <http://healthit.hhs.gov/portal/server.pt?open=512&mode=2&objID=3004&PageID=20395>)].
- Infrastructure is required for both subsets and value sets as defined here. From the standpoint of quality measurement, governance and standardization is most urgent for value sets. Therefore, my testimony will focus primarily on value sets.
Infrastructure requirements include:
- C. Authoring interface – a method for those requiring value sets to select from existing sets when available or define value sets as needed. The authoring interface requires delineation of responsibilities of the author to define, maintain and modify value sets. There must also be the capability to create value sets for concepts in different taxonomies based on the same intent (e.g., cross walk between value sets for diagnosis using ICD-9CM, ICD-10CM, SNOMED). For quality measurement, EHRs are not currently required to incorporate the terminologies provided in the published eMeasure / Health Quality Measure Format (HQMF) directly into their products. The value sets used and the logic define the criteria for each measure. Therefore, a locally employed terminology could be used to determine eligible patients (denominator), appropriate interventions (numerator) and exclusions – as long as the local concepts used are informed by the criteria set in the measure’s value sets. Where options are included within the measures, they can encourage implementation of the terminologies and provide guidance for those exploring the new terminologies for their vendor products (e.g., SNOMED, LOINC, ICD-10). Such options can encourage more rapid adoption of terminologies recommended for future adoption. To provide such options similar to the meaningful use measure retooling project requires a relatively easy to use interface with back-end heuristics to provide comparable results for the value sets created in each terminology.
 - D. Ability to search and reuse existing value sets and recommend modifications and create new ones for each new use *only* when existing sets are insufficient. Each quality measure identifies specific value sets required to define data elements in the denominator, the numerator and the exclusion criteria. Some of these value sets can be reused for other measures. In retooling 57 measures thus far, with three measure stewards, 802 value sets have been created. While each measure steward has reused value sets for their own measures, there is approximately 20% overlap with sets used by other stewards. A standard, centralized value set registry with support from terminology experts can encourage significantly more reuse.
 - E. Ability to version value sets. Standardized timing for applying new versions of each taxonomy to implementations (neither too often nor too infrequent) will be important.

- a. Taxonomies provide updates or new versions at different frequencies. Some taxonomies (e.g., RxNorm) can update weekly; others update at longer intervals. Standardization is required, by domain, for the timing of updates. Updates should also be created centrally by the terminology experts allowing specific domains to determine when to apply the updates for their own use.
- b. Ability to select from a group of value sets and extract all values into a subset to ease implementation into EHRs. Some comments on the published measures for Eligible Providers suggested that a complete set of terms for each context (e.g., diagnoses, procedures, encounters, laboratory tests, etc.) for all of the published measures would be helpful. The request is basically for a subset of all concepts used within the published measures.
- c. Enable a parsimonious set of value sets by domain (e.g., quality, research, public health). Similar to the validation and consensus process used by CDISC to engage comment and approve value sets for use, each domain should be able to tag value sets for use within that domain.
- d. Facilitate cross-domain harmonization.
As value sets for each domain will have application to other domains, the registry function needs to allow tagging of value sets by multiple domains. By using a central registry harmonization can be encouraged.
- e. Ability to request updates from underlying taxonomies with reasonable response times. As new concepts arise, taxonomy experts can provide input as to how they can be expressed with existing modeling. When new concepts are requested, a standard and timely response should be expected.

F. *Which requirements or functionalities are urgent, i.e., absolutely required to support “meaningful use”? Which would be most useful immediately? What would be a staged approach over time to get to the desired end state?*

As noted above, governance is the most urgent to support meaningful use. A central value set repository with the ability to create, update and preferably map to other taxonomies with infrastructure for retrieval is also required.

The process should move at the appropriate pace for each domain, balancing speed against harmonization. A carefully designed process can allow emergence of best practices and to encourage adoption in a measured approach. Begin decentralized within domain areas to allow innovation and tighten the infrastructure and harmonize over time.

Domain content areas include:

- i. Quality – measures, practice guidelines, clinical decision support rules
- ii. Research – clinical effectiveness research, pharmaceutical research, etc.
- iii. Public Health

Detailed Questions

G. *Where are you using value sets and subsets? For what domains? How many value sets and subsets?*

The measure retooling project for meaningful use has thus far created 802 value sets for 57 measures with an estimate of 20% overlap among measure stewards. To move the entire quality enterprise to electronic format as part of the NQF endorsement and maintenance process with approximately 600 measures and averaging 16 value sets per measure with 30% reuse will require 2880 value sets. Creating subsets from among the individual value sets to ease implementation among EHR vendors and local sites will require additional work. Many of the concepts within these measures are also reused in public health and safety reporting and in research. The Structured Care Recommendations for CDS (SCRCDs) project (<http://himssclinicaldecisionsupportwiki.pbworks.com/AHRQ%20eRecommendations%20Template%3A%20Feedback%20and%20Discussion>) sponsored by AHRQ has reused a number of the value sets in the preventive care measures as part of their clinical decision support project. That project highlights the ability to reuse value sets among clinical decision support and quality measures.

H. *In your experience with creating, disseminating, updating and/or using value sets, subsets, and entire vocabularies, what works and what does not work?*

- a. The findings identified in the measure retooling project for meaningful use are by no means new to the terminology community. The high volume, rapid turnaround, large scale (national) implementation of eMeasure (HQMf) format has highlighted the urgency for establishing governance and achieving parsimony.
- b. Value set implementation requires sufficient time for vetting and consensus by terminology experts and clinical domain experts to assure a level of accuracy and precision of the value sets. Testing will also be valuable for validity.
- c. Specific challenges
 - i. Taxonomies are created based on specific use cases. Establishing value sets based on the model and hierarchy of the taxonomy requires careful review to be sure to meet the needs of the intended use. The creator of the value set also needs to understand the structure and meaning of the taxonomy.
 - ii. Value set selection findings
 1. Errors of commission
Including potentially invalid codes due to taxonomy modeling
 - a. Example: Value set to determine all women with a current (active) pregnancy initially selected a branch from the SNOMED-CT hierarchy. Included were concepts for *gravid* (G) and *para* (P), e.g., G1, G2, G3...., P1, P2, P3.... These concepts are not indicative of current (active) pregnancy, therefore the value set had to be constrained.
 - b. Example: Value set to identify patients with a diagnosis of immunodeficiency included “Immunodeficiency of Arabian foals,” a veterinary disease in a human condition branch of SNOMED-CT. Care is required to avoid propagating modeling

issues in underlying taxonomies within value sets used for direct clinical care or secondary uses such as quality measurement.

2. Errors of omission
3. Variations in value set naming (or lack of naming convention)
 - a. Value set naming should sufficiently be consistent with the intended use and the purpose should be defined and include a text definition describing how the concepts in the value set are selected.
 - i. Example experience during the measure retooling project: Selecting a narrower value set name than intended.
Asymptomatic HIV infection – the value set included codes for symptomatic and asymptomatic infection, requiring a change to the name of the value set.
 - ii. Example experience during the measure retooling project: Selecting a broader name than intended.
Pneumococcal vaccination – the value set for pediatric immunization included only vaccine preparations appropriate for children ≤ 2 years; reuse for adult vaccination measure would have omitted valid vaccine preparations, requiring a change to the name of the value set.
 - b. The value set name should specify the information intended
 - i. Example experience during the measure retooling project: Selecting a name that is descriptive and not ambiguous
Conditions indicative of sexually active woman – since the use of the value set was to determine patients who should have Chlamydia screening, the value set name could have encouraged broader utilization as “conditions suggestive of STD risk.”
VZV (Varicella Zoster Vaccine) – the initial term, “Varicella Zoster medications,” included Varicella Zoster Immune Globulin (used for post-exposure prophylaxis but conferring no long term immunity) and inadvertently omitted Varicella Zoster vaccines; the issue has been corrected.

I. *What human resources does it take to implement and manage value sets, subsets, and entire vocabularies? Informaticists? Clinicians? IT people? How are you organized?*

The value sets have been created by the measure stewards. Feedback thus far is that informaticists with terminology expertise are required to appropriately manage value sets. This expertise must be coupled directly with clinical experts in the area of interest. Specific findings thus far in the process include:

- a. Multiple managers of the process add complexity and confusion
- b. The measure steward's input is required as the steward understands best the meaning intended
- c. A standardized source and consultation with content experts is essential – centralization is strongly preferred
- d. Terminology experts must participate to incorporate the appropriate use of the underlying taxonomies
- e. A consensus process for public vetting is preferred to be assured of face validity, accuracy and precision prior to testing
- f. Testing is needed to assure real usage provides the intended results with high positive predictive value and low negative predictive value
- g. Implementation issues must be addressed – if middleware or parsers are needed to map local terms to concepts in value sets, consistent output of these applications may not be standard and may require certification to support the process going forward
- h. Feedback and interaction with the stewards of underlying terminologies is required when modeling issues are identified and when new concepts are required

In closing, this is a complicated process but a very important one that must be done right if we are to produce meaningful information on quality. Thank you for the opportunity to provide this input.