Testimony to National Committee on Vital and Health Statistics (NCVHS) Subcommittee on Standards

ICD010 Planning, Testing, Implementation

New Uses for the ICD-10 Data

John Quinn on behalf of HL7

[Title Slide]

I have come today to explain and discuss with the Committee a set of topics around the use of HL7, SNOMED (and associated US Realm vocabulary standards) and the upcoming adoption and use of ICD-10-CM. I come with the primary perspective of HL7. I am an original founder of HL7 in 1987. I have been continuously active in the organization since then and have for the last five years served as the Organization's CTO f

We are coming close to the time when two events will occur about the same time:

- The US will adopt the use of ICD-10-CM in place of its long standing use of ICD-9 as it relates to both payer and provider communications as well as Public Health data, data analysis and reporting both with this country and out through the World Health Organization and the UN.
- Meaningful Use Stage 2 is now starting to see implementation guides (i.e., that is Draft Standards for Trial Use—DSTUs). This DSTUs will be vetted by "first adopters" and for the first time presents the opportunity for routine exchange of structured clinical information among many different participants in the research and delivery of healthcare. These exchanges will be occurring among healthcare providers, public health and potentially payers. What is new is that they will not be all from the same organizations and the exchange of information will occur routinely among IT systems, architectures and healthcare data producing and using organizations that are independent of each other

[Slide 2]

HL7 does have some core assumptions about its role and the nature of the data that it communicates among IT systems. They are:

HL7 makes the following assumptions of the systems that deal with health information and requires data that:

- Is recorded at the appropriate level of detail—not too general or too specific
- Is consistent over time and across jurisdictions
- Can be transmitted without loss of meaning
- Can be aggregated at more general levels
- Can be integrated by automated systems

[Slide 3]

HL7 does not typically represents patient data through methods that it creates. Rather it relies on terminologies and classification schemes that are developed and maintained by other standards development organizations who specialize in this area. HL7 presently has affiliates in about 34 other countries outside the United States. The set of patient data representation standards that are used in the United States is not for the most part the same in each country. In the Unites States we refer to this as the US Realm of terminology and coding standards as set by Office of the National Coordinator for Health Information Technologies (ONC) and the National Library of Medicine (NLM).

We expect (and design our standards to facilitate the representation of patient data via:

- Reference Terminologies such as SNOMED CT® and LOINC®
- High Level Classification Systems such as ICD-9-CM (soon to be ICD-10-CM), ICF (International Classification of Functioning, Disability and Health), etc.
- Specialized Classification/Taxonomy systems such as DPT and CMG
- Within the context of HL7 and other healthcare reference information models such as ISO, HL7, open HER, etc.

[Slide 4]

HL7 does not create its own or itself attempt to modify vocabularies, reference terminologies, taxonomies, etc. In fact a principle within our Vocabulary Work Group is to avoid creating our own standards to represent/hold patient data unless there is no suitable external vehicle available.

HL7 does have largerly internal and not normally visible internal code sets (i.e., “meta data”) that identifies HL7 modeling, messaging, document, functional standard, etc. components. (e.g., the message segments, data type names, information model pieces such as classes, , etc.) In many cases these components that we do create are part of ISO TC-215’s (medical Informatics) standards that were submitted and accepted into ISO TC-215 from HL7 and through the American National Standards Institute (ANSI)

[Slide 5]

HL7 relies on terminology standards development organizations (e.g., IHTSDO, LOINC, ISO) to supply and support the terminology standards that are necessary to unambiguously represent the content of an exchanged message, document, etc.

This is a two way path that we travel. Over the years, HL7 has also made significant investments in its standards to that they can properly accommodate the major coded international classification and terminology sets that are needed by our customers. (e.g., Common Terminology Services-2, HL7 Data Types, etc.

[Slide 6]

This slide gives a pictorial HL7 view of this environment as we present it in our Vocabulary Training tutorial.

[Slide 7]

The National Center for Health Statistics (NCHS) has developed what is known as a “General Equivalence Mappings” (GEMs) for the diagnosis codes. The Centers for Medicare & Medicaid Services (CMS) have developed the GEMs for the procedure codes. The GEMs are considered to be the authoritative source for cross-walking between ICD-10 and ICD-9.

The GEMS for ICD-9 and ICD-10 even at a high level gives us some level of pause and concern about ICD-9 and the implications of ICD-10 for anything beyond its historical and continues use.

[Slide 8]

The National Center for Health Statistics (NCHS) has developed what is known as a “General Equivalence Mappings” (GEMs) for the diagnosis codes. The Centers for Medicare & Medicaid Services (CMS) have developed the GEMs for the procedure codes. The GEMs are considered to be the authoritative source for cross-walking between ICD-10 and ICD-9. (CPT® codes remain the codes physicians will use to report services and procedure performed in the office and ambulatory settings).

[Slide 9]

The table shown here well points out what is happening. ICD-9-CM mapping to ICD-10 is going from a classification scheme of with n classifications to a scheme of n+x. The concept of “automatically” mapping does not appear to be good when taken in the context of “what if we used ICD for coding the physician observed clinical disease diagnosis.

When looking at “1 to 1 approximate match with one choice going in case of “many to one” that is ICD-10 to ICD-9 we see a “success rate” of 82.6%...while this might have been a “ok” grade in grammar school, I wouldn’t personally think too highly of my physician if that was his or hers’ track record. The real concern in looking at this is we are now using ICD-9 and going to ICD-10. The 1 to 1 “approximate” match with 1 choice is only 49%. In all cases when I talk to one of my terminologies in HL7 (and we have many of the most well published in the country) a discussion about ICD9 to ICD10 mapping and (even worse) ICD10 to ICD9 mapping (many to one) creates a “suffering look” on their face and we have a hard time moving the conversation forward.

Please understand I am not seriously concerned about this because, in general, HL7 is dealing with clinical data and no provider organization or individual practitioner that I have dealt with is using ICDx for clinical documentation.

[Slide 10]

In addition to the conundrum created by having one to many and many to one decisions to make, It is important to also understand that a many to one mapping produces a loss of information. Unless our software can save all information before a mapping occurs with the ability to "roll-back" to the pre-decision" machine state…data is lost! It is also not commutative (i.e., one to many transforms does not produce the same result as a many to one transform with the same value.

[Slide 11]

Let's remember that "Meaningful Use" Stage 2 creates an opportunity for significant use of coded clinical data expressed in SNOMED, LOINC and other terminology standards for uses such as Consolidated CDA and uses derived from required Consolidated CDA Template Documents, Laboratory Results Interface, Syndromic Surveillance, Immunization registry, etc. We are starting to see the availability of structured coded data that can provide a very high level of semantic interoperability (i.e., the computer systems can be programmed to essentially have the same "understanding" of the set of structured information.

(if you have an older copy of my slide deck, Slide 12 is a content duplication of slide 11 and should be skipped)

[Slide 12]

An example of a challenge in Public Health is that the current Meaningful Use Stage 1 and 2 biosurveillance activities could map the existing ICD-9 syndrome and sub-syndrome classification to ICD-10.
However, as the use of SNOMED becomes more widespread the sub-syndrome categories will likely become SNOMED instead of ICD. This will link better with the coded data in the EMR and biosurveillance will likely come to rely on the EMR data rather than the claims data as it does now.

It is reasonable to believe that where more data is available in the EMR it will likely start to supplement the existing claims sourced data and we will have to either deal with the stark differences between the classification coding structure coming from the claims based data vs. the EMR sourced data and its abilities to connect between the recorded observation ton to the Medical Knowledge representation that is also carried along in the SNOMED clinical terminology system.

[Slide 13]

Considerable recent conversations have occurred about enabling the "bigdata"/data analytics promise of EMR sourced anonymized data…predicated on the availability of structured clinical documentation of patient complaints and observations.

In other words in order to have data in the EMR that will present potential value to the proposed anonymized data mining, IT MUST FIRST BE ENTERED INTO THE EMR AS STRUCTURED CODED DATA. While is routinely possible in automated testing environments (e.g., lab tests), it will be a bigger hurdle

in itself to get the correct documentation entered into the EMR by the observing physician at the point of examination.

[Slide 14]

So it gets more complicated! It is not just an exchange of information from the Provider's Billing System with the Payer. The Payers will soon have HIPAA attachment enabled transactions for claims documentation and pre-authorization requests. The Payer systems are satisfied with a modest increase in complexity by moving from ICD9-CM to ICD10-CM. The clinical data available in the EMR system to support the Payer's need for clinical information will be arriving in SNOMED, LOINC, etc. form..

With the advent of discussion about interoperating SNOMED, LOINC Structured Clinical Documentation with Payers the mapping problem now includes SNOMED and LOINC (at a minimum) with ICD-9-CM and ICD-10-CM

[Slide 15]

The NLM has done a study and created a document on a potential CNOMED CT to ICD-10-CM mapping. The work appears to be a collaborative effort between Dr. James Campbell of Nebraska Medical Center (and an long-term developer of SNOMED) and Dr. Kin Wah Fung of the NLM

[Slide 16]

The NLM study looks into what does it mean to have "semi-automatic generation of ICD-10-CM classification codes sources from clinical data encoded in SNOMED CT. Could we use SNOMED CT data to generate a better ICD-10-CM code for a billing diagnosis.

Even though reimbursement is expected to be the main use case, the ICD-10-CM codes suggested by the Map are not guaranteed to be reimbursable, and the codes are not optimized to provide the highest level of reimbursement. The provider may have more information about the patient's condition and reasons for charges to the patient's payer but is their value if it doesn't result in a better payment experience?!

[Slide 17]

Going in Presumptions:

This is NOT:

- A completely automated ICD-10-CM coding from a SNOMED CT source
- A map that supports management of context beyond that found in the coded record and ICD-10-CM conventions and rules as noted in General approach and Heuristics

…Let's just assume it's about "what's possible"

[Slide 18]

Our Mapping Assumptions in this examination…

- Since ICD-10-CM is a classification, the semantic space of a particular classification code depends upon the definition of sibling codes and others within the same category. Since ICD-10-CM is designed for statistical, epidemiological and reimbursement purposes, one ICD-10-CM classification code may include many SNOMED CT concepts within its semantic space.
- *Does it make sense to assume that EMR systems software (or even payer systems software) will employ rules to take a given SNOMED coded concept to map it to a unique ICD-10-CM code?*

[Slide 19]

- Even if we solved all of the apparent ambiguity, is it reasonable to assume that something between adequate-to-excellent semantically correct computer communications could occur between provider and payer?

We might want to assume that if more information is an advantage than maybe the provider has an advantage over the payer in the discussion of reimbursement?...doesn't seem likely

[Slide 20]

Can the Map be used to automatically generate ICD-10-CM codes?

The Map is designed to find out what is possible with by two researchers that have a whole lot more understanding about the data they have to work with than is average for the industry. But it does raise the question and maybe there is value here in the future.

- The Map is intended to be used in a semi-automatic manner. The Map will suggest candidate ICD-10-CM codes based on SNOMED CT codes and, if applicable, additional information obtained from the electronic patient record or direct user input. Review of the candidate ICD-10-CM codes by either the healthcare provider or professional coder is recommended.

  The Map contains map rules and advice that can be used to highlight specific coding principles, or point to additional information required for coding. For the simple cases with one-to-one mappings, it is likely that a high proportion of the candidate ICD-10-CM codes will be adopted by the reviewer.

[Slide 21]

For your reference there are today 22 member countries in IHTSDO with a very wide variety in population size.

Are there any questions and discussion beyond what has already occurred?

Thank You