Clarification of Clinical Data Sets, Vocabularies, Terminologies, and Classification

by AHIMA's Coding Policy and Strategy Committee

[Please note: the content in this article is out of date. It is being updated by AHIMA's Clinical Terminology and Classification Practice Council]

Clinical data set, vocabulary, terminology, and classification system. We frequently hear, see, and speak these terms, but what do we know about them? How do they interrelate? How are they similar? How are they different?

One of the goals of the Coding Policy and Strategy Committee (CPSC) for 1998 was to clarify the distinction between clinical data sets, vocabularies, terminologies and classification systems. The committee believes that this is an extremely important goal, as many of these terms tend to be used interchangeably without thought to their true definition or meaning.

This article summarizes the CPSC's discussion of these entities, as well as a diagram that illustrates the interrelationships between these terms.

Data Set

A data set is a collection of data elements such as age, principal diagnosis, or level of functioning. Each data element should have a single stated definition and purpose. It should also have a unique name or data dictionary established. A data set should also have some interrelationship between the data elements that are defined. Examples of a data set could include something as simple as a master patient index to something more complex such as the Minimum Data Set for long term care.

Clinical Vocabulary

A clinical vocabulary is a list or collection of clinical words or phrases with their meanings. Examples of clinical vocabularies include the Systematized Nomenclature of Human and Veterinary Medicine (SNOMED) from the College of American Pathologists, the Read Codes from Britain's National Health Service, and the Unified Medical Language System (UMLS) from the National Library of Medicine.

Clinical Terminology

A clinical terminology provides the proper use of clinical words as names or symbols. It is a system of clinical terms of preferred terminology or a nomenclature. Examples include defined terms such as a neuroma or colitis, abbreviations such as SLE for systemic lupus erythematosus or CMV for cytomegalovirus, and synonyms of terms.

Classification Systems

A classification system is clinically descriptive and arranges or organizes like or related entities for easy retrieval. Examples include ICD-9-CM, CPT, ICD-10-PCS, and the North American Nursing Diagnosis Association Taxonomy (NANDA).

Interrelationships

Figure 1 illustrates the relationships between a clinical data set, vocabulary, terminology, and classification system. It demonstrates that a data set is used at all levels of the hierarchy, with the lowest level being a vocabulary and the highest level being a classification system. It also demonstrates that once a clinical vocabulary is established it may be used to form a clinical terminology, which then may be used to form a classification system, although this may not always be the case.

The accuracy of the translation from clinical vocabularies to clinical terminologies to classification systems has not been evaluated very well. This translation is highly error-prone. The CPSC recommends that a major role of the HIM professional is to examine the translation among these entities and to work closely with vendors to make sure that linkages between different levels of granularity are appropriate.

The CPR and Classification Systems

The computer-based patient record (CPR) will not cause classification systems to become obsolete. In fact, the CPR may actually promote the use of certain classification systems in the following areas:

- clinical pathways
- outcomes research
- epidemiology
- education
- cost of care

Aggregation of clinical data will remain valuable because it allows the data to be more manageable. Clinical data sets, vocabularies, terminologies, and classification systems will have unique roles in a computerized environment. It is the responsibility of the HIM professional to work toward promoting these entities in the computerized environment. With input from the HIM professional, these entities will not lose their true meaning but will be used appropriately and will eventually improve the overall quality of healthcare.

The Role of HIM Professionals

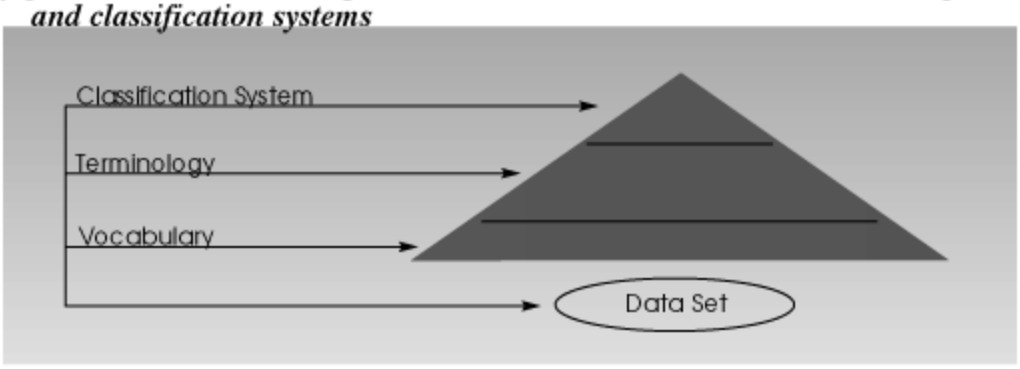
In the future, computers will perform coding and classification. However, the role of the HIM professional will become even more necessary, as he or she maintains, interprets, and analyzes data related to each of these entities (i.e., data sets, clinical vocabularies, terminologies, and classification systems). The first step in this process is for HIM professionals to become experts in how these entities are similar and unique and how they interrelate. Once this is mastered, one can effectively and efficiently manage all types of healthcare systems.

Seeking Compliance Plans

If you are an HIM professional whose facility or department has an HIM-oriented compliance plan in place, AHIMA would like to hear from you. Send us a copy of your plan and we will consider it for placement it on our Web site (with identifying characteristics removed) to use as a model for others seeking to design their own plans. Only plans submitted via e-mail in text or Microsoft Word format will be accepted. Please send all submissions to jsegal@ahima.org. If you have any questions, please call Jack Segal, AHIMA's director of public relations, at (312) 573-8508.

---

figure 1—interrelationships between data sets, vocabularies, terminologies, and classification systems



Article Citation:

AHIMA Coding Policy and Strategy Committee. "Clarification of Clinical Data Sets, Vocabularies, Terminologies, and Classification." *Journal of AHIMA* 70, no. 2 (1999): 72-73.