caBIG™ cancer Biomedical Informatics Grid ™

An Initiative of the National Cancer Institute

CTCAE version 3
Vocabulary Standards Review
"Review of the Review"

Presented by

Salvatore Mungal | Duke University
Stuart Turner | University of California, Davis

Based on work by Jim Cimino and the CTCAE Review Team (circa 2007)

Friday, 11 July 2008

Overview

Note: The evaluation of the CTCAE (Common Terminology for Adverse Events) version 3.0 was performed late 2006 and presented to the caBIG Workspace in January and April 2007. This is a summary of that review performed by **Jim Cimino** and the CTCAE review team. Much of the following information is derived from that report and the source documents can be referenced from Vocabulary Standardization section of GForge.

Overview

- The “native” version of CTCAE version 3.0 was available as a text document encoded in PDF and browsers/coders
- No clear delineation of “terms”
  - No term identifiers
  - Adverse Events (AEs):
    - terms in their own right?
    - only when postcoordinated with grades?
  - Are there 5 grades or ~ 5000?
    - Grades definitions (within the same grade) are inherently unique due to contextual dependency on the AE term
- Therefore, no “*Finite, enumerated set of terms to convey information unambiguously*”
- Evaluation using standard criteria is problematic

Vocabulary Review Criteria

caBIG cancer Biomedical Informatics Grid™
An Initiative of the National Cancer Institute

	A	B	C	D	E	F	G	H	I
1	Vocabulary Review Criteria - version 3.3				Terminology Assessment				
2					meets criterion	partially meets criterion	does not meet criterion	criterion not applicable	criterion not assessed
3	**A. Structure** – criteria related to the data model of the terminology	**A.1. Concept orientation** –Is terminologic information organized around meaning of terms?	**A.1.a. Does each concept have a single, coherent meaning?**	**A.1.a.i. Does each term correspond to at least one meaning?** (nonvagueness)					
4				**A.1.a.ii. Does each concept correspond to no more than one meaning?** (nonambiguity)					
5				**A.1.a.iii. Does each meaning correspond to no more than one concept?** (nonredundancy)					
6			**A.1.b. Does the terminology support synonyms and is synonymy explicitly represented?**						
7		**A.2. Concept permanence**	**A.2.a. Is the meaning of a concept, once created, inviolate?**						
8			**A.2.b. Does the data model accommodate name changes and retirement?**						
9		**A.3. Nonsemantic concept identifiers**	**A.3.a. Does each concept have a unique identifier?**						
10			**A.3.b. Are identifiers are free of hierarchical or other implicit meaning?**						
11			**A.3.c. Are identifiers NOT re-used when a concept is made obsolute or is superseded?**						
12		**A.4. Polyhierarchical organization** - Is it allowed? Is it appropriate?	**A.4.a. Is the basic principle for any hierarchical arrangement explicitily stated?**						
13			**A.4.b. In the case of a polyhierarchy, is any concept capable of having multiple semantic parents?**						
14			**A.4.c. In the case of a polyhierarchy, does each concept have the same meaning regardless of the parent from which it is reached?**						
15		**A.5. Graceful evolution** - How are updates applied to the content?	**A.5.a. Are there clear detailed descriptions of what changes occur and why?**						
16			**A.5.b. Are updates and modifications referable to consistent version identifiers?**						

NB: At the time of the CTCAE Review in 2006, the vocabulary standards review was in version 2.0.

Changes have been made and it is now in version 3.0 (*shown to the left*)

Vocabulary Review Criteria

caBIG cancer Biomedical Informatics Grid™
An Initiative of the National Cancer Institute

Structure
The overall “data model” of the terminology

Content
The extent of domain coverage within the terminology as well as representation via textual definitions, use of “not elsewhere classified” (NEC) terms, polyhierarchy, formalism regarding concept usage, etc.

Documentation
Purpose and scope; statement of intended use; description of usage of codes/identifiers, output formats, use of semantic relationships; tooling available, etc.

Editorial Process
Curation process; concept permanence; QA and QC; extensions to other terminologies, etc.

Vocabulary Review Approach

CTCAE v3.0 was *reified* for representation within the NCI Thesaurus

Three approaches were considered…

1. AEs are terms, 5 Grades are modifiers
    - Link AEs to allowable Grades
    - Most compact form
    - Loses context-specific meanings of Grades
2. AEs are terms, AE-specific Grades are terms
    - Link through “has-grade” relationships
    - Allows reuse of Grades
    - Useful with Supra-Ordinate groups
    - Care must be exercised (Death=Death?)
3. **Precoordinate AEs and allowable Grades**
    - Simplifies relationships to “is-a”
    - Largest possible version
    - No re-use, but reduces redundancy/ambiguity

Vocabulary Review Approach

caBIG cancer Biomedical Informatics Grid™
An Initiative of the National Cancer Institute

NCIt Reification

1. Concept is unit of discourse
2. Concepts have unique identifiers
3. Terms are from sources, mapped to concepts
4. Information available in various file formats
5. CTCAE has been incorporated into NCIT
6. Precoordination approach taken
7. NCIT Flat File, XML and OWL versions

NCIT Flat File

C57213<tab>CTCAE_Grade_1_Nausea<tab> Nausea_Adverse_Event<tab>
CTCAE Grade 1 Nausea| Grade 1 Nausea

NCIT XML

```
<conceptDef>
        <name>CTCAE_Grade_1_Nausea</name>
        <code>C57213</code>
        <id>57213</id>
        <namespace>NCI</namespace>
        <primitive/>
        <kind>Findings_and_Disorders_Kind</kind>
        <definingConcepts>
        <concept>Nausea_Adverse_Event</concept>
        </definingConcepts>
        <definingRoles></definingRoles>
        <properties>
        <property><name>Preferred_Name</name>
                <value>CTCAE Grade 1
Nausea</value></property>
        <property><name>Semantic_Type</name>

        <value>Finding</value></property>
        <property><name>Synonym</name>
                <value>CTCAE Grade 1
Nausea</value></property>
        <property><name>Synonym</name>
                <value>Grade 1
Nausea</value></property>
```

CTCAE in NCIt

- Naming conventions
  - Categories: “Adverse Event Associated with X”
  - Supra-Ordinates: “X Adverse Event”
  - AEs: “X Adverse Event”
  - Grades: “CTCAE Grade n X”
- Stats
  - 28 Categories
  - 5 Grade terms
  - 51 Supra-Ordinate terms
  - 1,043 Aes
  - 4,472 AE-Grade pre-coordinations
- Included content
  - Codes
  - Semantic types (all are “Finding”)
  - Labels and Preferred Names
  - Subclass_of
  - Synonyms and Definitions

Vocabulary Review Approach

Therefore, there were *two* instances of the terminology…

- Native CTCAE version as a text document
- NCIt version

The review uniquely performed the evaluation of ***both*** in parallel. This comparison is an important outcome of that review and the basis of comparative recommendations made.

These recommendations remain important to the evolution to version 4.0, although some may be less relevant or perhaps even obsolete given recent goals or objectives in the design of the next version.

Vocabulary Review Approach

caBIG cancer Biomedical Informatics Grid™
An Initiative of the National Cancer Institute

	A	B	C	D	E	F	G	H	I	J
1					Vocabulary Assessment					
2	Evaluation of CTCAE Using Revised Vocabulary Review Criteria -- VRC version 2.0				meets criterion	partially meets criterion	does not meet criterion	criterion not applicable	criterion not assessed	Documentation
3	**Understandability, Reproducibility and Usability (URU)** — Does the vocabulary conform to the highest degree practicable to the standards of good vocabulary practices laid out by the community?	**purpose and scope**	Is the purpose and scope of the vocabulary clearly stated in operational terms so that its fitness for particular purpose can be assessed and evaluated?		CTEP NCIT					See CTEP publications
4		**vocabulary content coverage**	Does the vocabulary provide comprehensive or explicit in-depth coverage of the domain of interest it claims to address as stated in purpose and scope of the vocabulary segment?		CTEP NCIT					Reasonable domain coverage; NEC terms cover gaps; but no formal evaluations available
5			Are there formal methods in place for expanding and refining the vocabulary?		NCIT		EP			CTEP has no plans to modify CTCAE; NCIT would encompass changes if they occurred
6			Are there explicit, reproducible methods for recognizing and filling gaps in				CTEP		NCIT	See 5J
7		**concept orientation**	Does each concept have a single, coherent meaning?	Does each term correspond to at least one meaning	NCIT		CTEP			NCIT is concept oriented; CTCAE is published coding system
8				correspond to no more than one meaning	NCIT		CTEP			See 8J
9				correspond to no more than one concept	NCIT		CTEP			See 8J
10			Does the vocabulary support synonyms and is synonymy explicitly represented?		CTEP NCIT					Available in published and electronic formats
11			Are relationships between concepts uniform within the vocabulary (internal		NCIT		CTEP			NCIT has logical representation; CTEP does not
12		**concept permanence**	Is the meaning of a concept, once created, inviolate?		NCIT		CTEP			See 8J
13		**nonsemantic concept identifiers**	Does each concept have a unique identifier?		NCIT		CTEP			See 8J
14			Are identifiers are free of hierarchical or other implicit meaning?		NCIT		CTEP			See 8J
15			Are identifiers NOT re-used when a concept is made obsolute or is superseded?		NCIT		CTEP			See 8J
16		**hierarchical organization**	Is the basic principle for any hierarchical arrangement explicitily stated?		NCIT		CTEP			Published version has no hierarchy; NCIT uses subclass of relationships
17			Is the vocabulary organized as a polyhierarchy?			NCIT	CTEP			Published version has no hierarchy; NCIT can support multihierarchy but uses it sparingly
18			In the case of a polyhierarchy, is any concept capable of having multiple			NCIT		CTEP		See 17J
19			In the case of a polyhierarchy, does each concept have the same meaning regardless of the parent from which it is reached?		NCIT			CTEP		See 17J
20		**formal definitions**	Does each term have a definitive set of relationships to other concepts that, taken together, are both individually necessary and collectively sufficient to distinguish the concept from all other concepts?				CTEP NCIT			No semantic relations except hierarchy
21			Does the vocabulary make explicit which concepts are atomic (i.e. non-composite)?				CTEP NCIT			See 20J
22			Does the vocabulary support the creation of composite concepts?				CTEP NCIT			See 20J

Understandability, Reproducability, Usability

caBIG cancer Biomedical Informatics Grid™
An Initiative of the National Cancer Institute

	CTEP	NCIT
Purpose and Scope	+	+
Vocabulary Content Coverage	+	+
Concept Orientation	-	+
Concept Permanence	+/-	+
Nonsemantic Concept Identifier	-	+
Polyhierarchy	-	+/-
Formal Definition	-	-
Explicitness of Relations	+/-	+
Rejection of NEC	-	-
Multiple Granularities	+	+
Multiple Consistent Views	+	+
Context Representation	-	-
Graceful Evolution	-	+
Recognize Redundancy	-	-

Silver Level Criteria

1. Describe the essential nature of the concept – Yes
2. Concise, precise and unambiguous – Generally Yes
3. However, some parts of definitions may themselves be ambiguous. For example, although used frequently, "major urgent intervention" is not explicitly described. So, for example, would biliary tree hemorrhage requiring transfusion of 10 units of blood be considered Grade 3 ("Transfusion, interventional radiology, endoscopic, or operative intervention indicated") or 4 ("major urgent intervention indicated")?
4. Avoidance of rationale, functional usage or procedural information - Yes
5. Consistent terminology and logical structure – Yes
6. Description logic relationships to other terms in the terminology – No

Review Summary

Evaluation of CTCAE Using Revised Vocabulary Review Criteria -- VRC version 2.0				Vocabulary Assessment				
				meets criterion	partially meets criterion	does not meet criterion	criterion not applicable	criterion not assessed
			CTEP	26	14	54	4	1
			NCIT	68	14	12	3	2
				criteria met	criteria partially met	criteria not met	criteria not applicable	criteria not assessed

Summary and Conclusions (2007)

- CTCAE is not a true controlled terminology
- CTEP version of CTCAE does not meet most criteria
- NCIT succeeds in providing CTCAE as a terminology
- NCIT construction of CTCAE *meets most* criteria
  - Polyhierarchy could use work
  - NEC is not so bad (if used sparingly in practice)
  - Formal definitions, context representation, & redundancy detection hard
- Some semantics are lost
- A few inconsistencies were found
- AE-Grade names unhelpful
- Codes should be used as pointers
- Separate files would be nice
- Content maintenance is an issue
- Formal evaluations of content lacking
- Reconciliation with MedDRA is an issue