Developing Effective Quality Systems in Clinical Trials: An Enlightened Approach

NCI Data Quality Initiatives & Enterprise Vocabulary Services

October 13, 2010

Edward Helton, NCI Imaging Program
Larry Wright, NCI Enterprise Vocabulary Services
Late 1990s cancer research enterprise data quality & tracking

- Trials diversity in running, reporting and analysis
  - Treatment, prevention, screening, supportive care, etc.
  - Many sponsors, sites, systems

- Related research activities even more diverse
  - Basic research, clinical care, cancer registries, etc.
  - Part of broader communities: federal, health care, etc.

- Dr. Klausner wanted
  - A broad, consistent understanding of research activity
  - Sharper, better integrated research effort
EVS born in 1997 as part of this effort, key thrusts:

1. Analyze and organize diverse coding
   - LexEVS and other terminology services

2. Map meanings between different terminologies
   - NCI Metathesaurus (NCIm)

3. Reference terminology with high quality meaning shared between people and computers
   - NCI Thesaurus (NCIt)

EVS provides the semantic foundation for sharing and re-use of data and services – interoperability.
Interoperability: The ability of two or more systems or components to exchange information and to use the information that has been exchanged.

NCI Design for Interoperability

- **Common API Integration**: Part of the syntactic component of interoperability.

- **Vocabularies/Terminologies/Ontologies**: Provides semantic interoperability, used to record information in and about systems and data.

- **Data Elements**: Metadata provides a description of the meaning of information in addition to its value. For example, *Performance Status* would describe both a meaning and what constitutes a valid value.

- **Information Models**: Describe the information design and structure of the data maintained in a system.
No Controlled Terminology?
No Interoperability

- Systems cannot exchange or use information if they use incompatible codes or tokens to signify meaning.
- Terminology services provide those tokens, codes and compatibility.
- Proper use of terminology assures consistent meaning across and among enterprises.
- Requires content and technical quality, with a big dose of harmonization and consensus building.
**Semantic Type: Quantitative Concept**
**Code:** C42539
**Definition:** A unit of amount of substance, one of the seven base units of the International System of Units (Systeme International d'Unites, SI). It is the amount of substance that contains as many elementary units as there are atoms in 0.012 kg of carbon-12. When the mole is used, the elementary entities must be specified and may be atoms, molecules, ions, electrons, other particles, or specified groups of such particles.

**Semantic Type: Mammal**
**Code:** C14876
**Definition:** A small, furry creature of the family Talpidae that lives underground and feeds on small invertebrates. The mole has tiny covered eyes that are believed to be able to distinguish night from day, and not much else.

**Semantic Type: Neoplastic Process**
**Code:** C7570
**Definition:** A neoplasm composed of melanocytes that usually appears as a dark spot on the skin.

**Semantic Type: Occupation or Discipline**
**Definition:** [No use case for this term yet, but welcome CIA inquiries].

**Semantic Type: Food or Food Product**
**Definition:** [No use case for this term yet, but welcome inquiries accompanied by samples].
NCI Thesaurus (NCIt)

- Reference terminology/ontology for clinical, research and other NCI and caBIG systems.
- Recognized standard for biomedical coding used both nationally and internationally.
- Public domain, open content license.
- Built by integrated team of subject matter and technology experts using description logics.
- Built in active collaboration with wide range of partners and stakeholders.
NCIt Community & Partnerships

- **NCI Internal Outreach & Integration**
  - PDQ Cancer Clinical Trials Registry & NCI Drug Dictionary
  - CTEP therapeutic trials
  - DCP prevention trials
  - Many others

- **External outreach**
  - Other NIH including NLM, NHLBI, NICHD, NIDCR, BTRIS
  - Interagency: FDA, VHA, CDC, other
  - SDOs: CDISC, NCPDP, HL7, WHO

- **caBIG drew in Cancer Centers & others, public & private**
  - Standards and frameworks for interoperability
  - Building applications and resources
**imatinib mesylate**

The mesylate salt of imatinib, a tyrosine kinase inhibitor with antineoplastic activity. Imatinib binds to an intracellular pocket located within tyrosine kinases (TK), thereby inhibiting ATP binding and preventing phosphorylation and the subsequent activation of growth receptors and their downstream signal transduction pathways. This agent inhibits TK encoded by the bcr-abl oncogene as well as receptor TKs encoded by the c-kit and platelet-derived growth factor receptor (PDGFR) oncogenes. Inhibition of the bcr-abl TK results in decreased proliferation and enhanced apoptosis in malignant cells of Philadelphia-positive (Ph+) hematological malignancies such as CML and ALL; effects on c-kit TK activity inhibit mast-cell and cellular proliferation in those diseases overexpressing c-kit, such as mastocytosis and gastrointestinal stromal tumor (GIST). Check for [active clinical trials](https://clinicaltrials.gov) or [closed clinical trials](https://clinicaltrials.gov) using this agent. ([NCI Thesaurus](https://ncit.nci.nih.gov/ncitbrowser/

**US brand name:** Gleevec  
**Foreign brand name:** Glivec  
**Code names:** CGP 57148  
**Chemical structure name:** 4-[[4-methyl-1-piperazinyl]methyl]-N-[4-methyl-3-[[4-(3-pyridinyl)-2-pyrimidinyl]amino]-phenyl]benzamide methanesulfonate
Anemia (Code E10010)

Terms and Properties

Definition: A disorder characterized by a reduction in the amount of hemoglobin in 100 ml of blood. Signs and symptoms of anemia may include pallor of the skin and mucous membranes, shortness of breath, palpitations of the heart, soft systolic murmurs, lethargy, and fatigability.

Preferred Name: Anemia

Label: Anemia

Common Terminology Criteria for Adverse Events Code: E10010

Synonyms & Abbreviations: (see Synonym Details)

Anemia

External Source Codes:

- NCIt Code: C2869 (see NCI Thesaurus info)
- MedDRA Code: 10002272 (see MedDRA info)
NCI-FDA Terminology Collaboration

• 2001- partnership and agreements in several terminology areas.
  - Structured Product Labeling (SPL)
  - Unique Ingredient Identifier (UNII)
  - Regulated Product Submission (RPS)
  - Individual Case Safety Report (ICSR)
  - Center for Devices and Radiological Health (CDRH)
  - Over 15,000 terms in more than 20 NCIt subsets

• FDA PDUFA IV IT Plan:
  “For terminology standards, the FDA partners with the National Cancer Institute Enterprise Vocabulary Services (EVS). The NCI EVS hosts the FDA terminologies and makes them freely available to the public.”

• FDA terminology resources are available on the NCI portal website: http://www.cancer.gov/cancertopics/terminologyresources/FDA
NCI-CDISC Terminology Collaboration

- **2002- partnership and agreements in several terminology areas.**

<table>
<thead>
<tr>
<th></th>
<th>Numbers</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CDISC Controlled Terminology</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDTM</td>
<td>3745</td>
<td>Published</td>
</tr>
<tr>
<td>CDASH</td>
<td>82</td>
<td>Published</td>
</tr>
<tr>
<td>ADaM</td>
<td>29</td>
<td>Published</td>
</tr>
<tr>
<td>SEND</td>
<td>1928</td>
<td>Pilot</td>
</tr>
<tr>
<td><strong>CDISC (SDTM+CDASH) Metadata</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domains</td>
<td>45</td>
<td>Published</td>
</tr>
<tr>
<td>Observation Classes</td>
<td>200</td>
<td>Harmonization</td>
</tr>
<tr>
<td>Variables</td>
<td>654</td>
<td>Harmonization</td>
</tr>
<tr>
<td><strong>CDISC Disease Standards</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>25</td>
<td>Mapping in EVS</td>
</tr>
<tr>
<td>Tuberculosis</td>
<td>130+</td>
<td>Mapping in EVS</td>
</tr>
<tr>
<td>Polycystic Kidney Disease</td>
<td>200+</td>
<td>Mapping in EVS</td>
</tr>
</tbody>
</table>

- **CDISC terminology resources are available on the NCI portal website:** [http://www.cancer.gov/cancertopics/terminologyresources/CDISC](http://www.cancer.gov/cancertopics/terminologyresources/CDISC)
Concept Code:
A unique, permanent identifier

<table>
<thead>
<tr>
<th>Term</th>
<th>Source</th>
<th>Type</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>mole(s)</td>
<td>ICH</td>
<td>PT</td>
<td>0196</td>
</tr>
<tr>
<td>MOLE</td>
<td>FDA</td>
<td>AB</td>
<td></td>
</tr>
<tr>
<td>MOLE</td>
<td>FDA</td>
<td>PT</td>
<td></td>
</tr>
<tr>
<td>mole</td>
<td>HL7</td>
<td>PT</td>
<td></td>
</tr>
<tr>
<td>Mole</td>
<td>NCI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mole</td>
<td>UCUM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mole</td>
<td>CDISC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mol</td>
<td>HL7</td>
<td>PT</td>
<td></td>
</tr>
<tr>
<td>moi</td>
<td>ICH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mol</td>
<td>NCI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mol</td>
<td>NCI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mol</td>
<td>UCUM</td>
<td>AB</td>
<td></td>
</tr>
<tr>
<td>mol</td>
<td>CDISC</td>
<td>PT</td>
<td></td>
</tr>
</tbody>
</table>

6 sources
Choose your source

NCImetathesaurus

Mole, unit of measurement (CUI C0439189)

11 Sources

Source  Type  Code
CDISC  PT  MTHU119366
ICH  AB  0196
MGED  PT  MO_988
NCI  AB  C42539
NCI  AB  C42539
NCI-HL7  AB  12856
SNOMEDCT  PT  258717005
UCUM  AB  MTHU119368
CDISC  SY  MTHU120619
FDA  AB  MTHU105387
FDA  AB  MTHU105387
HL7V3.0  PT  MOL
HL7V3.0  PT  mol
NCI  PT  C42539
NCI-HL7  PT  12856
UCUM  PT  MTHU119367
SNOMEDCT  SY  258717005
SNOMEDCT  FN  258717005
ICH  FT  0196
MTH  FN  NOCODE
Progress on Quality and Interoperability

- **EVS offers quality terminology infrastructure**
  - Consistent access to controlled terminology
  - Mapping between meanings
  - Reference terminology for both people and computers
  - Extensive internal and external quality measures
  - Being deployed throughout NCI trial & informatics systems

- **Community participation builds quality, adoption & standards**
  - Partnerships crucial to high-quality consensus on meaning
  - Interoperability a core component of quality
  - Consensus and interoperability build durable standards
EVS Products & Services
Are Open

- NCI Thesaurus is Open Content
  http://evs.nci.nih.gov/terminologies
- NCI Metathesaurus is Mostly Open Source
  (See Each Source’s License)
- NCI EVS Servers Are Freely Accessible
  - On the Web:
    http://nciterms.nci.nih.gov
    http://ncimeta.nci.nih.gov
  - Via API:
    https://cabig.nci.nih.gov/tools/LexEVS_API
  - On caGrid:
    https://cabig.nci.nih.gov/workspaces/Architecture/caGrid
- All Software Developed by NCI EVS is Public Open Source and Free for the Asking:
  http://ncicb.nci.nih.gov/download/#ETools
Methods of Data Retrieval

- **NCI ftp site:**
  

- **NCI partner web sites (CDISC, FDA, etc.)**

- **Request a report from NCI staff:**
  
  [http://ncit.nci.nih.gov/ncitbrowser/pages/contact_us](http://ncit.nci.nih.gov/ncitbrowser/pages/contact_us)

- **NCIt Browser by subset:**
  

- **Cancer.gov:**
  
Consolidated Content Services
Lawrence W Wright
Associate Director
Enterprise Vocabulary Services
National Cancer Institute
lwright@mail.nih.gov

Margaret Haber
Associate Director
Enterprise Vocabulary Services
National Cancer Institute
mhaber@mail.nih.gov