



SNOMED CT:

The role of terminology in standardizing electronic pathology information

Kent A. Spackman, MD PhD

10th European Congress on Telepathology
4th International Congress on Virtual Microscopy
Vilnius, Lithuania
July 1, 2010



Overview of presentation

- Standardizing pathology information:
 - Organizational activities and resources
 - IHTSDO
 - Governance, Special Interest Groups, Project Groups
 - Interactions with other standards
 - Harmonization agreements, joint work
 - Implementation resources for representing clinical information
 - Basic SNOMED CT components
 - Codes, terms, subsets/refsets, mappings, hierarchies
 - Terminological composition
 - Concept model, logic, qualifiers and modifiers
 - Term binding
 - Data structures, information models,



Overview of presentation

- Standardizing pathology information:
 - **Organizational activities and resources**
 - IHTSDO
 - Governance, Special Interest Groups, Project Groups
 - Interactions with other standards
 - Harmonization agreements, joint work
 - Implementation resources for representing clinical information
 - Basic SNOMED CT components
 - Codes, terms, subsets/refsets, mappings, hierarchies
 - Terminological composition
 - Concept model, logic, qualifiers and modifiers,
 - Term binding
 - Data structures, information models,



Overview of presentation

- Standardizing pathology information:
 - Organizational activities and resources
 - IHTSDO
 - Governance, Special Interest Groups, Project Groups
 - Interactions with other standards
 - Harmonization agreements, joint work
 - Implementation resources for representing clinical information
 - Basic SNOMED CT components
 - Codes, terms, subsets/refsets, mappings, hierarchies
 - Terminological composition
 - Concept model, logic, qualifiers and modifiers,
 - Term binding
 - Data structures, information models,



About IHTSDO

- International Health Terminology Standards Development Organization
- Formed in Denmark, 23rd March 2007
- Area of standardization:
 - Terminology for interoperability of electronic health information





The purpose of the IHTSDO

- To acquire, own and administer the rights to SNOMED CT and other relevant assets (collectively, the "Terminology Products");
- To develop, maintain, promote and enable the uptake and correct use of its Terminology Products around the world;
- To undertake activities required to achieve these purposes





Status of the IHTSDO

- The IHTSDO is a Danish Association
- The Association is a registered not-for-profit entity in Denmark [23rd March 2007]
- Articles of Association detail the who, what, where and how of the Association
 - <http://www.ihtsdo.org/about-us/governance/>
- The Association owns the intellectual property
- Intellectual property in SNOMED CT and antecedent works (SNOMED 3.5, RT etc.) transferred to the IHTSDO [26th April 2007]





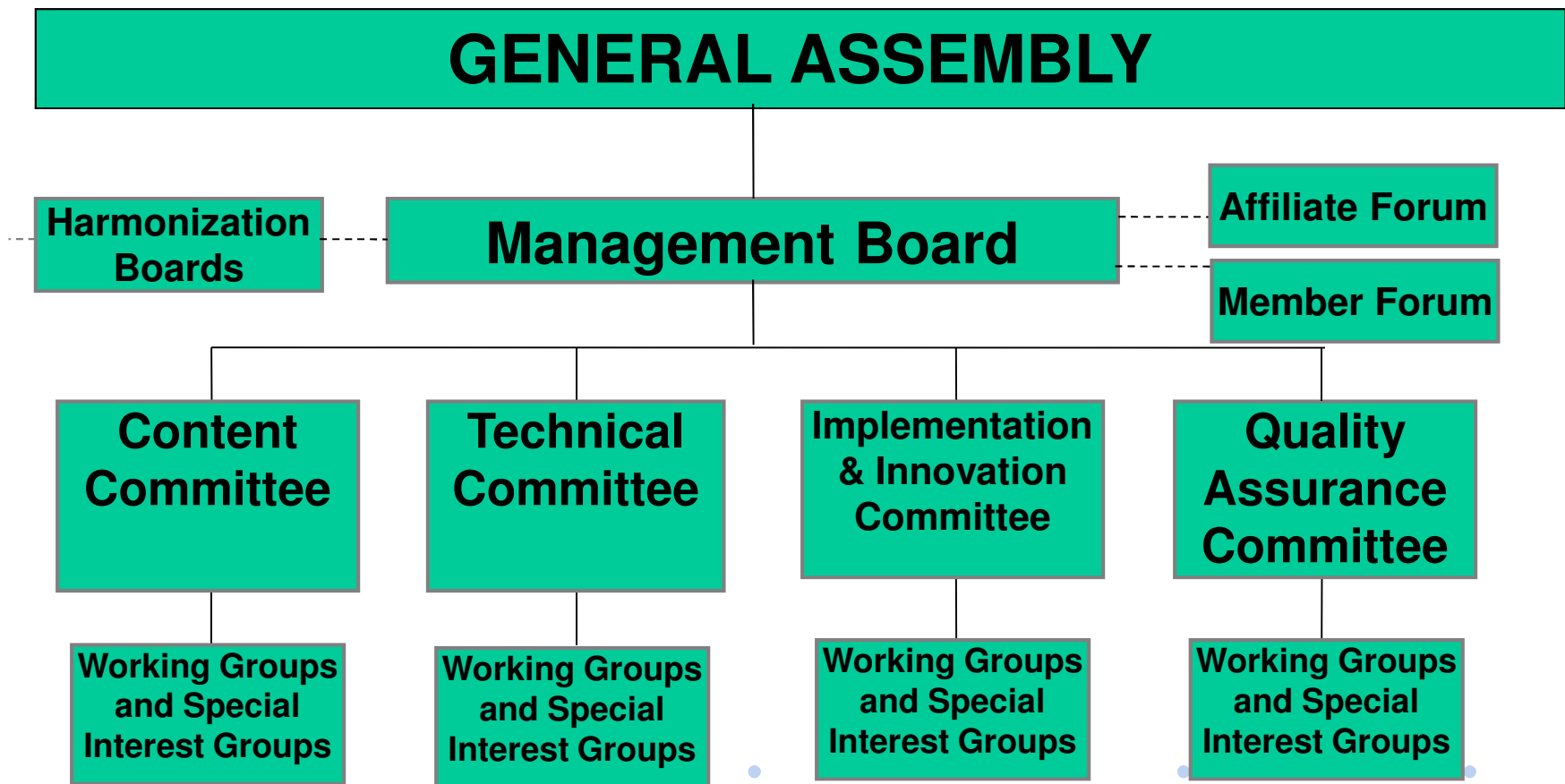
IHTSDO Members

- Members are countries
 - Eligible Members are all voting members of the United Nations
- The Members control the organization and the Articles of Association; [subject to Danish Law]
- Nine Charter [initial] Members in 2007:
 - Australia, Canada, Denmark, Lithuania, Netherlands, New Zealand, Sweden, United Kingdom, United States of America
- Six Members have subsequently joined:
 - Cyprus, Estonia, Singapore, Slovak Republic, Slovenia, Spain





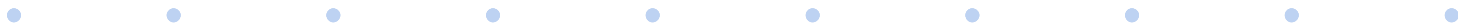
IHTSDO Organizational Structure





Working Groups and SIGs

- Two types of Group [both are entirely open]
 - Working Group: [Task focused]
 - Special Interest Group (SIG): [Domain focused e.g. profession (nursing), specialism (mapping)]
- First Global Profession SIG: International Pathology and Laboratory Medicine [in cooperation with the World Association of Societies of Pathology and Laboratory Medicine]
 - Chairman: Raj Dash, MD
 - IPaLM SIG reports to the Implementation & Innovation Committee





Overview of presentation

- Standardizing pathology information:
 - Organizational activities and resources
 - IHTSDO
 - Governance, Special Interest Groups, Project Groups
 - **Interactions with other standards**
 - **Harmonization agreements, joint work**
 - Implementation resources for representing clinical information
 - Basic SNOMED CT components
 - Codes, terms, subsets/refsets, mappings,
 - Terminological composition
 - Concept model, logic, qualifiers and modifiers,
 - Term binding
 - Data structures, information models,



Harmonization with other standards

- HL7
 - Terminfo work product: Initial guide on how to use SNOMED CT in HL7
- OpenEHR
 - Initial agreement to begin working together is complete
 - Closer working relationship is being developed
- WHO
 - Nearing agreement on cooperative work plan encompassing all WHO classifications (ICD-10, ICD-O, ICD-11, ICF, etc)
- LOINC and IFCC-IUPAC (NPU)
 - Nearing agreements on laboratory test terminology cooperation
- IHE
 - Work to date driven by individual IHTSDO Members:
 - E.g. Netherlands Nictiz, IHE/SNOMED in lab test reporting in 2007
- DICOM
 - Long-standing working relationship (next slide)



DICOM – SNOMED work

- SNOMED is the preferred coding system within DICOM for anatomy, clinical findings, procedures, pharmaceutical/biologic products (including contrast agents), and other clinical terms.
- Long history of cooperation
 - June 2000, SNOMED International (CAP) and NEMA entered into an agreement whereby SNOMED codes required for basic interoperability will be available on a pre-paid royalty basis. By agreement, terms continue to be added. SNOMED is an integral component to full content integration of digital images and text using DICOM's Structured Reporting standard.
- DICOM WG 26 (Pathology) – submits new terms periodically, for inclusion in SNOMED CT, and codes are used in DICOM standards





Overview of presentation

- Standardizing pathology information:
 - Organizational activities and resources
 - IHTSDO
 - Governance, Special Interest Groups, Project Groups
 - Interactions with other standards
 - Harmonization agreements, joint work
 - **Implementation resources for representing clinical information**
 - Basic SNOMED CT components
 - Codes, terms, subsets/refsets, mappings, hierarchies
 - Terminological composition
 - Concept model, logic, qualifiers and modifiers,
 - Term binding
 - Data structures, information models,



First Rule of *Coding*

Yesterday's data should be usable tomorrow

Corollary:

If no-one is going to re-use the data,
then no-one needs to code it.





Definitions: Clinical data re-use

- Primary use = “direct patient care”, or, the initial purpose
- Secondary use = anything beyond primary use

Example:

A pathologist sends a report of a biopsy to the surgeon who obtained the biopsy

What is the primary use?

What are the secondary uses of that report?





Data re-use

- Requires the ability to:
 - query databases containing coded clinical data
 - systematically retrieve patients based on general criteria
 - aggregate data in ways *not directly encoded*
 - Consider the following type of rule:
 - If the patient has had an MI but has no CHF, AV block, asthma, peripheral vascular disease, or Type 1 diabetes mellitus, and is not taking a beta blocker, you need to consider adding beta blocker therapy ...





Terminology standards

- How far does a reference terminology take us towards being able to use secondary data?
 - Permits common reference points for meaning
 - With appropriate history mechanism, sustains the value of previously recorded data
 - Does not (independently) solve the problem of data collection / data entry





Overview of presentation

- Standardizing pathology information:
 - Organizational activities and resources
 - IHTSDO
 - Governance, Special Interest Groups, Project Groups
 - Interactions with other standards
 - Harmonization agreements, joint work
 - Implementation resources for representing clinical information
 - **Basic SNOMED CT components**
 - Codes, terms, subsets/refsets, mappings
 - Terminological composition
 - Concept model, logic, qualifiers and modifiers
 - Term binding
 - Data structures, information models



Codes and Terms

- A code is any arbitrary string of characters
 - In SNOMED 3 (prior to SNOMED CT), the meaning hierarchy was represented in the code string:
 - D3-40000 subsumed D3-42000 subsumed D3-42100 ... etc
 - In SNOMED CT, the codes are integer strings
 - Hierarchy is represented in a separate table
- A term is the set of words that represent the meaning of the code
 - “adenocarcinoma”
 - “endoscopic retrograde choledochopancreatography”
 - “vermiform appendix”





SNOMED CT Encoding of Checklists

Checklist Identifier: [R-10142] College of American Pathologists Cancer Checklist; Lung: Biopsy (qualifier value)

***LUNG: Biopsy [P1-28300, 78603008] Biopsy of lung (procedure)**
(Note: Use of checklist for biopsy specimens is optional)

*Patient name: *[R-0025D, 371484003] Patient name (observable entity)*

*Surgical pathology number: *[R-002A2, 371482004] Surgical pathology identifier (observable entity)*

Note: Check 1 response unless otherwise indicated.

***MACROSCOPIC [F-048D6, 395526000] Macroscopic specimen observable (observable entity)**

*SPECIMEN TYPE *[R-00254, 371439000] Specimen type (observable entity)*

* ___ Fiberoptic bronchoscopic *[G-8328, 122611008] Specimen from lung obtained by fiberoptic bronchoscopic biopsy (specimen)*

* ___ Transbronchial *[G-8329, 122612001] Specimen from lung obtained by transbronchial biopsy (specimen)*

* ___ Mediastinoscopic biopsy *[G-832A, 122613006] Specimen from lung obtained by mediastinoscopic biopsy (specimen)*

* ___ CT-guided needle biopsy *[G-8348, 384747005] Specimen from lung obtained by CT guided needle biopsy (specimen)*

* ___ Wedge biopsy *[G-8347, 384746001] Specimen from lung obtained by wedge biopsy (specimen)*



Future of coded templates

INVASIVE CARCINOMA OF COLON

HISTOLOGIC TYPE	Adenocarcinoma	SNOMED CT: 371441004+35917007
TUMOR SIZE	3 X 2 cm	
TUMOR CONFIGURATION	Ulcerative	
HISTOLOGIC GRADE	High grade	Histologic type (observable entity)
EXTENT OF INVASION	Into (not through)	PLUS
NODE STATUS	Metastatic carcinoma	Adenocarcinoma, no subtype (morphologic abnormality)
	out of twenty-one regional lymph nodes (2/21)	
MARGIN STATUS		
PROXIMAL	Negative	
DISTAL	Negative	
RADIAL (DEEP)	Negative	
CLOSEST MARGIN	Radial (3 cm)	
VASCULAR (SMALL VESSEL) INVASION	Absent	
VASCULAR (LARGE VENOUS) INVASION	Present (Intramural)	
PERINEURAL INVASION	Present	
MESORECTAL INVASION	Not applicable	
DISTANT METASTASIS	Not identified	
TUMOR BORDER CONFIGURATION	Infiltrating (ragged)	
SITE	Transverse colon	
SPECIMEN TYPE	Transverse hemicolectomy (12 cm)	



SNOMED and ICD-O

M-9400/3 in ICD-O:

M-9400/3 Astrocytoma, NOS

Astrocytic glioma

Astrogloma

Diffuse astrocytoma

Astrocytoma, low grade

Diffuse astrocytoma, low grade

Cystic astrocytoma





SNOMED and ICD-O

- ICD-O captured both Astrocytoma, NOS and Diffuse astrocytoma with M-9400/3
- SNOMED created a separate concept for Diffuse astrocytoma, and placed it in the SNOMED hierarchy under M-94003

MICROSCOPIC [F-048D7, 395527009] Microscopic specimen observable (observable entity)

HISTOLOGIC TYPE [R-00257, 371441004] Histologic type (observable entity)

___ Astrocytoma, not otherwise characterized [M-94003, 38713004]

Astrocytoma, no ICD-O subtype (morphologic abnormality)

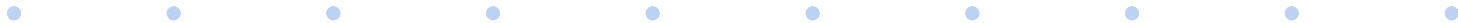
___ Astrocytoma, diffuse [R-1009A, 397381007] Diffuse astrocytoma (morphologic abnormality)





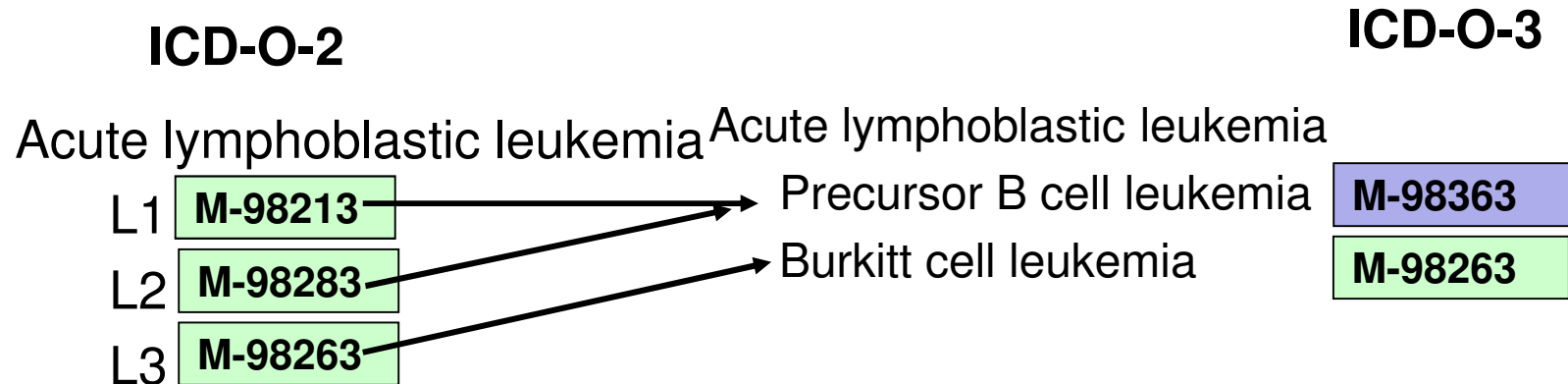
Secondary Use: Cancer Registry

- ICD-O version 2 was replaced in 2001 with ICD-O version 3.
- Example: Acute lymphoblastic leukemia
 - French-American-British classification had three subtypes: L1, L2, and L3
 - ICD-O-2 codes were:
 - L1: 9821/3
 - L2: 9828/3
 - L3: 9826/3





Acute lymphoblastic leukemia



SNOMED history table contains two rows:
M-98213 REPLACED-BY M-98363
M-98283 REPLACED-BY M-98363



Example: Burkitt's lymphoma/leukemia

Acute lymphoblastic leukemia

B-lineage ALL

Early pre-B cell ALL

Pre-B cell ALL

Transitional B-lineage ALL

Mature B cell ALL

T-lineage ALL

B-cell neoplasm

Mature (peripheral) B-cell neoplasm

Burkitt lymphoma/leukemia

Endemic

Sporadic

Atypical

Immunodeficiency assoc.



Example: Burkitt's lymphoma/leukemia

Acute lymphoblastic leukemia

B-lineage ALL

Early pre-B cell ALL

Pre-B cell ALL

Transitional B-lineage ALL

9826/3 Burkitt cell leukemia

T-lineage ALL

B-cell neoplasm

Mature (peripheral) B-cell neoplasm

9687/3 Burkitt lymphoma, NOS

Endemic


Sporadic

Atypical

Immunodeficiency assoc.

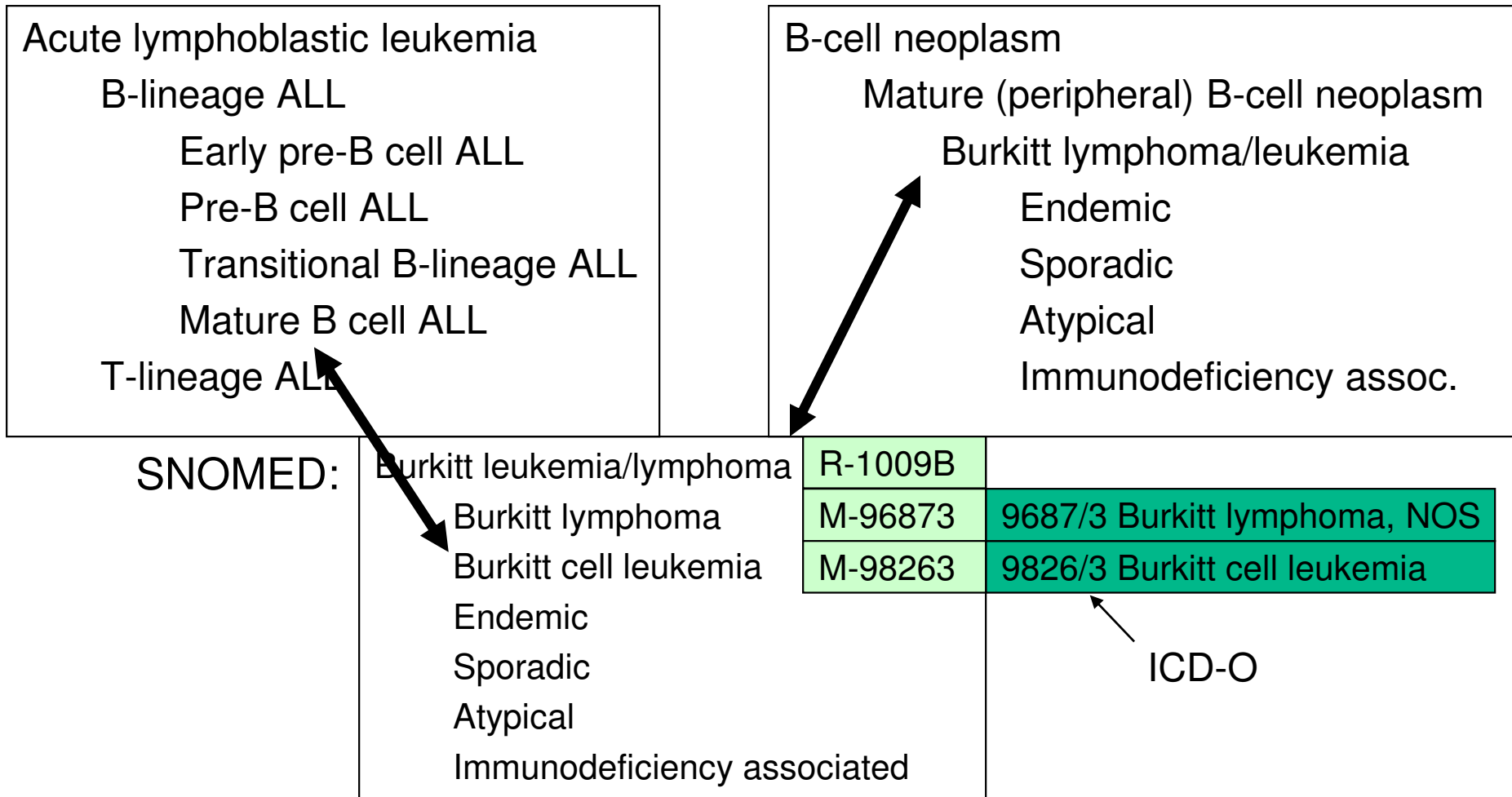


Example: Burkitt's lymphoma/leukemia

<p>Acute lymphoblastic leukemia B-lineage ALL Early pre-B cell ALL Pre-B cell ALL Transitional B-lineage ALL Mature B cell ALL T-lineage ALL</p>	<p>B-cell neoplasm Mature (peripheral) B-cell neoplasm Burkitt lymphoma/leukemia Endemic Sporadic Atypical Immunodeficiency assoc.</p>							
<p>SNOMED:</p>	<p>Burkitt leukemia/lymphoma Burkitt lymphoma Burkitt cell leukemia Endemic Sporadic Atypical Immunodeficiency associated</p>	<table border="1"> <tr> <td data-bbox="1066 1037 1287 1097">R-1009B</td> <td data-bbox="1287 1037 1929 1097"></td> </tr> <tr> <td data-bbox="1066 1097 1287 1157">M-96873</td> <td data-bbox="1287 1097 1929 1157">9687/3 Burkitt lymphoma, NOS</td> </tr> <tr> <td data-bbox="1066 1157 1287 1218">M-98263</td> <td data-bbox="1287 1157 1929 1218">9826/3 Burkitt cell leukemia</td> </tr> </table> <p style="text-align: right; margin-right: 50px;">  ICD-O </p>	R-1009B		M-96873	9687/3 Burkitt lymphoma, NOS	M-98263	9826/3 Burkitt cell leukemia
R-1009B								
M-96873	9687/3 Burkitt lymphoma, NOS							
M-98263	9826/3 Burkitt cell leukemia							



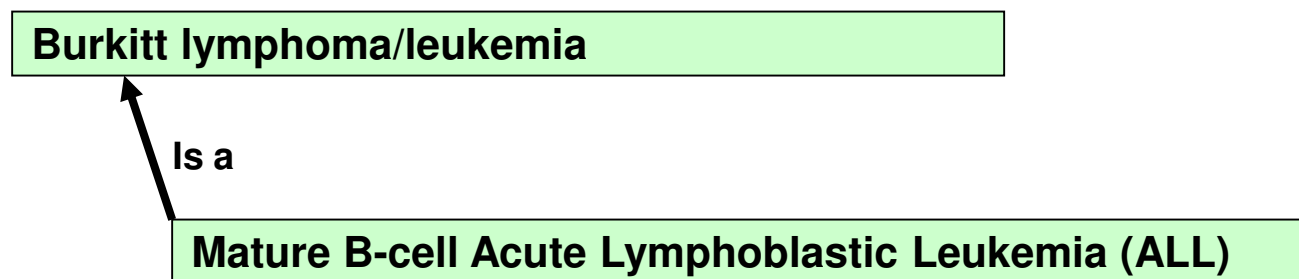
Example: Burkitt's lymphoma/leukemia





Data re-use and the value of a reference terminology

- A well-defined reference terminology permits the maximum re-use of data collected according to different views
 - This is one of the main purposes for a reference terminology such as SNOMED CT





Overview of presentation

- Standardizing pathology information:
 - Organizational activities and resources
 - IHTSDO
 - Governance, Special Interest Groups, Project Groups
 - Interactions with other standards
 - Harmonization agreements, joint work
 - Implementation resources for representing clinical information
 - Basic SNOMED CT components
 - Codes, terms, subsets/refsets, mappings
 - **Terminological composition**
 - **Concept model, logic, qualifiers and modifiers**
 - Term binding
 - Data structures, information models



Composition

- Composition is putting two or more codes together to represent your meaning
- The codes combinations can be
 - within the terminology itself (post-coordination)
 - can be done at run time or pre-compiled at user site
 - a terminology code within a standard data structure
 - HL7, OpenEHR, and other standards
 - proprietary structures also sometimes used





SNOMED Machine-Readable Concept Model (MRCM)

- Formal machine-readable constraints for SNOMED compositions
 - XML format
 - In process: integration with the IHTSDO Workbench
- MRCM in principle allows:
 - Uniform application of the concept model constraints
 - Validation of expressions
 - Support for creation of well-formed post-coordination expressions
 - Support for well-formed queries against data
- See MRCM Project Group
 - IHTSDO Collaborative Site (thecap.basecamphq.com)
 - Send email to info@ihtsdo.org to sign up





Description Logic

- SNOMED CT has a description logic foundation
 - It is distributed in OWL format
 - The OWL 2 profile is OWL EL
- There are many different classifiers available
 - Apelon Ontylog (currently used)
 - Snorocket (open source – being using in new Workbench)
 - FaCT++
 - others



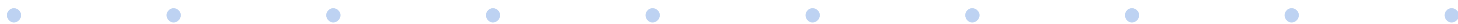
Overview of presentation

- Standardizing pathology information:
 - Organizational activities and resources
 - IHTSDO
 - Governance, Special Interest Groups, Project Groups
 - Interactions with other standards
 - Harmonization agreements, joint work
 - Implementation resources for representing clinical information
 - Basic SNOMED CT components
 - Codes, terms, subsets/refsets, mappings
 - Terminological composition
 - Concept model, logic, qualifiers and modifiers
 - **Term binding**
 - **Data structures, information models**



How does terminology interact with the record structure?

- Terminology model
 - terminology codes and expressions
- Information model
 - elements of the record structure, and inter-relationships of those elements
- Codes & expressions populate the record structure
- Interoperability depends on all these





Balance, overlaps, gaps

- Consider how to record the fact that the patient's blood type is "RH positive".
- What is the information model (field in the record)?
- What is the terminology (value put in the field)?

Field or question	Terminology value
Blood type	RH positive
RH D antigen	positive
Lab test result	Blood type RH positive
RH positive	Yes





Blood types: questions and answers

- Question? = Answer
- Blood type? = Rh positive.
- Rh positive? = Yes
- Rh type? = D positive
- D positive? = Yes
- Rh type? = Rh positive
- Rh type? = positive
- Allowable values
- Rh type
 - Rh(o) D positive
 - Rh(o) D negative
- What about:
 - Weak D
 - Partial D



Balance, overlaps, gaps

- Record the fact that “malignant mesothelial cells were found in a pleural fluid aspirate”:

Field or question	Terminology value
Pleural fluid finding	Malignant mesothelial cells
Site of malignant mesothelial cells	Pleural fluid
Lab test result	Malignant mesothelial cells in pleural fluid
Type of mesothelial cells in pleural fluid	Malignant
Type of malignant cells in pleural fluid	Mesothelial





Balance, overlaps, gaps

- There is no single best way to split assertions between the information model and the terminology model (or between the observables and the other values in the terminology, or between data slots in an archetype and values in the terminology)
- The best we can do is recognize equivalence
- The best tools for recognizing equivalence (by machine) are logic-based
- Therefore, a logic-based model of semantics is the foundation not just for the terminology but also for the *combination*





Some resources for term binding

- HL7 Terminfo
 - Using SNOMED CT in HL7 version 3; Implementation Guide, Release 1.5. <http://www.hl7.org/v3ballot/html/welcome/environment/index.htm>
 - Navigate to: Foundation->Using SNOMED CT
- Clinical Document Architecture (CDA)
 - *JAMIA, Oct 2005 (doi: 10.1197/jamia.M1888)*
- UK NHS Logical Record Architecture (LRA)
 - Register with TRUD
- EN13606 and openEHR
 - openEHR Architecture Overview
 - <http://www.openehr.org/releases/1.0.1/architecture/overview.pdf>
- Clinical Element Models (CEM)
 - Intermountain Health Care & General Electric
 - Presentation by Stan Huff to IHTSDO CM-SIG, winter 2010



Conclusion

- IHTSDO is the organization that maintains SNOMED CT
- Member countries run the IHTSDO
- SNOMED CT has resources for implementing electronic pathology information
- Collaboration and cooperation among standards groups is essential, and IHTSDO is actively working on multiple harmonization fronts
- You are welcome to get involved!
 - www.ihtsdo.org
 - thecap.basecamphq.com (email info@ihtsdo.org)

