

Terminology, Classification and Clinical Coding

Centre for Health Informatics

Prof John Chelsom

Runs: 3 hours
Tutor: Prof John Chelsom
Mode of attendance: Classroom

With thanks for some material
developed by Omid L. Shabestari M.D

Learning Objectives

- This session provides an overview of terminology, classification and clinical coding, with a more detailed look at the major coding and classification systems in use today.
- Specific learning objectives are to:
 - 1 Define terminology, classification and clinical coding schemes
 - 2 Understand the purpose and requirements for clinical coding
 - 3 Identify the major clinical coding schemes in use today
 - 4 Assess the structure, utility, strengths and weaknesses of these schemes
 - 5 Explore general concepts of coding for specific clinical requirements

Terminology, Classification and Clinical Coding

- Definition, purpose and requirements
- Coding Schemes
- Closer look at coding schemes
- Potential problems with clinical coding
- References and Further Reading

Information Sources

- Sources are listed in the references at the end of these slides



Some definitions and descriptions have been taken from quoted resources.

Retrieved October 2010.

Where consensus on definitions or descriptions is required, these have been taken from Wikipedia.

Retrieved October 2010.

Terminology, Classification and Clinical Coding

Definition, purpose and requirements

Clinical Information

- Clinical information can be accessed (electronically) from many different sources
 - Electronic health and patient records
 - Holding both structured and unstructured information
 - Guidelines and protocols
 - Medical textbooks
 - Web pages

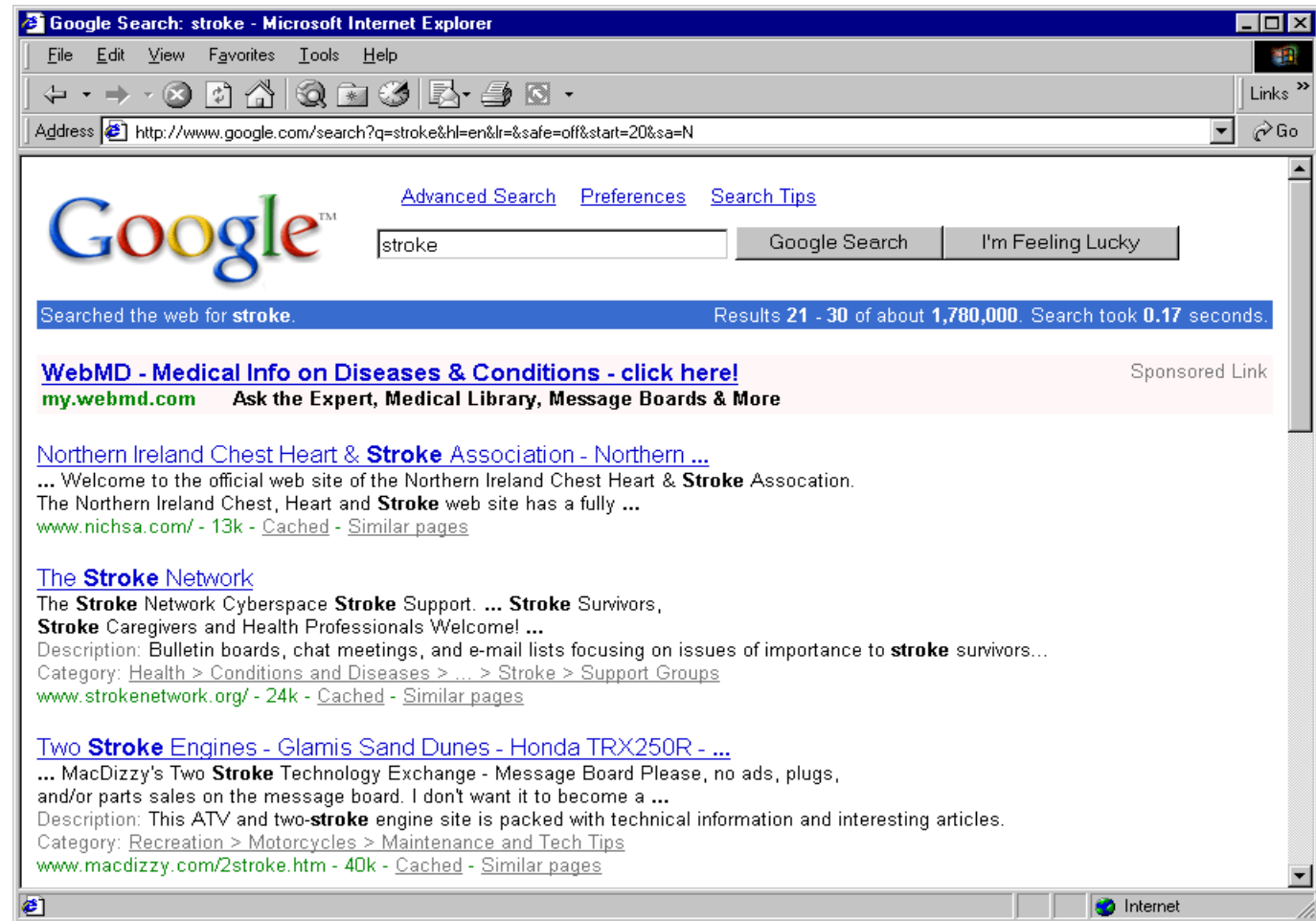
Free Text Search

Some problems with free text search of clinical information

- Different terms (or synonyms) point to the same concept. e.g. *myocardial infarction*, *coronary thrombosis*, *heart attack*, etc.
- What about other languages and language variants e.g. anemia vs. anaemia, appendectomy vs. appendicectomy
- What about spelling mistakes? e.g. anaemea, apendectomy
- Little or no regard to context: is the term relevant to the search we are performing? (e.g. try searching for Asthma in the EPR)

Problems with Free Text Search

You might be doing a search on the term '*stroke*' (cerebral infarction) and end up with documents that teach you about the workings of the two-stroke motorcycle engine. The non-discriminatory free-text method of document retrieval inevitably produces a number of irrelevant leads or noise (Kiley, 1996).



Constraining Free Text in Clinical Information Sources

- Free text searching doesn't work, what is the alternative?
- A simple coding scheme should do –
 - *"Medical language should be particularly easy for a computer to understand: there are relatively few terms, and these are well-understood with clear meanings"*
- Clinical information/knowledge is intrinsically difficult to pin down
 - Large variation in ways to record similar (identical?) concepts
 - Personal style
 - Importance of purpose and context

Vocabularies as Meta Data

A **vocabulary** is a set of words known to a person or other entity, or that are part of a specific language.

Technical **terminology** is the specialised vocabulary of a profession

A **lexicon** is usually a list of words together with additional word-specific information, i.e., a dictionary. Lexicon is a word of Greek origin (λεξικόν) meaning vocabulary.

A **dictionary** is a list of words with their definitions (type of lexicon)

A **controlled vocabulary** is a carefully selected list of words and phrases, which are used to tag units of information so that they may be more easily retrieved by a search

Fully developed controlled vocabulary systems, such as the Library of Congress Subject Headings, are often published in a reference work that is called a **thesaurus** that maps synonyms to concepts.

Controlled vocabularies form part of a larger universe of nomenclatural approaches to data classification called **meta data**

Ontology

A vocabulary of terms, together with the relationships between the terms can be called an Ontology.

A body of formally represented knowledge is based on a conceptualization: the objects, concepts, and other entities that are assumed to exist in some area of interest and the relationships that hold among them.

A conceptualization is an abstract, simplified view of the world that we wish to represent for some purpose. Every knowledge base, knowledge-based system, or knowledge-level agent is committed to some conceptualization, explicitly or implicitly.

An ontology is an explicit specification of a conceptualization.

Ontologies are often equated with taxonomic hierarchies of classes, class definitions, and the 'is a' relation, but ontologies need not be limited to these forms.

Tom Gruber of Stanford University, Knowledge Systems Lab

Payment By Results

The Department of Health (DH) published 'Reforming NHS Financial Flows' in October 2002. This outlined plans to introduce Payment by Results - a new funding system for the care provided to NHS patients in England.

The aim of Payment by Results (PbR) is to provide a transparent, rules-based system for paying trusts. It will reward efficiency, support patient choice and diversity and encourage activity for sustainable waiting time reductions.

<http://www.ic.nhs.uk/services/the-casemix-service/new-to-this-service/background-to-payment-by-results-pbr-and-healthcare-resource-group-hrg-development>

- PbR is activity-based payment, that uses Healthcare Resource Groups 4 (HRG4) as a basis for categorising clinical services.

Healthcare Resource Group

Within the English National Health Service (NHS), a Healthcare Resource Group (HRG) is a grouping consisting of patient events that have been judged to consume a similar level of resource.

For example, there are a number of different knee-related procedures that all require similar levels of resource; they may all be assigned to one HRG.

http://en.wikipedia.org/wiki/Healthcare_Resource_Group

- Published and maintained by
 - The National Casemix Service, part of the NHS Health and Social Care Information Centre
 - <http://www.hscic.gov.uk/casemix>

HRG and Clinical Coding

Healthcare Resource Groups 4 (HRG4) will support PbR by providing a classification framework that represents current clinical practice.

In addition, HRG4 will support service planning, costing and commissioning between PCTs and trusts by providing reliable and consistent presentation of activity data to:

- support the focus on patient-centred care, to enable patient choice
- support the analysis of healthcare needs and monitoring of service provision, to inform service planning.

<http://www.ic.nhs.uk/services/the-casemix-service/new-to-this-service/background-to-payment-by-results-pbr-and-healthcare-resource-group-hrg-development>

HRG4 uses a combination of ICD codes for disease and OPCS-4 codes for surgical / clinical procedures

Diagnosis-related Group

Diagnosis-related group (DRG) is a system to classify hospital cases into one of approximately 500 groups, also referred to as DRGs, expected to have similar hospital resource use, developed for Medicare as part of the prospective payment system.

DRGs are assigned by a "grouper" program based on ICD diagnoses, procedures, age, sex, discharge status, and the presence of complications or comorbidities.

DRGs have been used in the US since 1983 to determine how much Medicare pays the hospital, since patients within each category are similar clinically and are expected to use the same level of hospital resources. DRGs may be further grouped into Major Diagnostic Categories (MDCs).

http://en.wikipedia.org/wiki/Diagnosis-related_group

- Published and maintained by
 - Centers for Medicare and Medicaid Services (CMS)
 - <http://cms.hhs.gov>

Features of a Clinical Terminology

- Some desirable features of a clinical terminology
 - Concept based with a semantic definition of concepts
 - Completeness
 - All synonyms of a concept point to it and are semantically associated with it
 - Hierarchical
 - Multi-axial (Multiple classification and multiple parentage)
 - Compositional (see later)
 - Mapped to classifications
 - » usually some detail is lost as classifications have coarser granularity compared to terminologies
 - Language-independent model
 - Machine processable / executable

Coding Schemes

Clinical Coding Schemes

- There are hundreds of clinical coding schemes in existence
 - See http://en.wikipedia.org/wiki/Clinical_coding for a partial list
- Classified by specialist domain
 - All medicine SNOMED CT, UMLS
 - Nursing
 - Disease
 - Laboratory medicine
- Classified by usage
 - Record Keeping (i.e. EHR/EPR)
 - Billing
 - Epidemiology and clinical trials/studies

SNOMED CT

SNOMED CT (Systematized Nomenclature of Medicine - Clinical Terms), is a systematically organized computer processable collection of medical terminology covering most areas of clinical information such as diseases, findings, procedures, micro-organisms, pharmaceuticals etc.

It allows a consistent way to index, store, retrieve, and aggregate clinical data across specialties and sites of care.

It also helps organizing the content of medical records, reducing the variability in the way data is captured, encoded and used for clinical care of patients and research.

http://en.wikipedia.org/wiki/SNOMED_CT

- Published and maintained by
 - The International Health Terminology Standards Development Organization
 - IHTSDO
 - www.ihtsdo.org



ICD-10

The International Statistical Classification of Diseases and Related Health Problems (most commonly known by the abbreviation ICD) provides codes to classify diseases and a wide variety of signs, symptoms, abnormal findings, complaints, social circumstances, and external causes of injury or disease.

Under this system, every health condition can be assigned to a unique category and given a code, up to six characters long. Such categories can include a set of similar diseases.

http://en.wikipedia.org/wiki/International_Statistical_Classification_of_Diseases_and_Related_Health_Problems

- Published and maintained by
 - The World Health Organisation
 - WHO
 - <http://www.who.int/classifications/icd>



MeSH

Medical Subject Headings (MeSH) is a comprehensive controlled vocabulary for the purpose of indexing journal articles and books in the life sciences; it can also serve as a thesaurus that facilitates searching.

Created and updated by the United States National Library of Medicine (NLM), it is used by the MEDLINE/PubMed article database and by NLM's catalogue of book holdings.

http://en.wikipedia.org/wiki/Medical_Subject_Headings

- Published and maintained by
 - US National Library of Medicine
 - NLM
 - <http://www.nlm.nih.gov/mesh/>

MEDICAL SUBJECT HEADINGS



LOINC

Logical Observation Identifiers Names and Codes (LOINC) is a database and universal standard for identifying medical laboratory observations. It was developed and is maintained by the Regenstrief Institute, Inc., a US non-profit medical research organization, in 1994.

LOINC is one of the standards for use in U.S. Federal Government systems for the electronic exchange of clinical health information.

In 1999, it was identified by the HL7 Standards Development Organization as a preferred code set for laboratory test names in transactions between health care facilities, laboratories, laboratory testing devices, and public health authorities.

<http://en.wikipedia.org/wiki/LOINC>

- Published and maintained by
 - Regenstrief Institute, Inc
 - <https://loinc.org/>



UMLS

The Unified Medical Language System (UMLS) is a compendium of many controlled vocabularies in the biomedical sciences (created 1986).

It provides a mapping structure among these vocabularies and thus allows one to translate among the various terminology systems; it may also be viewed as a comprehensive thesaurus and ontology of biomedical concepts.

UMLS further provides facilities for natural language processing. It is intended to be used mainly by developers of systems in medical informatics.

<http://en.wikipedia.org/wiki/UMLS>

- Published and maintained by
 - US National Library of Medicine
 - NLM
 - <http://www.nlm.nih.gov/research/umls/>

OPCS-4

OPCS-4 is an abbreviation of the Office of Population, Censuses and Surveys Classification of Surgical Operations and Procedures (4th revision).

It translates operations, procedures and interventions carried out on a patient during an episode of health care in the NHS into alphanumeric code usually done by trained health care professionals working in an area called clinical coding.

<http://en.wikipedia.org/wiki/OPCS-4.5>

- Published and maintained by
 - NHS Classifications Service
 - part of NHS Connecting for Health
 - NCS
 - <http://www.connectingforhealth.nhs.uk/systemsand services/data/clinicalcoding/codingstandards/opcs4>

Terminology Reference Data Update Distribution Service



NHS Connecting for Health TRUD Service - Welcome

Release Sub Packs ?

Click on the categories below to expand/retract pack listings. Clicking on the pack will display its associated licence(s).

- All
- Clinical Imaging Procedures
- Health Language
- IHTSDO
- LRA
- MIM
- NHS Classifications
- NHS Data Model and Dictionary
- NHS Interoperability Toolkit
- NHS dm+d
- NHS eLearning
- ODS and SDS
- Terminology Tools
- UK READ
- UKTC SNOMED

Terminology Reference Data Update Distribution Service

This service is hosted by the UK Terminology Centre, part of the Data Standards and Products Directorate of NHS Connecting for Health.

Registration

You must register to use this site, using the 'click here to register' link on the right.

After registration you'll be emailed login details. Use these to access the site in future in order to maintain your registration and licence details, using the other links on the right.

Note: Details of your registration may be shared with relevant central representatives from your home country for the business of Terminology within their geographical healthcare boundaries

Trud Resources

Information and licensing details on RELEASE_SUB_PACKS (a compressed archive of all necessary files for a given reference dataset) can be obtained using the links on the left. Details of how to download RELEASE_SUB_PACKS will be sent to you after registration.

About Trud

NHS Connecting for Health is an agency of the Department of Health and is delivering the National Programme for Information Technology in the NHS (NPfIT).

Additionally NHS Connecting for Health is the host of the International Health Terminology Standards Development Organisation UK Terminology Centre (UKTC).

The Terminology Reference-data Update Distribution Service (TRUD) provides a mechanism for the UK Terminology Centre to license and distribute reference-data to interested parties.

This is the preferred Terminology distribution method of the UK Terminology Centre and NHS Connecting for Health.

Registered users of this service will be granted access to download reference data for which they are licensed.

Login:

[Problems logging in?](#)

Username:

Password:

[Log In](#)

[Click here to register](#)

[Registration Instructions]

[Forgotten/Change Password](#)

[External Admin Login](#)

Useful Links

[NHS N3 FTP Server](#)

[WWW FTP Server](#)

[Requests Submission Portal](#)

TRUD News:

Week 39 Release of the NHS BSA dm+d files and dm+d supplementary files are now available

Week 39 Release of the NH...

2010-09-27 09:23:46

The 4-weekly Release of SNOMED CT UK Drug Extension for 22nd September 2010 is now available

The 4-weekly Release of S...

2010-09-21 14:37:14

The SNOMED CT Core International July 2009 Release is now available

The SNOMED CT Core Intern...

<https://www.uktcregistration.nss.cfh.nhs.uk/trud/>

Closer Look at Coding Schemes

LOINC
ICD
SNOMED-CT
UMLS

Classifying Clinical Coding Schemes

By technical properties:

- Enumerated vs compositional schemes
- Lexical schemes
- Classifications

Enumerated Approach

- Pre-coordinated systems
 - Every possible concept is listed explicitly
 - However, result in compositional explosion
 - can never be sure that you have listed all the possibilities
 - “too big and too small”
 - need to list a large number of concepts so it’s hard to find what you want
 - Hard to predict every possible clinical subtlety, so often the exact concept required isn’t there

Compositional Approach

- Compositional systems seek to construct concepts from primitive building blocks, governed by validation rules (that prevent nonsense combinations)
- Many are based on compositional OAV (Object-Attribute-Value) triples e.g.
femur whichHasFractureType SimpleFracture

Lexical Approach

- Aim to use the rules of natural language e.g. sentence structure, syntactical rules, etc. to extract concepts and classify them.
- Requires
 - language-specific interpreter
 - Language independent concept model

Classification

- A list of concepts or codes grouped according to their similarities
- May be
 - Pre-classified i.e. the classification author decides on the most useful / interesting way to organise the concepts and hardwires it in
 - Classified “on the fly”

Closer Look at Coding Schemes

LOINC
ICD
SNOMED-CT
UMLS

LOINC – Identifiers

LOINC applies universal code names and identifiers to medical terminology

- The database currently has over 58,000 observation terms that can be accessed and understood universally.
- A unique code (format: nnnnn-n) is assigned to each entry upon registration.

LOINC – Names

Each database record includes six fields for the unique specification of each identified single test, observation, or measurement

- 1 Component - what is measured, evaluated, or observed (example: urea,...)
- 2 Kind of property - characteristics of what is measured, such as length, mass, volume, time stamp and so on
- 3 Time aspect - interval of time over which the observation or measurement was made
- 4 System - context or specimen type within which the observation was made (example: blood, urine,...)
- 5 Type of scale- the scale of measure. The scale may be quantitative, ordinal, nominal or narrative
- 6 Type of method- procedure used to make the measurement or observation

Closer Look at Coding Schemes

LOINC
ICD
SNOMED-CT
UMLS

International Classification of Diseases





International Statistical Classification of Diseases and Related Health Problems

- 1st edition in 1900, updates every 10 years (ish)
- The 10th edition, ICD-10 was developed in 1992
- ICD-11 is planned for 2015
 - to be revised using Web 2.0 principles.
- Core use is for WHO mortality statistics
- Expanded to include other medical contexts
- Divided into chapters based on disease classification

ICD Adaptations

Additional coding, focused on specific clinical/research requirements

ICD ADAPTATIONS

- International Classification of Diseases for Oncology, 3rd Edition (ICD-O-3)
- International Classification of External Causes of Injury (ICECI)
- International Classification of Primary Care, Second edition (ICPC-2)
-  The ICD-10 for Mental and Behavioural Disorders Diagnostic Criteria for Research
 pdf, 732kb
-  The ICD-10 for Mental and Behavioural Disorders Clinical Descriptions and
 Diagnostic Guidelines
pdf, 1.35Mb

International Classification of Diseases

ICD-10 has 22 chapters for classification

Chapter	Blocks	Title
I	A00-B99	Certain infectious and parasitic diseases
II	C00-D48	Neoplasms
III	D50-D89	Diseases of the blood and blood-forming organs...
IV	E00-E90	Endocrine, nutritional and metabolic diseases
V	F00-F99	Mental and behavioural disorders
VI	G00-G99	Diseases of the nervous system
VII	H00-H59	Diseases of the eye and adnexa
VIII	H60-H95	Diseases of the ear and mastoid process
IX	I00-I99	Diseases of the circulatory system
X	J00-J99	Diseases of the respiratory system
XI	K00-K93	Diseases of the digestive system
XII	L00-L99	Diseases of the skin and subcutaneous tissue
XIII	M00-M99	Diseases of the musculoskeletal system and connective tissue
XIV	N00-N99	Diseases of the genitourinary system
XV	O00-O99	Pregnancy, childbirth and the puerperium
XVI	P00-P96	Certain conditions originating in the perinatal period
XVII	Q00-Q99	Congenital malformations, deformations and chromosomal abnormalities
XVIII	R00-R99	Symptoms, signs and abnormal clinical and laboratory findings...
XIX	S00-T98	Injury, poisoning and certain other consequences of external causes
XX	V01-Y98	External causes of morbidity and mortality
XXI	Z00-Z99	Factors influencing health status and contact with health services
XXII	U00-U99	Codes for special purposes

Classification Markup Language

ICD-10 is distributed in XML using ClaML

ClaML is an acronym for Classification Markup Language. This is an xml data format specification meant for the exchange of medical classifications, which are code numbers for of medical diagnoses and procedures.

The ClaML specification has first been published as Technical Specification CEN/TS 14463:2003. The 2007 revision of ClaML has been accepted as European Norm EN 14463:2007.

The ClaML standard has been prepared by Working Group 3 (Semantic Content) of the International Organization for Standardization's (ISO) Technical Committee (TC) on health informatics, known as ISO/TC 215 WG3.

ClaML has been adopted by the WHO to distribute their family of international classifications..

<http://en.wikipedia.org/wiki/ClaML>

Check Out SNOB

<http://snob.eggbird.eu/manual/clamlexample.html>

eggbird
software development and
knowledge engineering



- eggbird
- Introduction
- Download
- statistics
- Manual
- . Installing
- . Importing
- . Exporting
- . Scripting
- . Indexing
- . Browsing
- . Searching
- . Colours
- . Subsets
- . Crossmappings
- . MetaModel
- . ClaML
 - . import
 - . example
 - . Notes
 - . Comment
 - . Post-coordinate
 - . Changelog
 - . Editing
 - . Lexicon
 - . Decorations
- Settings
- SNOB Server
- Screenshots
- Users
- FAQ
- TCM
- External links

[SNOB Wiki](#)

I support SNOB with

a small amount €10,00 EUR

Support

SNOB - A SNOMED Browser

ClaML Example

```
<?xml version="1.0"?>
<!DOCTYPE ClaML SYSTEM "http://terminologies.eu/download/en14463-2007.dtd">
<ClaML version="2.0.0">
  <Title name="example" date="20070926" version="1">Example ClaML</Title>
  <ClassKinds>
    <Kind name="category"/>
    <Kind name="chapter"/>
  </ClassKinds>
  <RubricKinds>
    <Kind name="preferred"/>
  </RubricKinds>
  <Class code="A" kind="chapter">
    <SubClass code="A.1"/>
    <SubClass code="A.2"/>
    <SubClass code="A.3"/>
    <Rubric kind="preferred">
      <Label xml:lang="en">chapter A</Label>
    </Rubric>
  </Class>
  <Class code="B" kind="chapter">
    <Rubric kind="preferred">
      <Label xml:lang="en">chapter B</Label>
    </Rubric>
  </Class>
  <Class code="C" kind="chapter">
    <Rubric kind="preferred">
      <Label xml:lang="en">chapter C</Label>
    </Rubric>
  </Class>
  <Class code="A.1" kind="category">
    <SuperClass code="A"/>
    <Rubric kind="preferred">
      <Label xml:lang="en">example class</Label>
    </Rubric>
  </Class>
  <Class code="A.2" kind="category">
    <SuperClass code="A"/>
    <Rubric kind="preferred">
      <Label xml:lang="en">example class</Label>
    </Rubric>
  </Class>
  <Class code="A.3" kind="category">
    <SuperClass code="A"/>
    <Rubric kind="preferred">
      <Label xml:lang="en">example class</Label>
    </Rubric>
  </Class>
</ClaML>
```

The classification in this ClaML file would look something like this in a browser.

A chapter A
 A.1 example class
 A.2 example class
 A.3 example class
B chapter B
C chapter C

iPhot

Closer Look at Coding Schemes

LOINC
ICD
SNOMED-CT
UMLS

Overview of SNOMED-CT

SNOMED CT consists of over a million medical Concepts.

For example 22298006 means myocardial infarction (MI).

Basic structure with

- Concepts table
- Descriptions table
- Relationships table

Check out SNOMED CT Browsers at

- http://www.nlm.nih.gov/research/umls/Snomed/snomed_browsers.html

Try out <http://www.snomedbrowser.com/>

Structure of SNOMED-CT

The Concepts are arranged in a type or IS-A hierarchy.

- For example, Viral pneumonia IS-A Infectious pneumonia IS-A Pneumonia IS-A Lung disease.

Concepts may have multiple parents

- for example Infectious pneumonia is also an Infectious disease. The Concept graph must be acyclic — a parent cannot be its own child.

Concepts can have Roles

- e.g. Viral pneumonia has a role Causitive Agent which must be a Virus.

Some Concepts can be Primitive, such as Virus. But SNOMED also allows concepts to be Defined by a predicate.

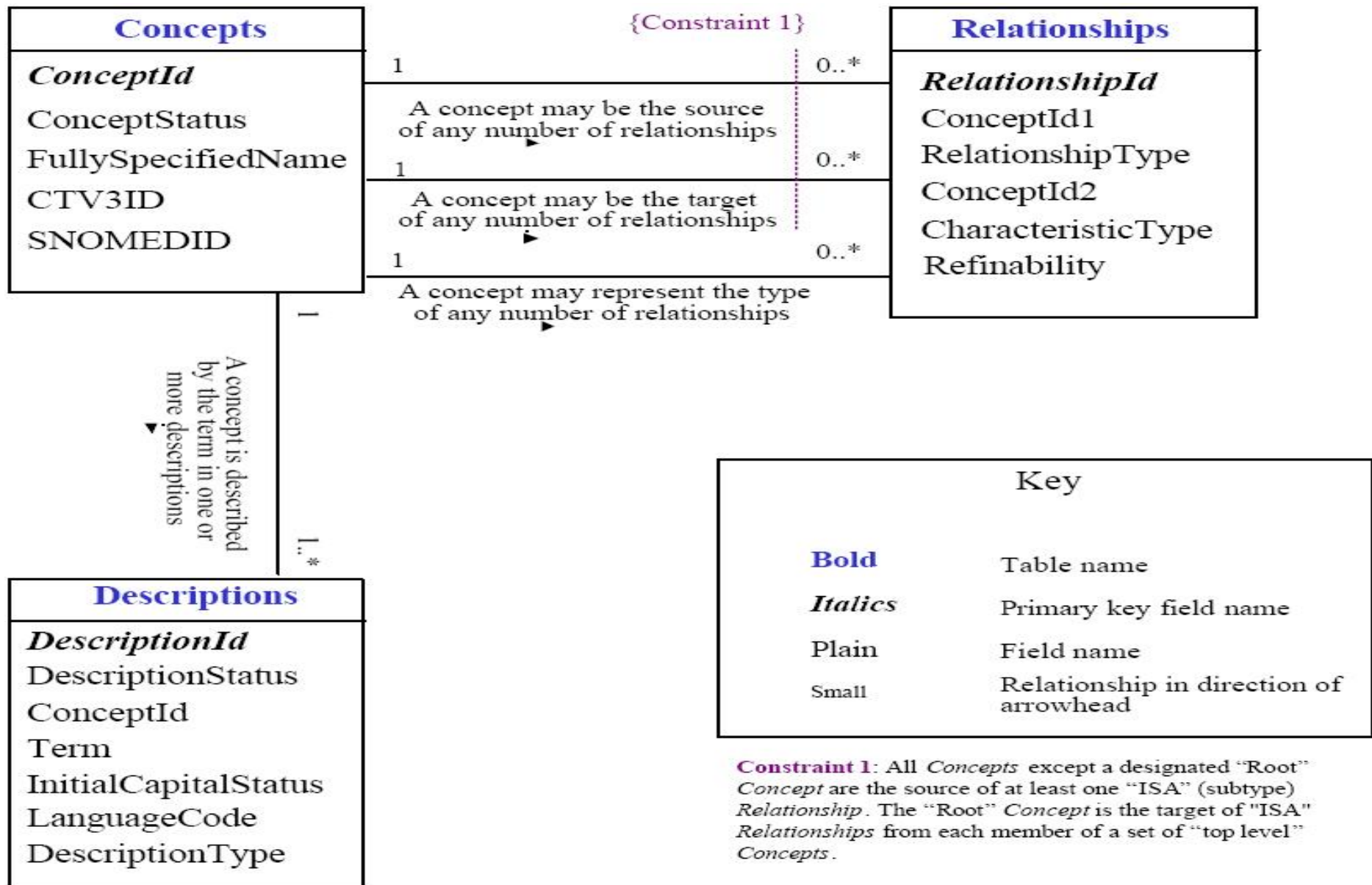
- For example Viral pneumonia might be defined as Pneumonia that is caused by a Virus. Defined concepts are based on Description Logic.

Descriptions are Terms or names (synonyms) assigned to a concept. Concepts often have several Descriptions, and a description may sometimes refer to more than one concept.

- For example, Immunosuppression might be a Therapy or a Finding.

Upper level concepts include Procedure, Pharmaceutical/biologic product, Clinical finding, Event, Body structure, Organism.

Entity-Relationship Diagram



SNOMED-CT Post-Coordination

For example, there might not be an explicit concept for a burn between the toes.

We can construct this using other SNOMED-CT terms.

```
284196006|Burn of skin|:  
  246112005|Severity|=24484000|severe,  
  363698007|Finding Site|=  
    (113185004|Structure of skin between fourth and fifth toes|:272741003|Laterality|=7771000|left)
```

Closer Look at Coding Schemes

LOINC
ICD
SNOMED-CT
UMLS

UMLS – Unified Medical Language System

Aims to allow intelligent retrieval from / integration of information from a wide range of disparate electronic biomedical information sources.

Aims to overcome variations in the way similar concepts are expressed in different sources to enable:

- linking of information from patient record systems, bibliographic databases, factual databases, expert systems, etc.
- indexing applications

UMLS Components

UMLS consists of three main components

Metathesaurus

- Concepts
- Inter-concept relationships

Semantic Network

- Semantic types
- Semantic network relationships

Lexical resources

- SPECIALIST Lexicon
- Lexical tools

UMLS – Not a Coding Scheme

- Includes machine-readable "Knowledge Sources" that can be used by a wide variety of applications programs to compensate for differences in the way concepts are expressed in different sources and by different users, to identify the information sources most relevant to a user inquiry.
- The Metathesaurus contains mappings to MeSH, ICD, SNOMED and a number of other coding systems.
- The UMLS is not itself a coding standard; it is a cross-referenced collection of standards and other data and knowledge sources.

UMLS – Metathesaurus

- Contains information about biomedical concepts and terms from a large number of controlled terminologies and thesauri
- Preserves the information encoded in the source vocabularies, such as the hierarchical contexts of the terms, their meanings and other attributes.
- Organized by concepts, which means that alternate names (synonyms, lexical variants, and translations) for the same meaning are all linked together as one concept
- Adds information to the concepts, including semantic types, definitions, and inter-concept relationships.

UMLS – Metathesaurus

- ◆ Synonymous terms clustered into a concept
- ◆ Preferred term
- ◆ Unique identifier (CUI)

Adrenal gland diseases	MeSH	D000307
Adrenal disorder	AOD	0000005418
Disorder of adrenal gland	Read	C15z.
Diseases of the adrenal glands	SNOMED	DB-70000

C0001621

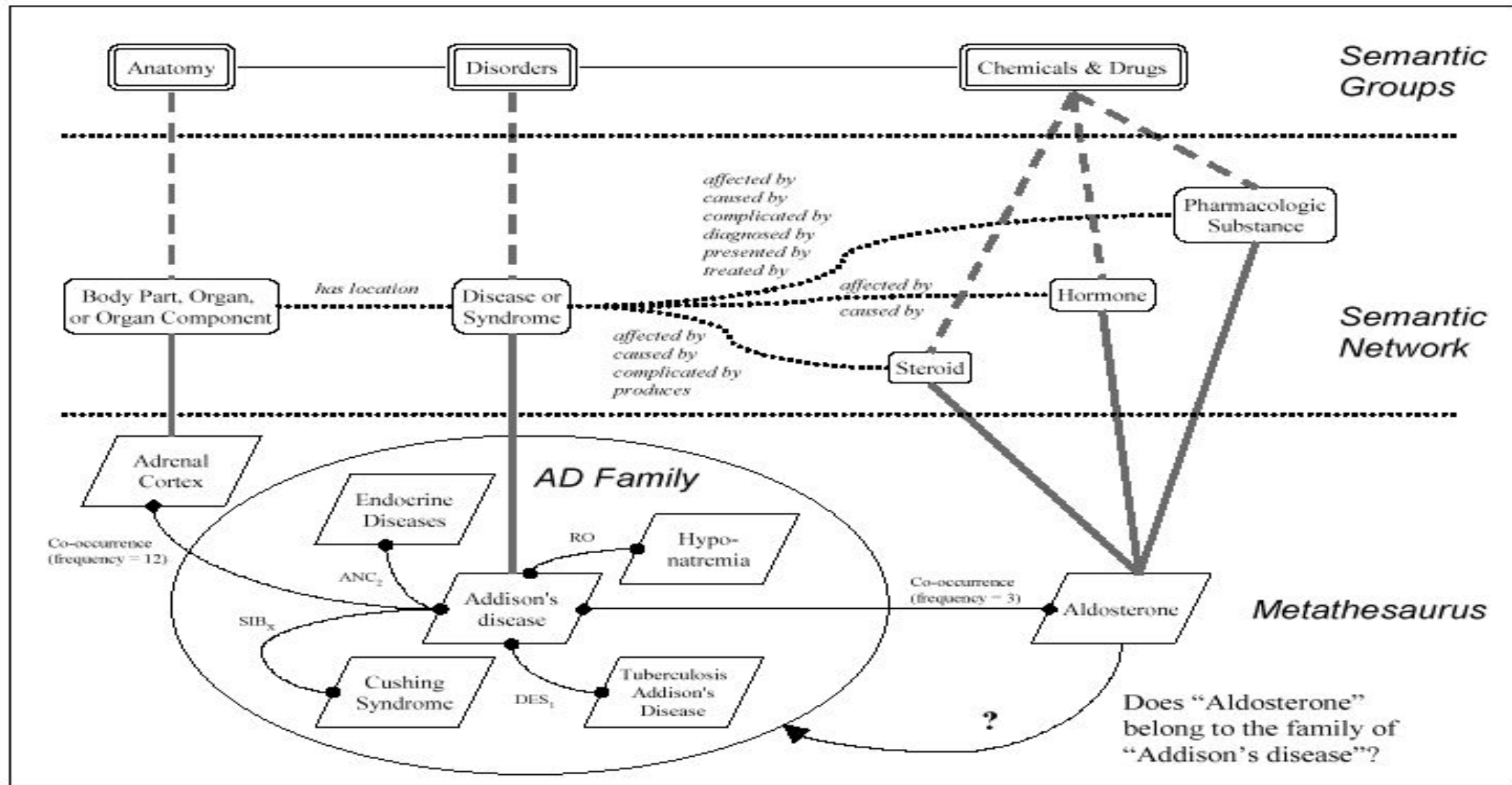
Adrenal Gland Diseases

Source: Bodenreider, Olivier; Willis, Jan; Hole, William. (2004)

UMLS – Semantic Network

- Provides a consistent categorization of all concepts represented in the Metathesaurus through semantic typing
- Links between the semantic types provide the structure for the Network and represent important relationships in the biomedical domain.
 - E.g. semantic types for organisms, anatomical structures, biologic function, chemicals, events, physical objects, and concepts or ideas.
- The primary relationship is the "is_a" link, and there are five major categories of additional relationships
 - Physical
 - Spatial
 - Temporal
 - Functional
 - Conceptual

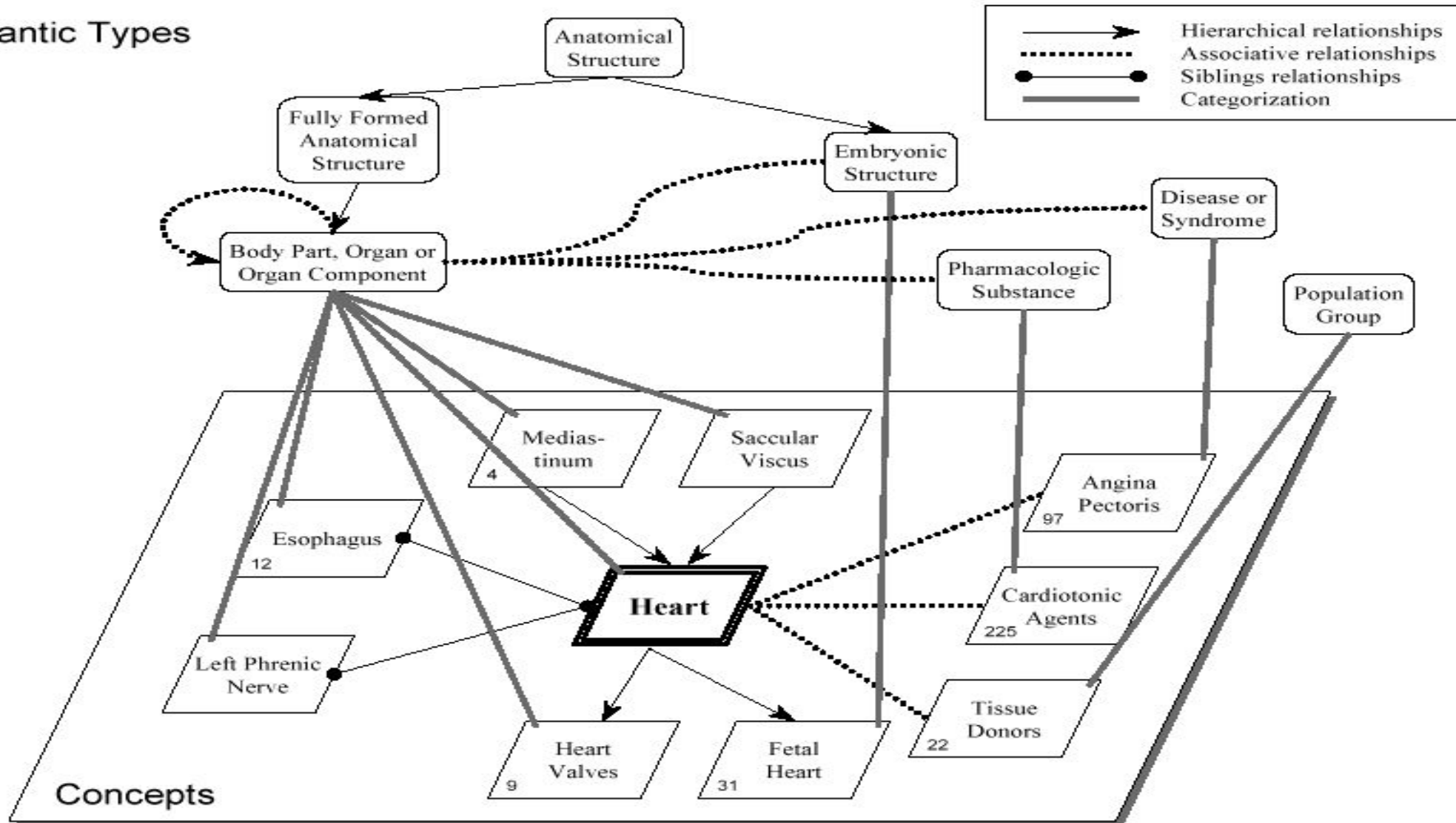
UMLS – Organisation



Overview of the methodology applied to the relationships of "Aldosterone" to "Addison's disease"

UMLS – Semantic Types

Semantic Types



UMLS SPECIALIST Lexicon

- Provides the lexical information needed for the SPECIALIST Natural Language Processing System (NLP).
- It is intended to be a general English lexicon that includes many biomedical terms.
- Coverage includes both commonly occurring English words and biomedical vocabulary. The lexicon entry for each word or term records the syntactic, morphological, and orthographic information needed by the SPECIALIST NLP System.
- The lexical programs or tools are designed to address the high degree of variability in natural language words and terms. Words often have several inflected forms which would properly be considered instances of the same word. The verb "treat", for example, has three inflectional variants:
 - treats — the third person singular present tense form
 - treated — the past and past participle form
 - treating — the present participle form

Potential Problems with Coding Schemes

Potential Problems with Coding Schemes

Understand the problems in creating / using medical terminologies

- Scalability
- Usability
- Ambiguity & Redundancy

Scalability

- (Particularly enumerative schemes) tend to become very large - millions of terms
- Hard to remember what's in there
- Hard to organise
- Hard to analyse
- Different people use the same code to mean different things

Usability

- Underlying collection of terms and concepts is large and complex in any (useful) scheme
- This needs to be organised for different purposes to make it usable
 - Navigation (how do people work?)
 - Statistical analysis (what reports do we need?)
 - Assigning meanings to codes
 - may mislead the user – misses the subtleties of the meaning of the code
 - Having limited code lengths (READ2)
 - Hard for the user – similar codes may have different meanings
 - “true universal” classification – depends on purpose

No Ambiguity (Disjoint Set)

Each concept should have one unambiguous preferred term (although it may have any number of synonyms)

- Synonyms may be shared with other concepts
 - e.g., "Ventricle" is a synonym (but cannot be the preferred term) of both "Cardiac ventricle" and "Brain ventricle".

No Redundancy (Disjoint Set)

Duplicate concepts should not be allowed

- i.e., cannot allow two different ways of coding the same thing or concept
- e.g., "Heart attack" and "Myocardial infarction" are synonyms, and the system should recognise this.
- If there are separate ways of representing the same concept this may lead to confusion on the part of someone entering the code, and leads to inaccuracy / omissions in search & retrieval
- Multi-axial systems can lead to redundancies & duplications in statistical accounting

Seemingly Redundant Concepts

- There may be *apparently* different ways of recording the same concept (e.g. to improve usability) but these in reality lead to the same underlying concept
- E.g. both “MI” and “heart attack” may exist as prompts on the same computer system, but they both point to the same underlying concept
- A clinician may record “femoral fracture” or “fracture of the femur” but the system recognises that these point to the same concept.

Compositional Nonsense

- Systems must not allow meaningless or ambiguous juxtapositions e.g. “compound fracture of the eyebrow”
- Either:
 - prevent nonsense combinations by attaching properties to concepts stating which combinations are/are not allowed
 - have tools that constrain the user
- In either case requires extrinsic knowledge about the world to be modelled

Keeping Up-to-date

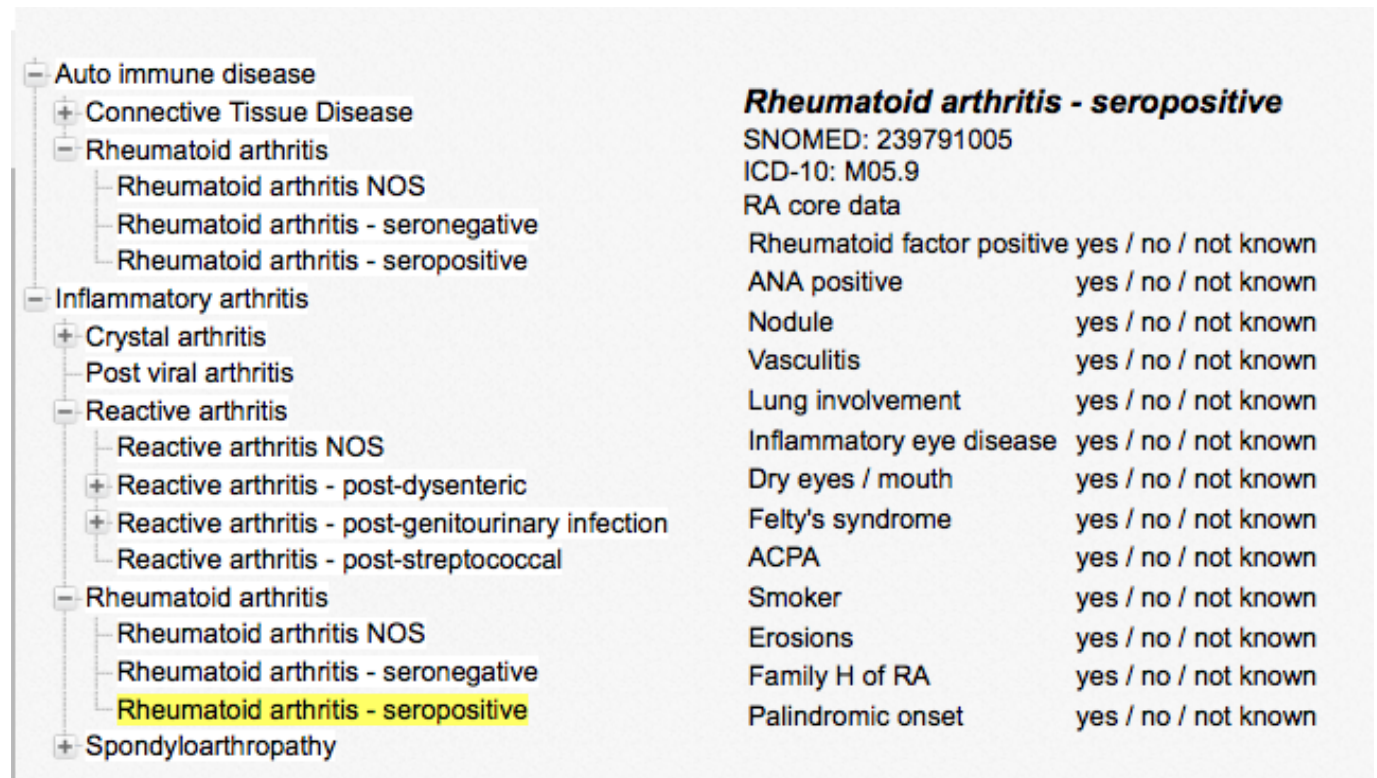
- Classifications become less accurate with time and need revision e.g., when new diseases are discovered (where to put these diseases, and if we put them under existing categories the meaning of these categories will drift with time making comparisons with previous years statistics done using the same classification less accurate and reliable).
- Updating a classification also implies preparing an equivalence-mapping table (to compare statistics done using different versions of the classification).

Too Many, Or Too Few?

Some people are of the opinion that there are already too many clinical coding schemes.

Why would anyone want to create a new system of coding?

- ORCHID class hierarchies are Directed Acyclic Graphs
- Used for diagnosis, medications, laboratory tests, etc
- Three-level, with Core Data Sets to represent phenotypes
- Nodes can also carry clinical coding (SNOMED, ICD-10, etc)



References and Further Reading

References

1. Kiley R. (1996) Medical information on the Internet. a guide for health professionals. New York, NY: Churchill Livingstone, 1996.

Further Reading

These resources have been downloaded to the City University Moodle site.

Kent Spackman (2007). SNOMED Clinical Terms Fundamentals. IHTSDO, December 14, 2007. Available online at:

http://www.ihtsdo.org/fileadmin/user_upload/Docs_01/SNOMED_Clinical_Terms_Fundamentals.pdf

World Health Organisation (2004). ICD-10: international statistical classification of diseases and related health problems: tenth revision. 2nd ed. ISBN 92 4 154653 0. Available online at: http://www.who.int/entity/classifications/icd/ICD-10_2nd_ed_volume2.pdf

Bodenreider, Olivier; Willis, Jan; Hole, William. (2004) The Unified Medical Language System: What is it and how to use it? MEDINFO, September 8, 2004; San Francisco, CA. Available online at:

http://www.nlm.nih.gov/research/umls/presentations/2004-medinfo_tut.pdf

Further Reading

Nelson, Stuart J.; Powell, Tammy; Humphreys, Betsy L. (2002) The Unified Medical Language System (UMLS) Project. In: Kent, Allen; Hall, Carolyn M., editors. Encyclopedia of Library and Information Science. New York: Marcel Dekker, Inc.; 2002. p.369-378. Available online at: <http://www.nlm.nih.gov/mesh/umlsforelis.html>

Clem McDonald, Stan Huff, Kathy Mercer, Jo Anna Hernandez, Daniel J. Vreeman (eds) (2010) Logical Observation Identifiers Names and Codes (LOINC®) Users' Guide. Regenstrief Institute, Inc. Available online at: <http://loinc.org/downloads/files/LOINCManual.pdf>

Chartered Society of Physiotherapy. (2005) Payment By Results. The new funding system for the NHS in England: practical support for Allied Health Professionals. Available online at: www.rcslt.org/docs/free-pub/payment_results.pdf

National Network of Libraries of Medicine (2010). Searching PubMed with MeSH. Available online at: <http://nmlm.gov/training/resources/meshtri.pdf>