NLM Georgia
Biomedical Informatics Course

Spring 2016 Course

Controlled Vocabularies
(aka Biomedical Terminologies/Ontologies)

Olivier Bodenreider

Lister Hill National Center
for Biomedical Communications
Bethesda, Maryland - USA

NIH U.S. National Library of Medicine

Learning objectives

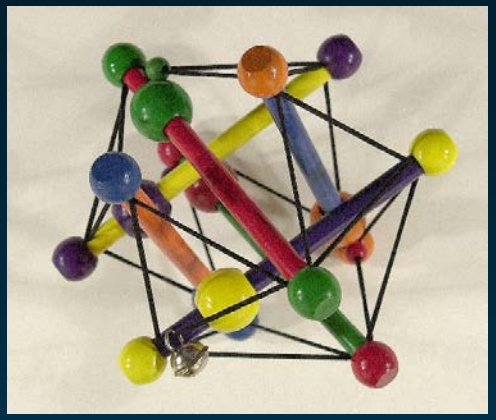
- Describe the history of biomedical ontologies
- Explain how clinical features are reflected in disease names
- List and describe the main biomedical ontologies used in 21st century healthcare
- Discuss the purpose of biomedical ontologies in knowledge management, clinical decision support and analytics

References Review articles

- Bodenreider O, Stevens R. Bio-ontologies: current trends and future directions. Brief Bioinform. 2006 Sep;7(3):256-74.
- Cimino JJ, Zhu X. The practical impact of ontologies on biomedical informatics. Yearb Med Inform. 2006:124-35.
- Bodenreider O. Biomedical ontologies in action: role in knowledge management, data integration and decision support. Yearb Med Inform. 2008:67-79.

Additional references

- Cimino JJ. Desiderata for controlled medical vocabularies in the twenty-first century. Methods Inf Med. 1998 Nov;37(4-5):394-403.
- Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. Nucleic Acids Res. 2004 Jan 1;32(Database issue):D267-70.



Medical Ontology Research

Contact: olivier@nlm.nih.gov
Web: https://mor.nlm.nih.gov

Olivier Bodenreider

Lister Hill National Center
for Biomedical Communications
Bethesda, Maryland - USA

NIH U.S. National Library of Medicine

Outline

- Historical perspective
- Introduction to biomedical terminologies through an example

---

- "High-Impact" Biomedical Ontologies
  - Structural perspective

- Biomedical Ontologies "in Action"
  - Functional perspective

Controlled Vocabularies
Part 1

Historical perspective

To support a theory of diseases

- Hippocrates
  - Dismisses superstition
  - Four humors
    - Blood
    - Phlegm
    - Yellow bile
    - Black bile
- Thomas Sydenham (1624-1689)
  - *Medical observations on the history and cure of acute diseases* (1676)

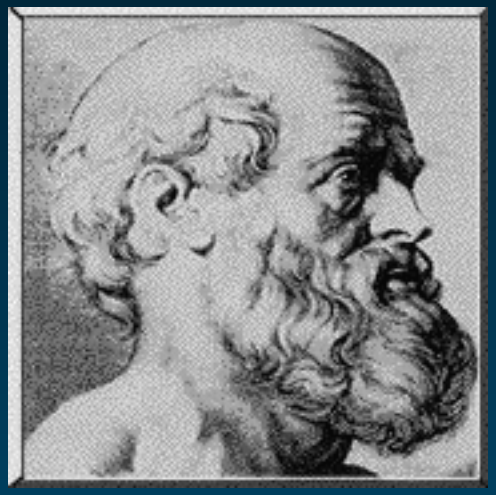
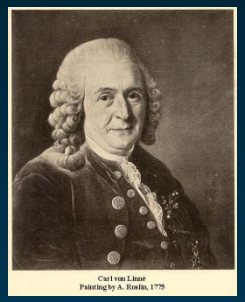


Figure 36 Thomas Sydenham (1624–1689)

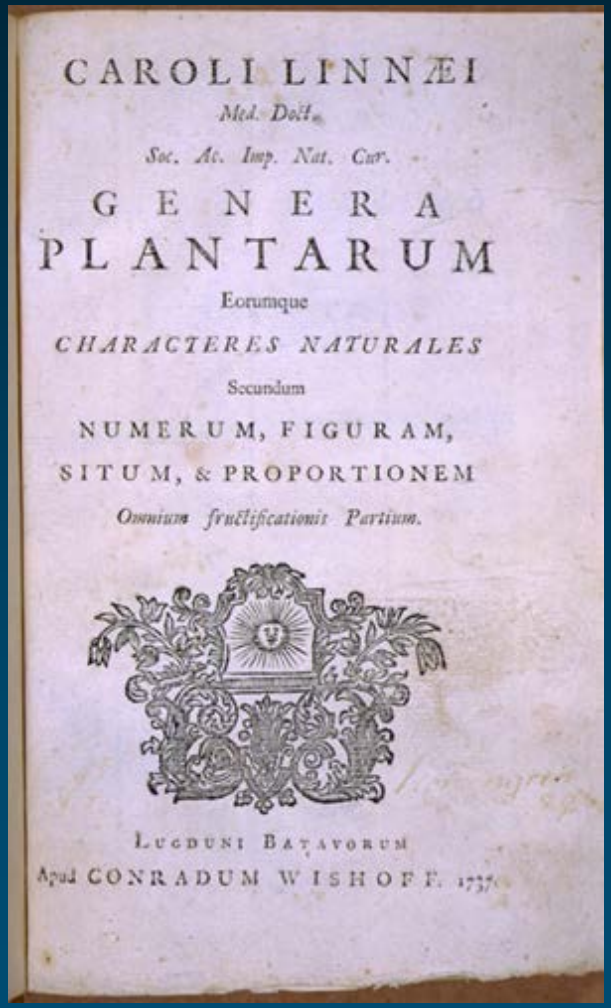
NLM
Lister Hill National Center for Biomedical Co

To classify diseases (and plants)

- Carolus Linnaeus (1707-1778)
  - *Genera Plantarum* (1737)
  - *Genera Morborum* (1763)
- François Boissier de La Croix a.k.a. F. B. de Sauvages (1706-1767)
  - *Methodus Foliorum* (1751)
  - *Nosologia Methodica* (1763/68)
- William Cullen (1710-1790)
  - *Synopsis Nosologiae Methodicae* (1785)



From plants…



CAROLI LINNÆI

Med. Doct.

Soc. Ac. Imp. Nat. Cur.

GENERA PLANTARUM

Eorumque

CHARACTERES NATURALES

Secundum

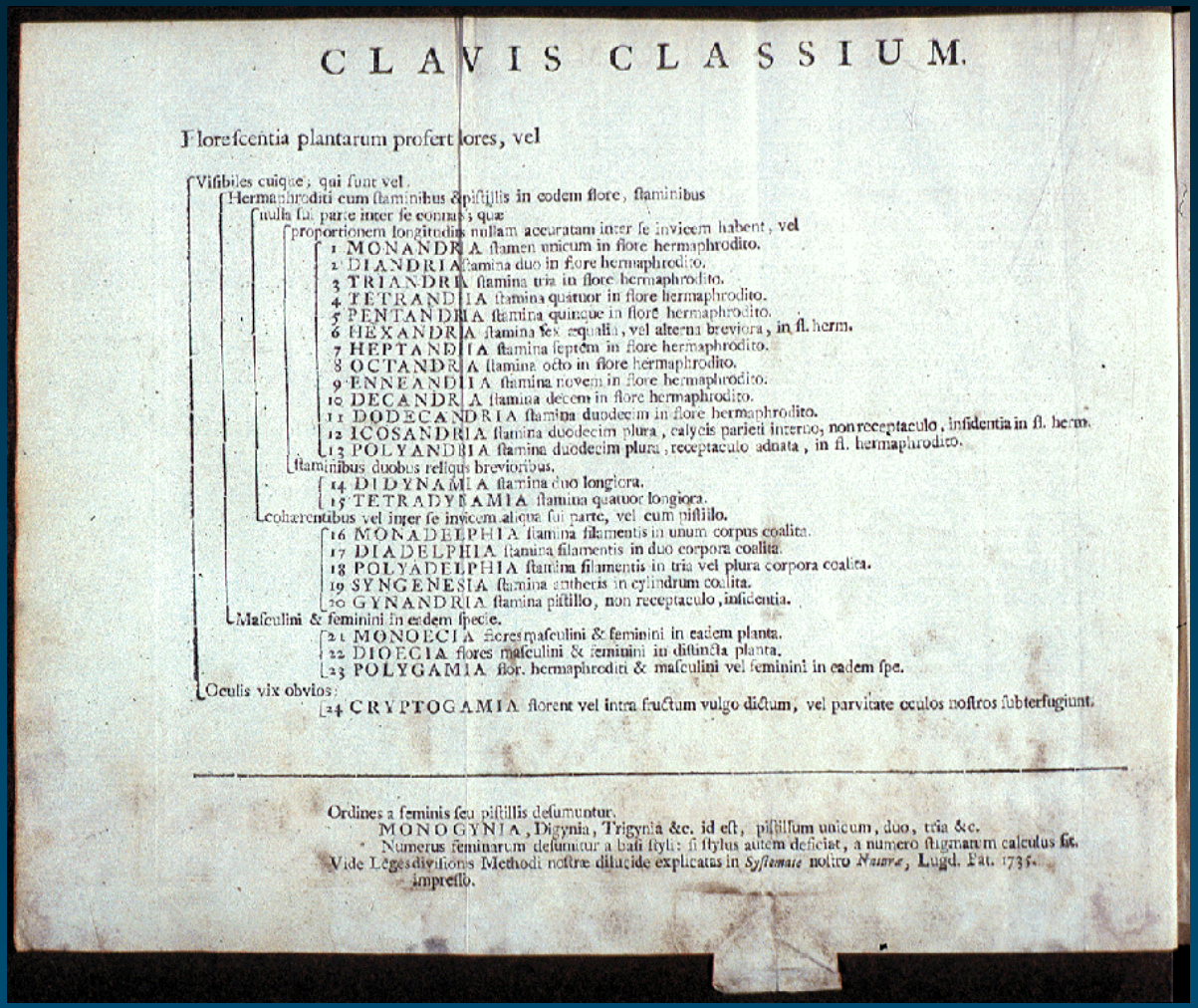
NUMERUM, FIGURAM,

SITUM, & PROPORTIONEM

Omnium fructificationis Partium.

LUGDUNI BATAVORUM

Apud CONRADUM WISHOFF. 1737.



CLAVIS CLASSIUM.

Florescentia plantarum profert flores, vel

Visibiles cuique; qui sunt vel:

Hermaphroditi cum staminibus & pistillis in eodem flore, staminibus

nulla sui parte inter se connatis; quæ

proportionem longitudinis nullam accuratam inter se invicem habent, vel

1 MONANDRIA stamen unicum in flore hermaphrodito.
2 DIANDRIA stamina duo in flore hermaphrodito.
3 TRIANDRIA stamina tria in flore hermaphrodito.
4 TETRANDRIA stamina quatuor in flore hermaphrodito.
5 PENTANDRIA stamina quinque in flore hermaphrodito.
6 HEXANDRIA stamina sex æqualia, vel altera breviora, in fl. herm.
7 HEPTANDRIA stamina septem in flore hermaphrodito.
8 OCTANDRIA stamina octo in flore hermaphrodito.
9 ENNEANDRIA stamina novem in flore hermaphrodito.
10 DECANDRIA stamina decem in flore hermaphrodito.
11 DODECANDRIA stamina duodecim in flore hermaphrodito.
12 ICOSANDRIA stamina duodecim plura, calycis parieti interno, non receptaculo, insidentia in fl. herm.
13 POLYANDRIA stamina duodecim plura, receptaculo adnata, in fl. hermaphrodito.

staminibus duobus reliqui brevioribus.

14 DIDYNAMIA stamina duo longiora.
15 TETRADYNAMIA stamina quatuor longiora.

cohærentibus vel inter se invicem aliqua sui parte, vel cum pistillo.

16 MONADELPHIA stamina filamentis in unum corpus coalita.
17 DIADELPHIA stamina filamentis in duo corpora coalita.
18 POLYADELPHIA stamina filamentis in tria vel plura corpora coalita.
19 SYNGENESIA stamina antheris in cylindrum coalita.
20 GYNANDRIA stamina pistillo, non receptaculo, insidentia.

Masculini & feminini in eadem specie.

21 MONOECIA flores masculini & feminini in eadem planta.
22 DIOECIA flores masculini & feminini in distincta planta.
23 POLYGAMIA flor. hermaphroditi & masculini vel feminini in eadem spe.

Oculis vix obvios:

24 CRYPTOGAMIA florent vel intra fructum vulgo dictum, vel parvitate oculos nostros subterfugiunt.

Ordines a feminis seu pistillis desumuntur.
MONOGYNIA, Digynia, Trigynia &c. id est, pistillum unicum, duo, tria &c.
Numerus feminarum desumitur a basi styli: si stylus autem deficiat, a numero stigmatum calculus fit.
Vide Leges divisionis Methodi nostræ dilucide explicatas in *Systemate* nostro *Naturæ*, Lugd. Bat. 1735. impresso.

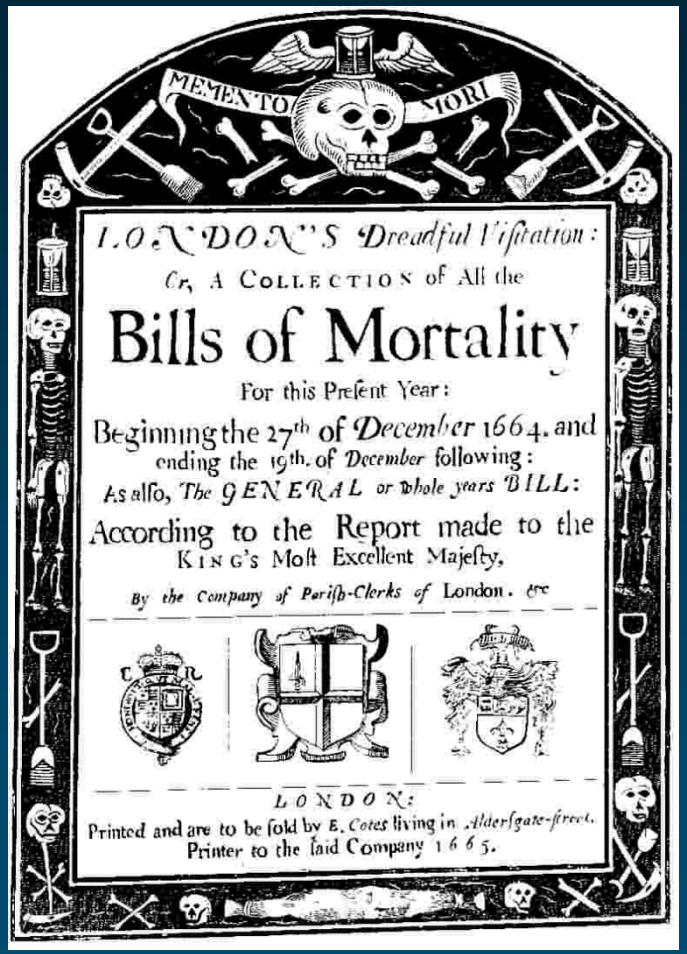
… to diseases

- Four categories (W. Cullen)
  - Fevers
  - Nervous disorders
  - Cachexias
  - Local diseases

> "The distinction of the genera of diseases, the distinction of the species of each, and often even that of the varieties, I hold to be a necessary foundation of every plan of physic, whether dogmatical or empirical."
> – William Cullen, Edinburgh, 1785
> *Synopsis Nosologia Methodicae*

(Cited by Chris Chute)

London Bills of Mortality



A generall Bill for this present year, ending the 19 of *December* 1665. according to the Report made to the KINGS most Excellent Majesty.

By the Company of Parish Clerks of *London*, &c.

The Diseases and Casualties this year.

Abortive and Stillborne	617	Executed	21	Palsie	50
Aged	1545	Flox and Small Pox	655	Plague	68596
Ague and Feaver	5257	Found dead in Streets, fields, &c.	2	Planet	6
Appoplex and Suddenly	116	French Pox	86	Pleurisie	15
Bedrid	12	Frighted	23	Poysoned	4
Blasted	5	Gout and Sciatica	27	Quinsie	35
Bleeding	16	Grief	46	Rickets	557
Bloody Flux, Scowring & Flux	185	Griping in the Guts	1288	Rising of the Lights	597
Burnt and Scalded	8	Hang'd & made away themselves	7	Rupture	34
Calenture	3	Headmould-shot & Mould-fallen	14	Scurvy	17
Cancer, Gangrene and Fistula	56	Jaundies	120	Shingles and Swine pox	2
Canker, and Thrush	211	Impostume	227	Sores, Ulcers, broken and bruised	
Childbed	625	Kill'd by severall accidents	46	Limbs	82
Chrisomes and Infants	1258	Kings Evill	86	Spleen	14
Cold and Cough	63	Leprosie	2	Spotted Feaver and Purples	1929
Collick and Winde	134	Lethargy	14	Stopping of the Stomack	334
Consumption and Tisick	4808	Livergrown	21	Stone and Strangury	98
Convulsion and Mother	1036	Measom and Headach	12	Suddat	2
Distracted	5	Meazles	7	Teeth and Worms	2614
Dropsie and Tympany	1478	Murthered and Shot	9	Vomiting	54
Drowned	53	Overlaid & Starved	45	Worm	7

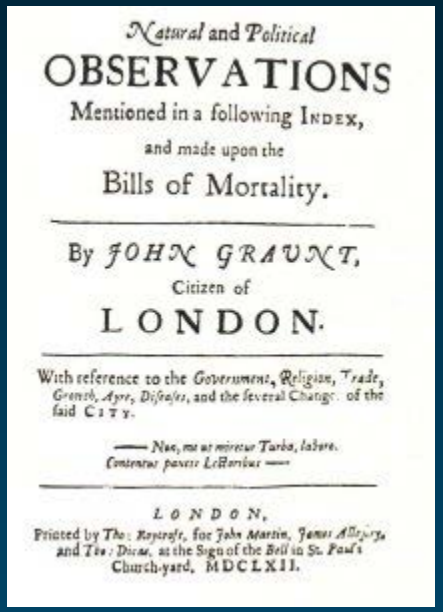
Christned	Males	5114	Buried	Males	48568
	Females	4853		Females	48737
	In all	9967		In all	97305

Of the Plague — 68596

Increased in the Burials in the 130 Parishes and at the Pest-house this year — 79609
Increased of the Plague in the 130 Parishes and at the Pest-house this year — 68590

To support epidemiology

- John Graunt (1620-1674)
  - Analyzes the vital statistics of the citizens of London
- William Farr (1807-1883)
  - Medical statistician
  - Improves Cullen's classification
  - Contributes to creating ICD
- Jacques Berthillon (1851-1922)
  - Chief of the statistical services (Paris)
  - Classification of causes of death (161 rubrics)



Natural and *Political*

OBSERVATIONS

Mentioned in a following INDEX,

and made upon the

Bills of Mortality.

By JOHN GRAUNT,

Citizen of

LONDON.

With reference to the *Government*, *Religion*, *Trade*, *Growth*, *Ayre*, *Diseases*, and the several Changes of the said CITY.

—— *Nec, me ut miretur Turba, laboro,*
Contentus paucis Lectoribus ——

LONDON,

Printed by *Tho: Roycroft*, for *John Martin*, *James Allestry*, and *Tho: Dicas*, at the Sign of the *Bell* in St. *Paul's* Church-yard, MDCLXII.

Limitations of existing classifications

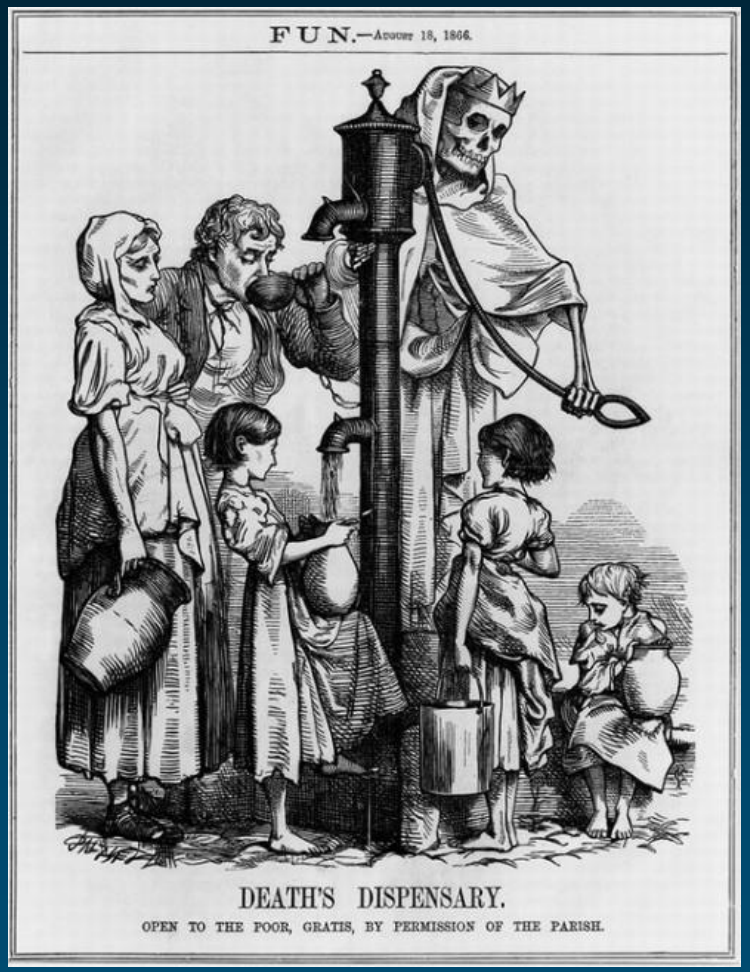
"The advantages of a uniform statistical nomenclature, however imperfect, are so obvious, that it is surprising no attention has been paid to its enforcement in Bills of Mortality. Each disease has, in many instances, been denoted by three or four terms, and each term has been applied to as many different diseases: vague, inconvenient names have been employed, or complications have been registered instead of primary diseases. The nomenclature is of as much importance in this department of inquiry as weights and measures in the physical sciences, and should be settled without delay."

– William Farr

First annual report.

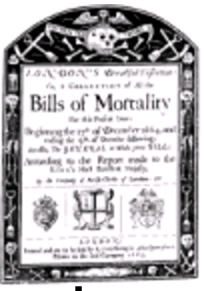
London, Registrar General of England and Wales, 1839, p. 99.

From “bad air” to “bad water” (John Snow)



[Bodenreider, BIB 2006]

History of Medical Ontologies



Synopsis Nosologiae Methodicae

SNOP
CPT
OPCS
EmTree
MeSH
ICD
ICD9

1603 1700 1785 1855 1900 1975

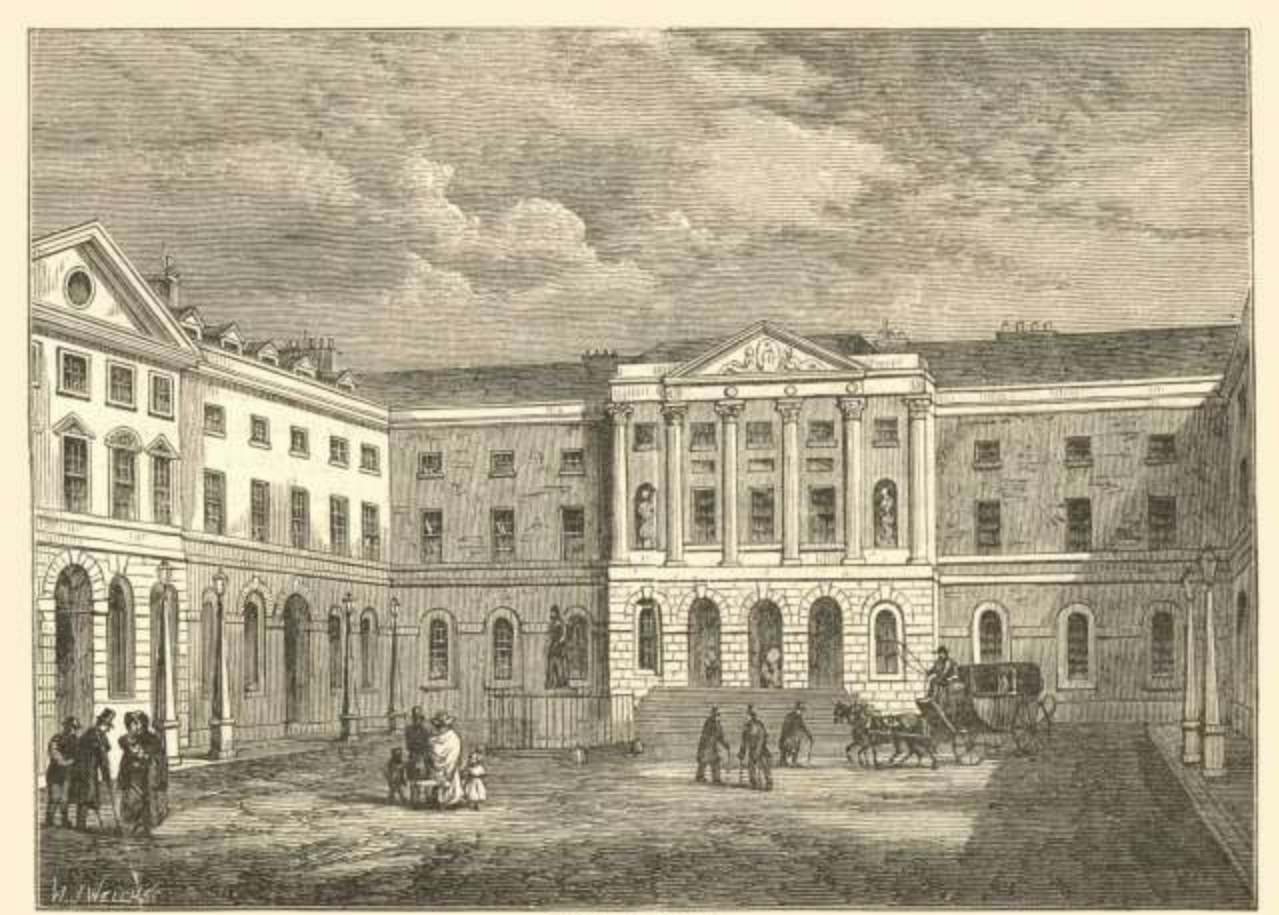
FMA
GALEN
DM&D
UMLS
OPCS
OPCS3
OPCS4
OPCS4.3
CTV3
READ
ICPC
SNOMED International
SNOMED-2
SNOMED-RT
SNOMED-CT
SNOP
CPT
MESH
ICD

1975 1985 1995 2005

Controlled Vocabularies
Part 2

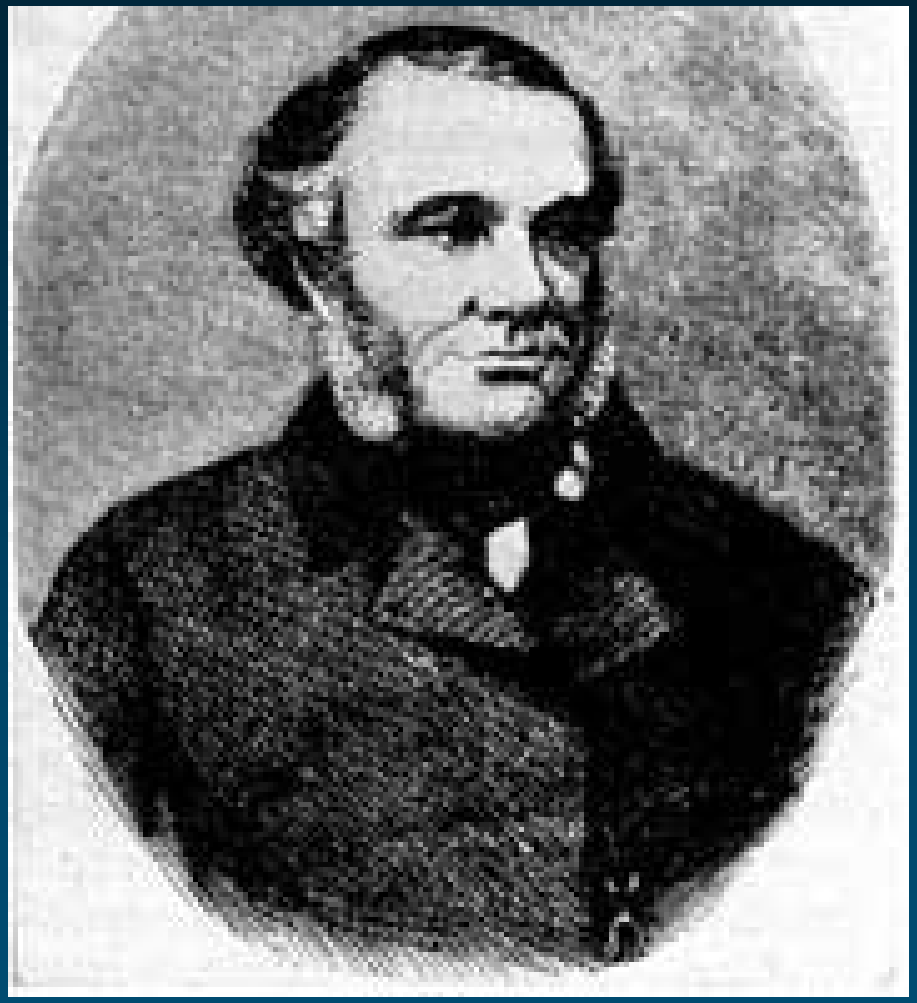
Introduction to biomedical terminologies through an example

Guy's Hospital, London



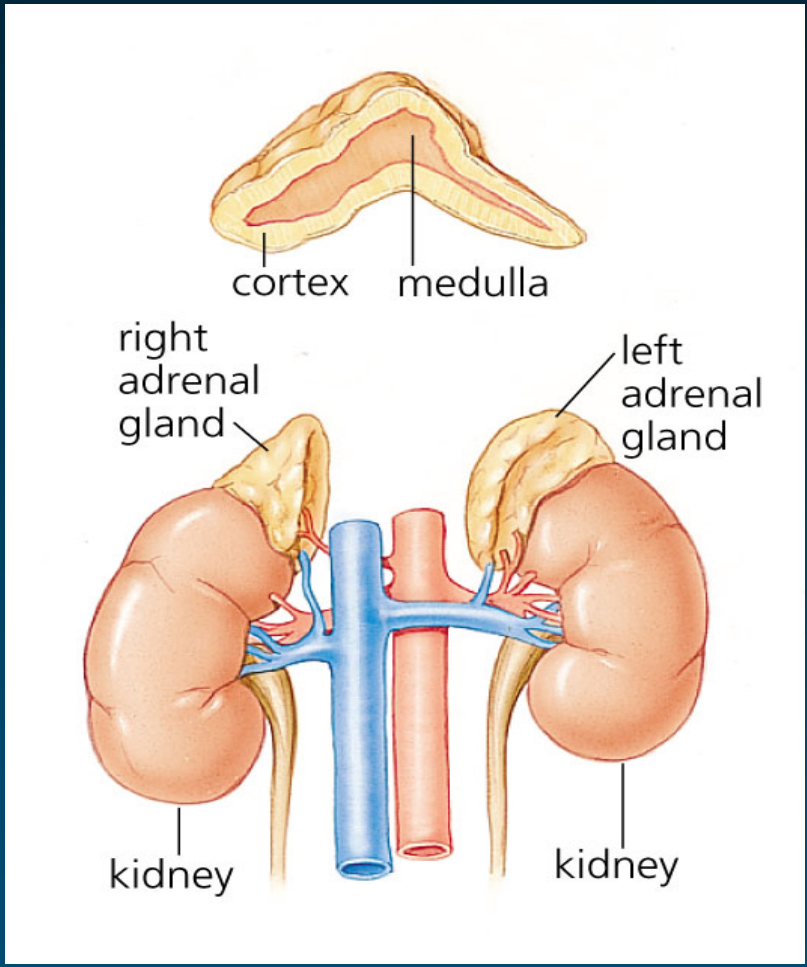
GUY'S HOSPITAL.

Thomas Addison (1795-1860)



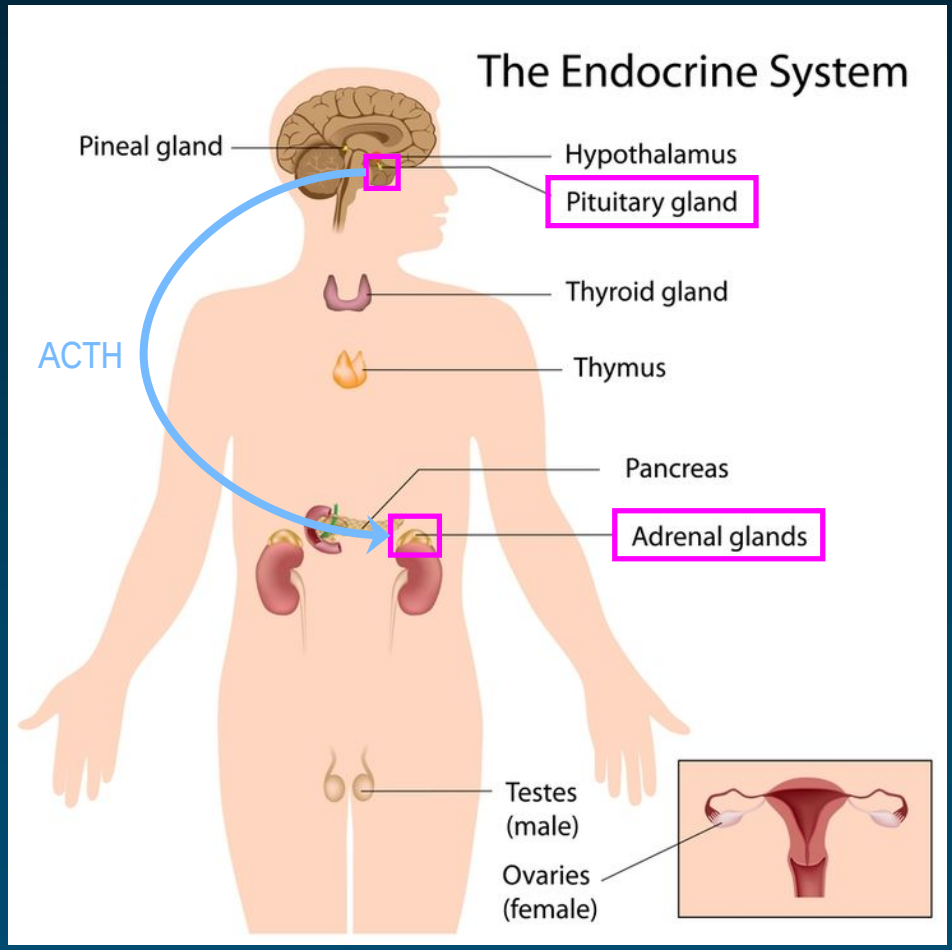
Addison’s disease

- Addison's disease is a rare endocrine disorder
- Addison's disease occurs when the adrenal glands do not produce enough of the hormone cortisol
- For this reason, the disease is sometimes called chronic adrenal insufficiency, or hypocortisolism



Adrenal insufficiency Clinical variants

- Primary / Secondary
  - Primary: lesion of the adrenal glands themselves
  - Secondary: inadequate secretion of ACTH by the pituitary gland
- Acute / Chronic
- Isolated / Polyendocrine deficiency syndrome



Addison's disease: Symptoms

- Fatigue
- Weakness
- Low blood pressure
- Pigmentation of the skin (exposed and non-exposed parts of the body)
- …

AD in medical vocabularies

- Synonyms: different terms
  - Addisonian syndrome — eponym
  - Bronzed disease — symptoms
  - Addison melanoderma — symptoms
  - Asthenia pigmentosa — symptoms
  - Primary adrenal deficiency — clinical variants
  - Primary adrenal insufficiency — clinical variants
  - Primary adrenocortical insufficiency — clinical variants
  - Chronic adrenocortical insufficiency — clinical variants
- Contexts: different hierarchies

Internal Classification of Diseases

- IV Endocrine, nutritional and metabolic diseases
  - E00-E07 Disorders of thyroid gland
  - E10-E14 Diabetes mellitus
  - E15-E16 Other disorders of glucose regulation and pancreatic internal secretion
  - E20-E35 Disorders of other endocrine glands
    - E20 Hypoparathyroidism
    - E21 Hyperparathyroidism and other disorders of parathyroid gland
    - E22 Hyperfunction of pituitary gland
    - E23 Hypofunction and other disorders of pituitary gland
    - E24 Cushing syndrome
    - E25 Adrenogenital disorders
    - E26 Hyperaldosteronism
    - E27 Other disorders of adrenal gland
      - E27.0 Other adrenocortical overactivity
      - E27.1 Primary adrenocortical insufficiency
      - E27.2 Addisonian crisis
      - E27.3 Drug-induced adrenocortical insufficiency
      - E27.4 Other and unspecified adrenocortical insufficiency
      - E27.5 Adrenomedullary hyperfunction
      - E27.8 Other specified disorders of adrenal gland
      - E27.9 Disorder of adrenal gland, unspecified
    - E28 Ovarian dysfunction
    - E29 Testicular dysfunction
    - E30 Disorders of puberty, not elsewhere classified
    - E31 Polyglandular dysfunction
    - E32 Diseases of thymus
    - E34 Other endocrine disorders
    - E35 Disorders of endocrine glands in diseases classified elsewhere

E27 Other disorders of adrenal gland

E27.0 Other adrenocortical overactivity

Overproduction of ACTH, not associated with Cushing disease

Premature adrenarche

Excl.: Cushing syndrome (E24.-)

E27.1 Primary adrenocortical insufficiency

Addison disease

Autoimmune adrenalitis

Excl.: amyloidosis (E85.-)
tuberculous Addison disease (A18.7)
Waterhouse-Friderichsen syndrome (A39.1)

E27.2 Addisonian crisis

Adrenal crisis

Adrenocortical crisis

E27.3 Drug-induced adrenocortical insufficiency

Use additional external cause code (Chapter XX), if desired, to identify drug.

E27.4 Other and unspecified adrenocortical insufficiency

Adrenal:
- haemorrhage
- infarction

Adrenocortical insufficiency NOS

Hypoaldosteronism

Excl.: adrenoleukodystrophy [Addison-Schilder] (E71.3)
Waterhouse-Friderichsen syndrome (A39.1)

E27.5 Adrenomedullary hyperfunction

Adrenomedullary hyperplasia

Catecholamine hypersecretion

E27.8 Other specified disorders of adrenal gland

Abnormality of cortisol-binding globulin

E27.9 Disorder of adrenal gland, unspecified

Medical Subject Headings

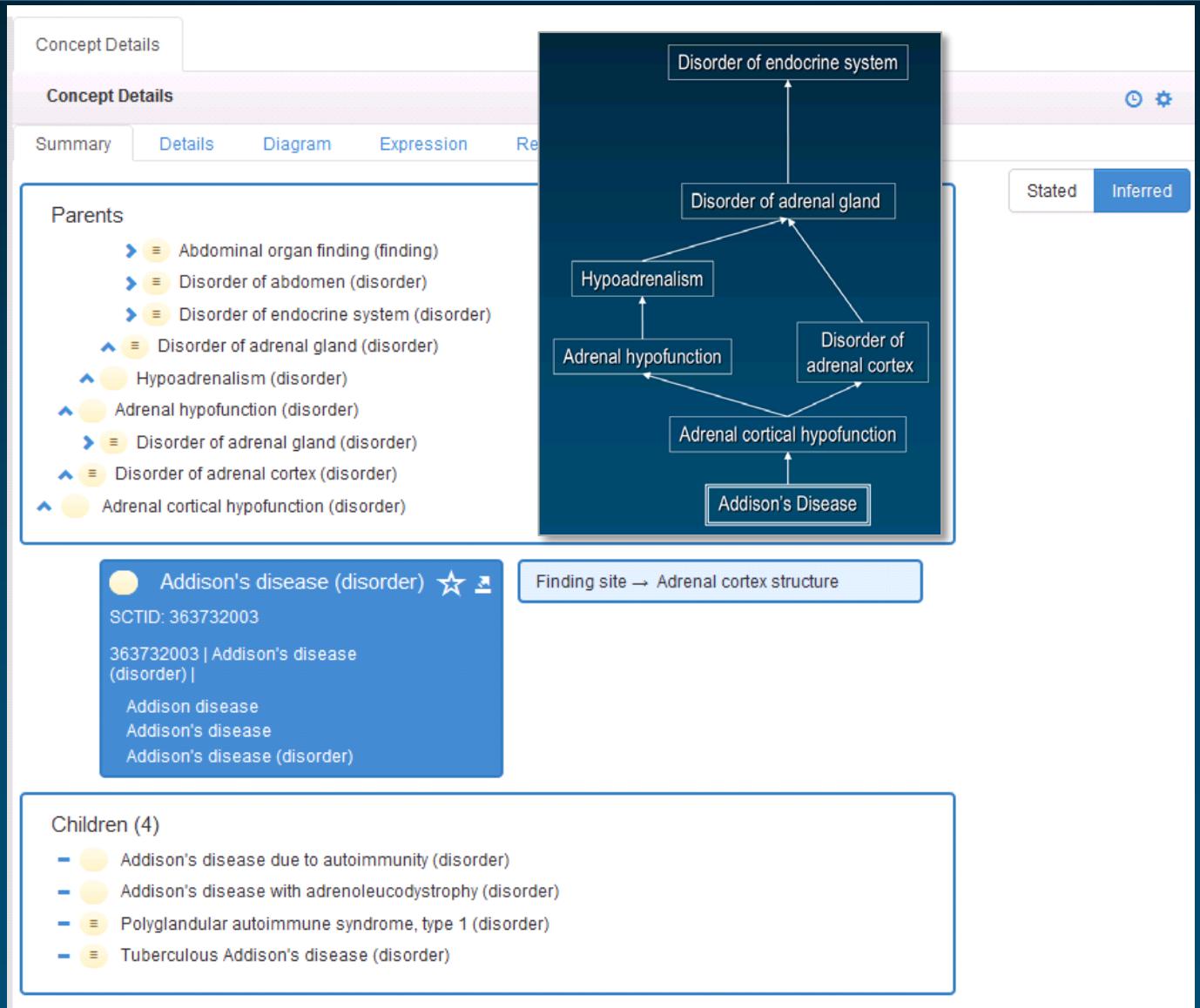
MeSH Tree Structures

Endocrine System Diseases [C19]
- Adrenal Gland Diseases [C19.053]
  - Adrenal Insufficiency [C19.053.500]
    - Addison Disease [C19.053.500.263]
    - Adrenoleukodystrophy [C19.053.500.270]
    - Hypoaldosteronism [C19.053.500.480]
    - Waterhouse-Friderichsen Syndrome [C19.053.500.740]

Immune System Diseases [C20]
- Autoimmune Diseases [C20.111]
  - Addison Disease [C20.111.163]
  - Anemia, Hemolytic, Autoimmune [C20.111.175]
  - Anti-Glomerular Basement Membrane Disease [C20.111.190]
  - Anti-Neutrophil Cytoplasmic Antibody-Associated Vasculitis [C20.111.193] +
  - Antiphospholipid Syndrome [C20.111.197]
  - Arthritis, Juvenile [C20.111.198]
  - Arthritis, Rheumatoid [C20.111.199] +
  - Autoimmune Diseases of the Nervous System [C20.111.258] +

[...]

SNOMED CT


Concept Details
Concept Details
Summary
Details
Diagram
Expression
Stated
Inferred
Parents
Abdominal organ finding (finding)
Disorder of abdomen (disorder)
Disorder of endocrine system (disorder)
Disorder of adrenal gland (disorder)
Hypoadrenalism (disorder)
Adrenal hypofunction (disorder)
Disorder of adrenal gland (disorder)
Disorder of adrenal cortex (disorder)
Adrenal cortical hypofunction (disorder)
Disorder of endocrine system
Disorder of adrenal gland
Hypoadrenalism
Adrenal hypofunction
Disorder of adrenal cortex
Adrenal cortical hypofunction
Addison's Disease
Addison's disease (disorder)
SCTID: 363732003
363732003 | Addison's disease (disorder) |
Addison disease
Addison's disease
Addison's disease (disorder)
Finding site → Adrenal cortex structure
Children (4)
Addison's disease due to autoimmunity (disorder)
Addison's disease with adrenoleucodystrophy (disorder)
Polyglandular autoimmune syndrome, type 1 (disorder)
Tuberculous Addison's disease (disorder)

Controlled Vocabularies
Part 3

"High-Impact" Biomedical Ontologies

A Structural Perspective

Overview

- Structural perspective
  - What are they (vs. what are they for)?
- "High-impact" biomedical ontologies [J. Cimino, YBMI 2006]
  - International Classification of Diseases (ICD)
  - Logical Observation Identifiers, Names and Codes (LOINC)
  - SNOMED Clinical Terms
  - Foundational Model of Anatomy
  - Gene Ontology
  - RxNorm
  - Medical Subject Headings (MeSH)
  - NCI Thesaurus
  - Unified Medical Language System (UMLS)

International Classification of Diseases

ICD Characteristics (1)

- Current version: ICD-10 (2016)
  - Annual updates
- Type: Classification
- Domain: Disorders
- Developer: World Health Organization (WHO)
- Funding: WHO
- Publicly available: Yes
- Used for: Mortality and morbidity statistics worldwide
- URL: http://www.who.int/classifications/icd/en/

ICD Characteristics (2)

- Number of
  - Concepts: 12,320 (ICD-10, 2004)
  - Terms: 1 per concept (tabular)
- Major organizing principles:
  - Tree (single inheritance hierarchy)
  - No explicit classification criteria
    - Idiosyncratic inclusion/exclusion mechanism
  - .8 slots for *Not elsewhere classified* (NEC)
  - .9 slots for *Not otherwise specified* (NOS)
- Specific coding rules
- Distribution: Proprietary format

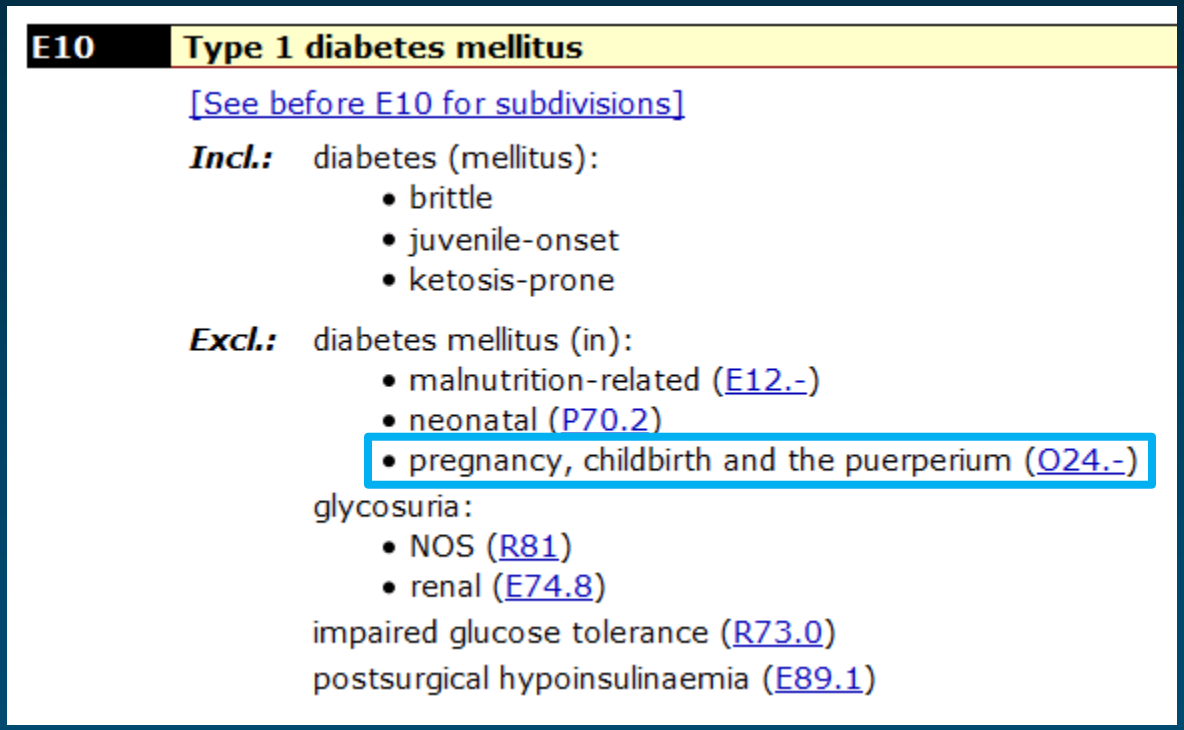
http://apps.who.int/classifications/icd10/browse/

ICD Top level

- **ICD-10 Version:2016**
  - I Certain infectious and parasitic diseases
  - II Neoplasms
  - III Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
  - IV Endocrine, nutritional and metabolic diseases
  - V Mental and behavioural disorders
  - VI Diseases of the nervous system
  - VII Diseases of the eye and adnexa
  - VIII Diseases of the ear and mastoid process
  - IX Diseases of the circulatory system
  - X Diseases of the respiratory system
  - XI Diseases of the digestive system
  - XII Diseases of the skin and subcutaneous tissue
  - XIII Diseases of the musculoskeletal system and connective tissue
  - XIV Diseases of the genitourinary system
  - XV Pregnancy, childbirth and the puerperium
  - XVI Certain conditions originating in the perinatal period
  - XVII Congenital malformations, deformations and chromosomal abnormalities
  - XVIII Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified
  - XIX Injury, poisoning and certain other consequences of external causes
  - XX External causes of morbidity and mortality
  - XXI Factors influencing health status and contact with health services
  - XXII Codes for special purposes

ICD Example

- Idiosyncratic inclusion/exclusion criteria



E10 Type 1 diabetes mellitus

[See before E10 for subdivisions]

Incl.: diabetes (mellitus):
- brittle
- juvenile-onset
- ketosis-prone

Excl.: diabetes mellitus (in):
- malnutrition-related (E12.-)
- neonatal (P70.2)
- pregnancy, childbirth and the puerperium (O24.-)

glycosuria:
- NOS (R81)
- renal (E74.8)

impaired glucose tolerance (R73.0)

postsurgical hypoinsulinaemia (E89.1)

ICD Example

- *Not elsewhere classified* (NEC)
- *Not otherwise specified* (NOS)

E84	**Cystic fibrosis**
	Incl.: mucoviscidosis
E84.0	**Cystic fibrosis with pulmonary manifestations**
E84.1	**Cystic fibrosis with intestinal manifestations**
	Distal intestinal obstruction syndrome
	Meconium ileus in cystic fibrosis† (P75*)
	Excl.: meconium obstruction (ileus) in cases where cystic fibrosis is known not to be present (P76.0)
E84.8	**Cystic fibrosis with other manifestations**
E84.9	**Cystic fibrosis, unspecified**

ICD-10-CM

- Derived from: ICD-10
  - Finer-grained (both clinically and administratively)
- Type: Classification
  - 92,042 codes (2015)
  - Terms: 1.2 per concept
- Domain: Disorders
- Developer: National Center for Health Statistics (NCVHS)
- Funding: U.S. Government
- Publicly available: Yes
- Used for: Billing
- URL: http://www.cdc.gov/nchs/icd/icd10cm.htm

ICD-10 vs. ICD-10-CM

E72 **Other disorders of amino-acid metabolism**

Excl.: abnormal findings without manifest disease (R7
disorders of:
- aromatic amino-acid metabolism (E70.-)
- branched-chain amino-acid metabolism (
- fatty-acid metabolism (E71.3)
- purine and pyrimidine metabolism (E79.-

gout (M10.-)

E72.0 **Disorders of amino-acid transport**

Cystine storage disease† (N29.8*)
Cystinosis
Cystinuria
Fanconi(-de Toni)(-Debré) syndrome
Hartnup disease
Lowe syndrome

Excl.: disorders of tryptophan metabolism (E70.8)

x6

E72 **Other disorders of amino-acid metabolism**

Excludes1: disorders of:
aromatic amino-acid metabolism (E70.-)
branched-chain amino-acid metabolism (E71.0-E71.2)
fatty-acid metabolism (E71.3)
purine and pyrimidine metabolism (E79.-)
gout (M1A.-, M10.-)

E72.0 **Disorders of amino-acid transport**

Excludes1: disorders of tryptophan metabolism (E70.5)

E72.00 **Disorders of amino-acid transport, unspecified**

E72.01 **Cystinuria**

E72.02 **Hartnup's disease**

E72.03 **Lowe's syndrome**

Use additional code for associated glaucoma (H42)

E72.04 **Cystinosis**
Fanconi (-de Toni) (-Debré) syndrome with cystinosis

Excludes1: Fanconi (-de Toni) (-Debré) syndrome wit

E72.09 **Other disorders of amino-acid transport**
Fanconi (-de Toni) (-Debré) syndrome, unspecified

ICD-10 vs. ICD-10-CM

W58 Bitten or struck by crocodile or alligator

W58 Contact with crocodile or alligator

The appropriate 7th character is to be added to each code from category W58
- A - initial encounter
- D - subsequent encounter
- S - sequela

W58.0 Contact with alligator

- **W58.01 Bitten by alligator**
  - W58.01A Bitten by alligator, initial encounter
  - W58.01D Bitten by alligator, subsequent encounter
  - W58.01S Bitten by alligator, sequela
- **W58.02 Struck by alligator**
- **W58.03 Crushed by alligator**
- **W58.09 Other contact with alligator**

W58.1 Contact with crocodile

- **W58.11 Bitten by crocodile**
- **W58.12 Struck by crocodile**
- **W58.13 Crushed by crocodile**
- **W58.19 Other contact with crocodile**

x24

Logical Observation Identifiers, Names and Codes (LOINC)

LOINC Characteristics (1)

- Current version: 2.54 (Dec. 2015)
  - 2 annual releases
- Type: Controlled terminology*
- Domain: Laboratory and clinical observations
- Developer: Regenstrief Institute
- Funding: NLM and other sources
- Publicly available: Yes
- Used for: information exchange
- URL: https://loinc.org/

LOINC Characteristics (2)

- Number of
  - Concepts: 73,958 active codes (2.52, June 2015)
  - Terms: 1 per concept (“long name”)
- Major organizing principles:
  - No hierarchical structure among the main codes
  - 6 axes
    - Component (analyte [+ challenge] [+ adjustments])
    - Property
    - Timing
    - System
    - Scale
    - [Method]
- Distribution: proprietary database format

LOINC Example

- *Sodium [Moles/volume] in Serum or Plasma*
  [the molar concentration of sodium is measured in the plasma (or serum), with quantitative result]

Axis	Value
Component	Sodium
Property	SCnc – Substance Concentration (per volume)
Timing	Pt – Point in time (Random)
System	Ser/Plas – Serum or Plasma
Scale	Qn – Quantitative
Method	--

2951-2 Sodium [Moles/volume] in Serum or Plasma

NAME

Fully-Specified Name:

Component	Property	Time	System	Scale	Method
Sodium	SCnc	Pt	Ser/Plas	Qn	

PART DEFINITION/DESCRIPTION(S)

Sodium is an essential nutrient that regulates blood volume, blood pressure, osmotic equilibrium and electrolyte balance. Sodium chloride is the principal source of sodium in the diet, and is used for seasoning and as a preservative. Increased levels of sodium intake can cause hypertension and reportedly leads to 7.6 million premature deaths worldwide. Sodium is also important in neuron function and osmoregulation between cells and the extracellular fluid.

Copyright: Text is available under the Creative Commons Attribution/Share-Alike License. See http://creativecommons.org/licenses/by-sa/3.0/ for details.

Source: Wikipedia, **URL:** Sodium (Wikipedia)

BASIC ATTRIBUTES

Class/Type:	CHEM/Lab
CDISC Lab Test:	Y
Common Lab Results Rank:	#5
Common SI Lab Results Rank:	#5
Common Orders Rank:	#107
Last Updated in Version:	2.34
Order vs. Obs.:	Both
Status:	Active

EXAMPLE UNITS

Unit	Source Type
mmol/L	EXAMPLE UCUM UNITS
mmol/L	REGENSTRIEF
mmol/L	eCHN

UNITS AND RANGE

Range	Units Type
mmol/L:[136,145]	

Copyright © 2015 Regenstrief Institute, Inc. All Rights Reserved. To the extent included herein, the LOINC table and LOINC codes are copyright © 1995-2015, Regenstrief Institute, Inc. and the Logical Observation Identifiers Names and Codes (LOINC) Committee.

SNOMED Clinical Terms

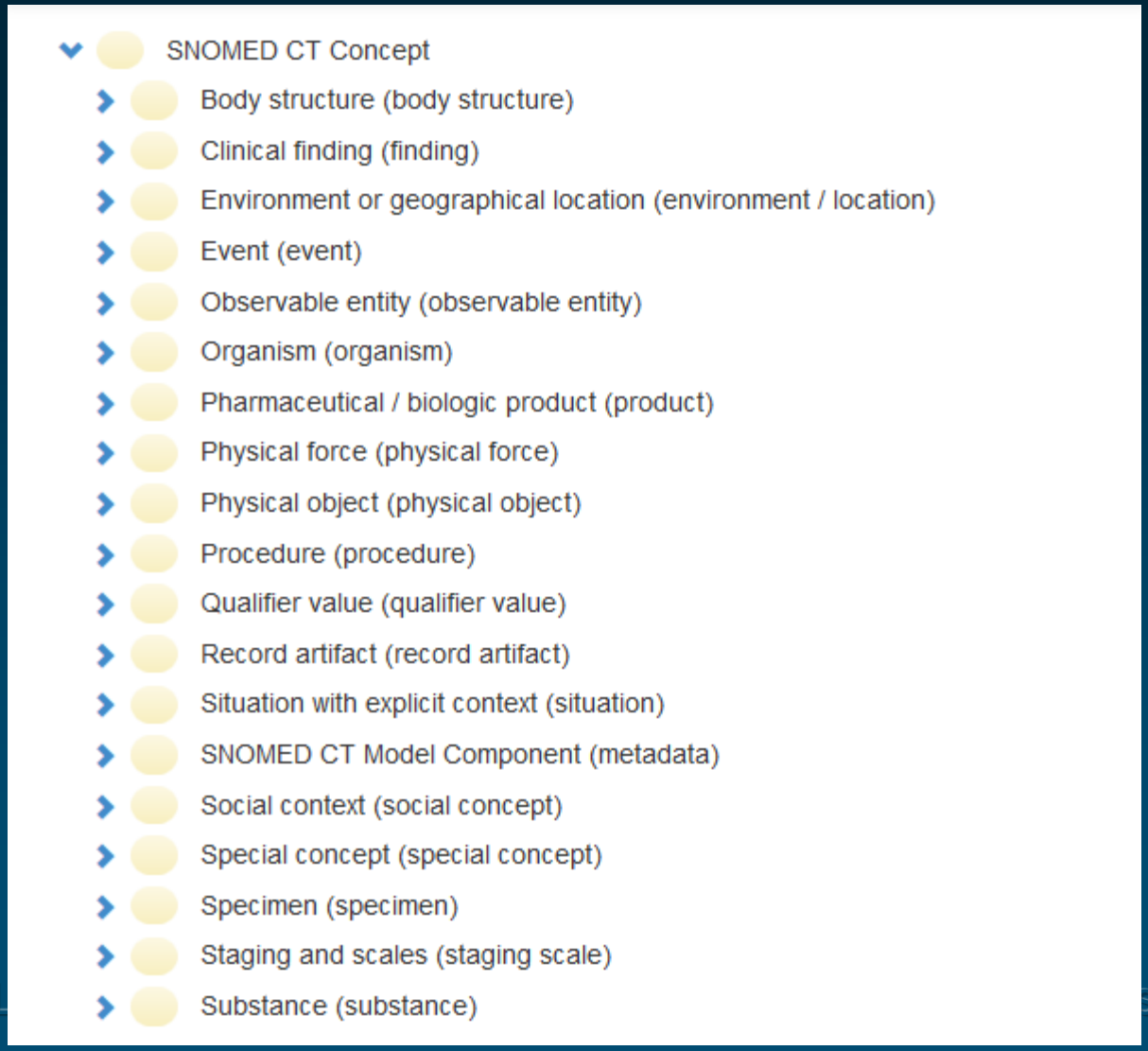
SNOMED CT Characteristics (1)

- Current version: January 31, 2016
  - 2 annual releases
- Type: Reference terminology / ontology
- Domain: Clinical medicine
- Developer: IHTSDO
- Funding: IHTSDO member countries
- Publicly available: Yes*
- Used for: clinical documentation, information exchange, analytics
- URL: http://www.ihtsdo.org/

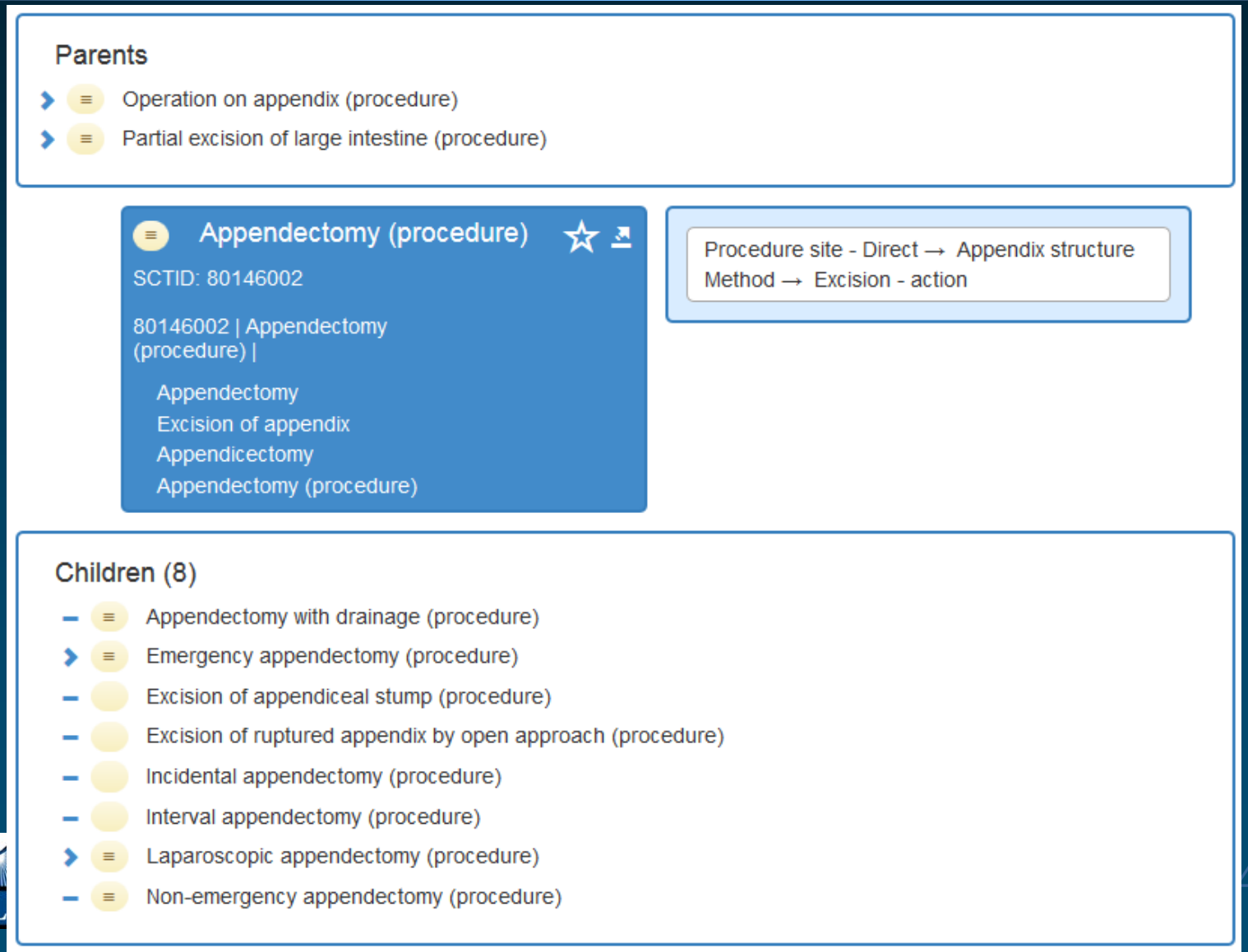
SNOMED CT Characteristics (2)

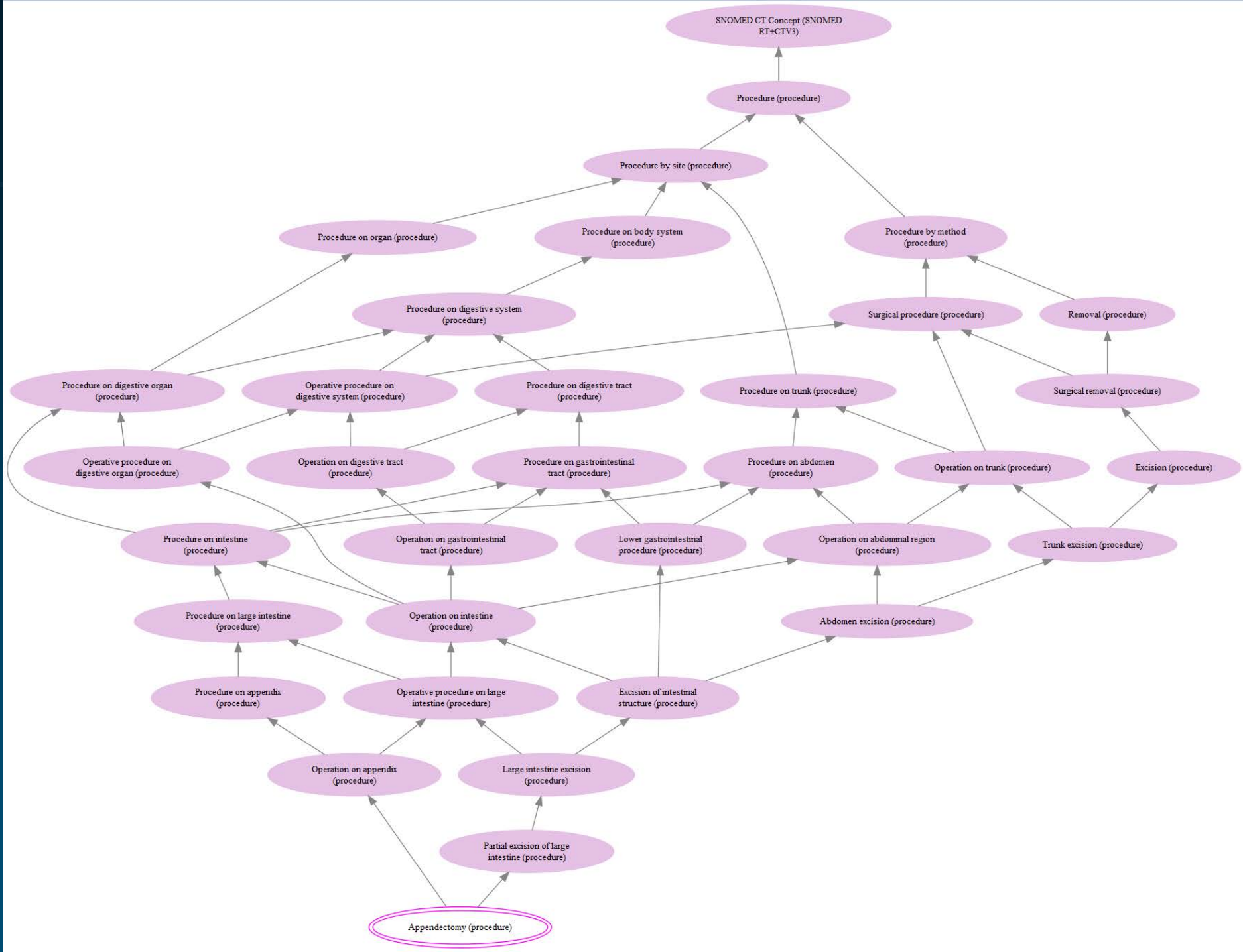
- Number of
  - Concepts: 320,912 active concepts (Sept. 2016)
  - Terms: 2.6 per concept (“descriptions”)
- Major organizing principles:
  - Polyhierarchy
  - Rich set of associative relationships
  - Logical definitions (incomplete: many primitives)
  - Built using description logics (EL++)
- Distribution: RF2 (proprietary)

SNOMED CT Top level

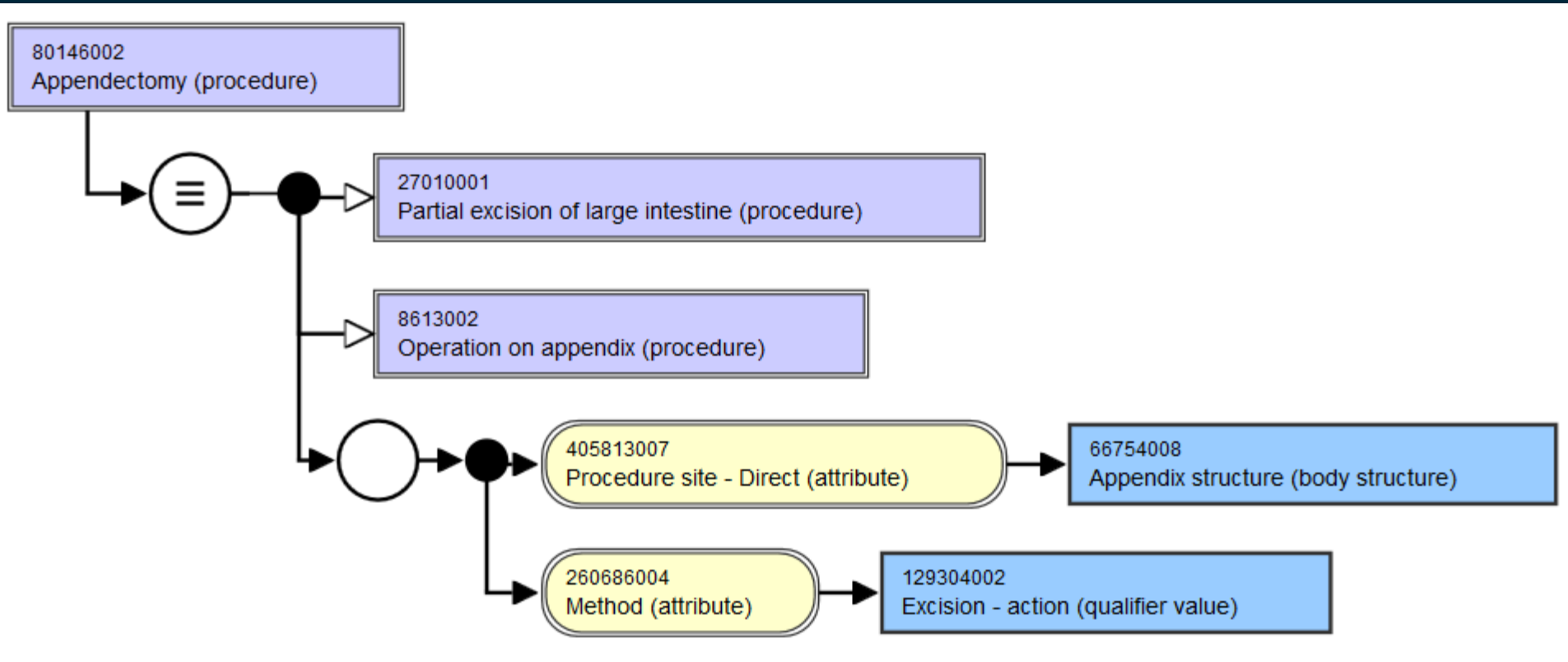


SNOMED CT Example




SNOMED CT Concept (SNOMED RT+CTV3)
Procedure (procedure)
Procedure by site (procedure)
Procedure on organ (procedure)
Procedure on body system (procedure)
Procedure by method (procedure)
Procedure on digestive system (procedure)
Surgical procedure (procedure)
Removal (procedure)
Procedure on digestive organ (procedure)
Operative procedure on digestive system (procedure)
Procedure on digestive tract (procedure)
Procedure on trunk (procedure)
Surgical removal (procedure)
Operative procedure on digestive organ (procedure)
Operation on digestive tract (procedure)
Procedure on gastrointestinal tract (procedure)
Procedure on abdomen (procedure)
Operation on trunk (procedure)
Excision (procedure)
Procedure on intestine (procedure)
Operation on gastrointestinal tract (procedure)
Lower gastrointestinal procedure (procedure)
Operation on abdominal region (procedure)
Trunk excision (procedure)
Procedure on large intestine (procedure)
Operation on intestine (procedure)
Abdomen excision (procedure)
Procedure on appendix (procedure)
Operative procedure on large intestine (procedure)
Excision of intestinal structure (procedure)
Operation on appendix (procedure)
Large intestine excision (procedure)
Partial excision of large intestine (procedure)
Appendectomy (procedure)

SNOMED CT Example



RxNorm

RxNorm Characteristics (1)

- Current version: April 2016
  - Monthly releases (+weekly updates)
- Type: Controlled terminology
- Domain: Drug names
- Developer: NLM
- Funding: NLM
- Publicly available: Yes*
- Used for: e-prescribing, information exchange, analytics
- URL: http://www.nlm.nih.gov/research/umls/rxnorm/

RxNorm Characteristics (2)

- Number of
  - Concepts: 117,774 (March 2016)
  - Terms: 1.5 per concept
- Major organizing principles:
  - Generic vs. brand
  - Ingredient + Strength + Dose form
  - No hierarchical structure; rich graph of associative relations
  - Integrates all major US drug information sources
  - No clinical information
- Distribution: similar to UMLS RRF format

RxNorm Normalized form

Strength **Ingredient** **Dose form**

4mg/ml Fluoxetine Oral Solution

Strength **Ingredient**

Semantic clinical drug component

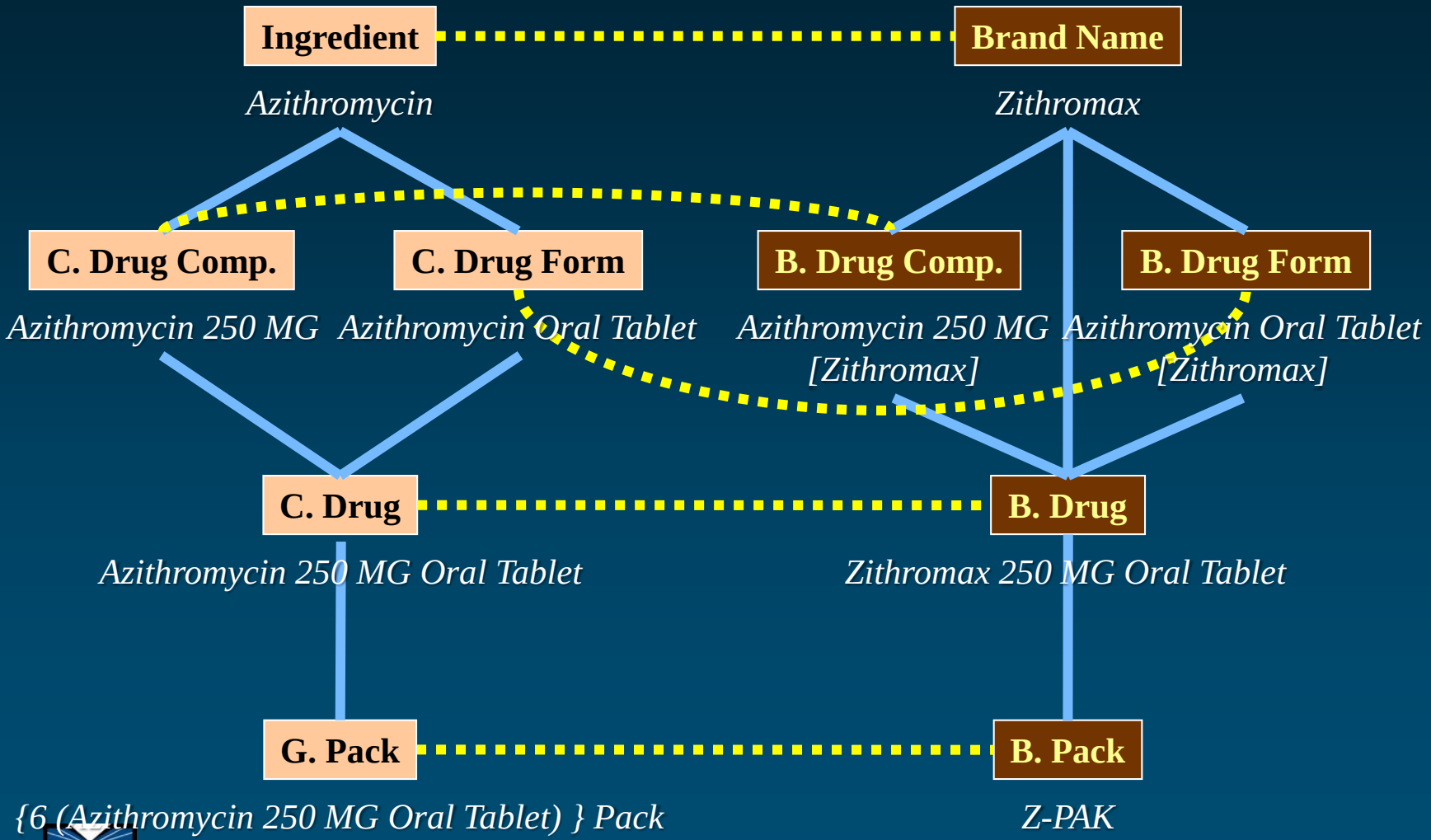
Ingredient **Dose form**

Semantic clinical drug form

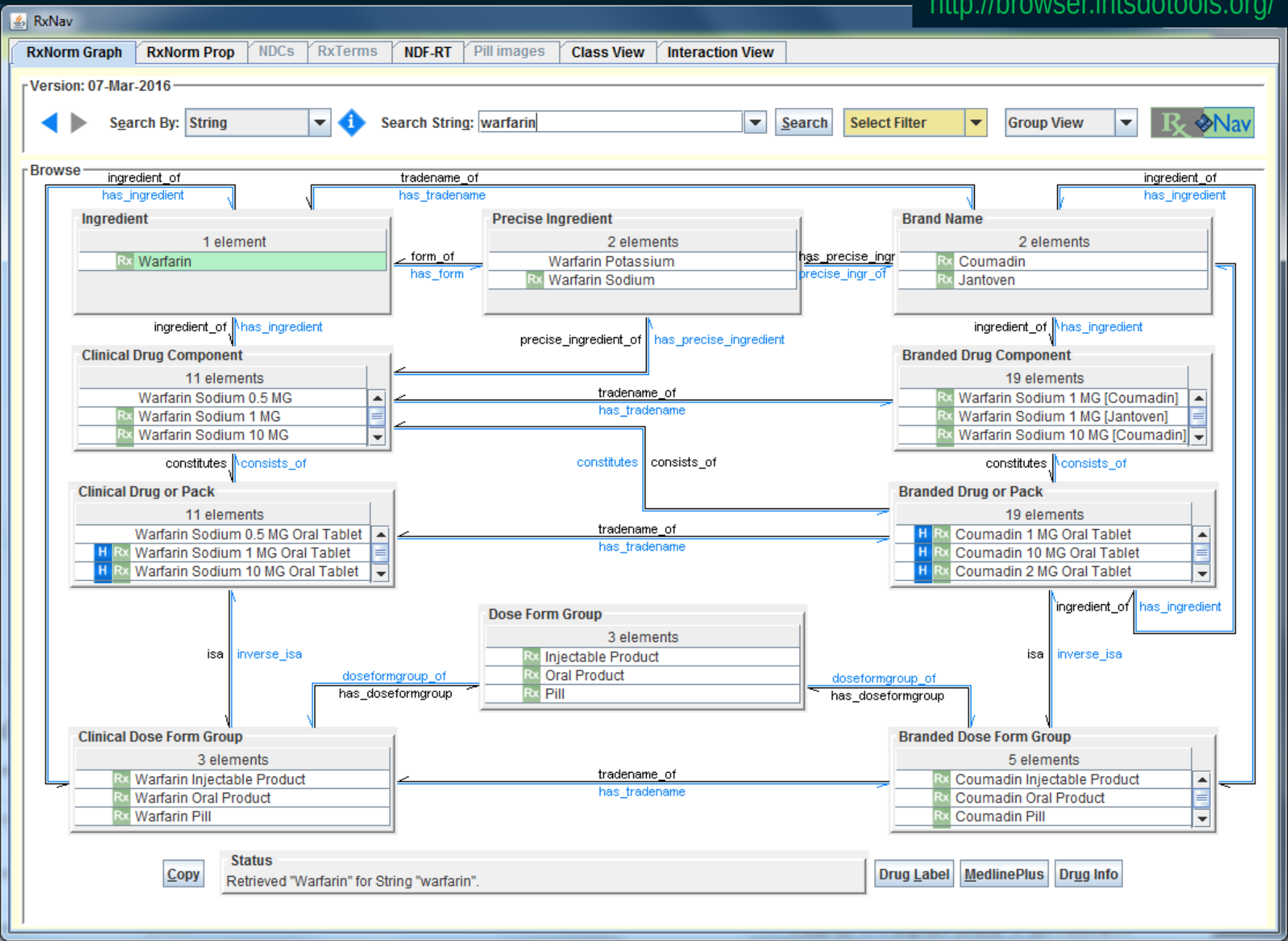
Strength **Ingredient** **Dose form**

Semantic clinical drug

RxNorm Example



http://browser.ihtsdotools.org/



Controlled Vocabularies
Part 4

Biomedical Ontologies “in Action”

A Functional Perspective

Overview

- Functional perspective [Bodenreider, YBMI 2008]
  - What are they for (vs. what are they)?
- "High-impact" biomedical ontologies
- 3 major categories of use
  - Knowledge management (indexing and retrieval of data and information, access to information, mapping among ontologies)
  - Data integration, exchange and semantic interoperability
  - Decision support and analytics (data selection and aggregation, decision support, natural language processing applications, knowledge discovery)

Knowledge management

Knowledge management

Annotating data and resources

Terminology in ontology

- Ontology as a source of vocabulary
  - List of names for the entities in the ontology (ontology vs. terminology)
- Most ontologies have some sort of terminological component
- Not all surface forms represented
  - Often insufficient for NLP applications
  - Large variation in number of terms per concept across ontologies

Annotating data

- Gene Ontology
  - Functional annotation of gene products in several dozen model organisms
- Various communities use the same controlled vocabularies
- Enabling comparisons across model organisms
- Annotations
  - Assigned manually by curators
  - Inferred automatically (e.g., from sequence similarity)

GO Annotations across species

ALDH2 aldehyde dehydrogenase 2 family (mitochondrial) [*Homo sapiens* (human)]

Gene ID: 217, updated on 13-Mar-2016

⊟ Gene Ontology Provided by GOA

Function	Evidence Code	Pubs
aldehyde dehydrogenase (NAD) activity	EXP	
aldehyde dehydrogenase (NAD) activity	IDA	PubMed
aldehyde dehydrogenase [NAD(P)+] activity	TAS	PubMed
electron carrier activity	TAS	PubMed

Aldh2 aldehyde dehydrogenase 2, mitochondrial [*Mus musculus* (house mouse)]

Gene ID: 11669, updated on 26-Jan-2016

⊟ Gene Ontology Provided by MGI

Function	Evidence Code	Pubs
NADH binding	ISO	
aldehyde dehydrogenase (NAD) activity	IBA	
aldehyde dehydrogenase (NAD) activity	ISO	
identical protein binding	ISO	
oxidoreductase activity	IEA	
oxidoreductase activity, acting on the aldehyde or oxo group of donors, NAD or NADP as acceptor	IEA	
protein binding	IPI	PubMed

Indexing the biomedical literature

- MeSH
  - Used for indexing and retrieval of the biomedical literature (MEDLINE)
- Indexing
  - Performed manually by human indexers
    - With help of semi-automatic systems (suggestions) e.g., Indexing Initiative at NLM
  - Specific indexing rules

MeSH MEDLINE indexing

Anesth Analg. 2008 Jun;106(6):1813-9. doi: 10.1213/ane.0b013e318172fdba.

PubMed.gov
US National Library of Medicine
National Institutes of Health

Free cortisol in sepsis and septic shock.

Bendel S[1], Karlsson S, Pettilä V, Loisa P, Varpula M, Ruokonen E; Finnsepsis Study Group.

⊕ Collaborators (26)

⊕ Author information

Abstract

BACKGROUND: Severe sepsis activates the hypothalamopituitary axis, increasing cortisol production. In some studies, hydrocortisone substitution based on an adrenocorticotropic hormone-stimulation test or baseline cortisol measurement has improved outcome. Because only the free fraction of cortisol is active, measurement of free cortisol may be more important than total cortisol in critically ill patients. We measured total and free cortisol in patients with severe sepsis and related the concentrations to outcome.

METHODS: In a prospective study, severe sepsis was defined according the American College of Chest Physicians/Society of Critical Care Medicine criteria. Blood samples were drawn within 24 h of study entry. Serum cortisol was analyzed by electrochemiluminescence immunoassay. The Coolens method was used for calculating serum free cortisol concentrations.

RESULTS: Blood samples were collected from 125 patients, of whom 62 had severe sepsis and 63 septic shock. Hospital mortality was 21%. Calculated free serum cortisol correlated well with serum total cortisol ($r = 0.90$, $P < 0.001$). There was no difference in the total cortisol concentrations in patients with sepsis and septic shock (728 +/- 386 nmol/L vs 793 +/- 439 nmol/L, $P = 0.44$). Nonsurvivors had higher calculated serum free (209 +/- 151 nmol/L) and total (980 +/- 458 nmol/L) cortisol concentrations than survivors (119 +/- 111 nmol/L, $P = 0.002$, and 704 +/- 383 nmol/L, $P = 0.002$). Depending on the definition, the incidence of adrenal insufficiency varied from 8% to 54%.

CONCLUSIONS: Clinically, calculation of free cortisol does not provide essential information for identification of patients who would benefit from corticoid treatment in severe sepsis and septic shock.

PMID: 18499615 [PubMed - indexed for MEDLINE]

MeSH MEDLINE indexing

MeSH Terms

Adrenal Cortex Function Tests
Adrenal Insufficiency/blood*
Adrenal Insufficiency/drug therapy
Adrenal Insufficiency/mortality
Adult
Biomarkers/blood
Female
Finland/epidemiology
Hospital Mortality
Humans
Hydrocortisone/blood*
Hydrocortisone/therapeutic use
Kaplan-Meier Estimate
Male
Predictive Value of Tests
Prospective Studies
Sepsis/blood*
Sepsis/drug therapy
Sepsis/mortality
Severity of Illness Index
Shock, Septic/blood*
Shock, Septic/drug therapy
Shock, Septic/mortality
Treatment Outcome

SNOMED CT/ICD Coding clinical data

- SNOMED CT
  - Used for clinical documentation
  - E.g., problem lists
- ICD-10-CM
  - Used for coding clinical data for billing purposes
  - Other uses of ICD
    - Morbidity and mortality reporting worldwide
  - Specific coding rules

Knowledge management

Accessing biomedical information

Resources for biomedical search engines

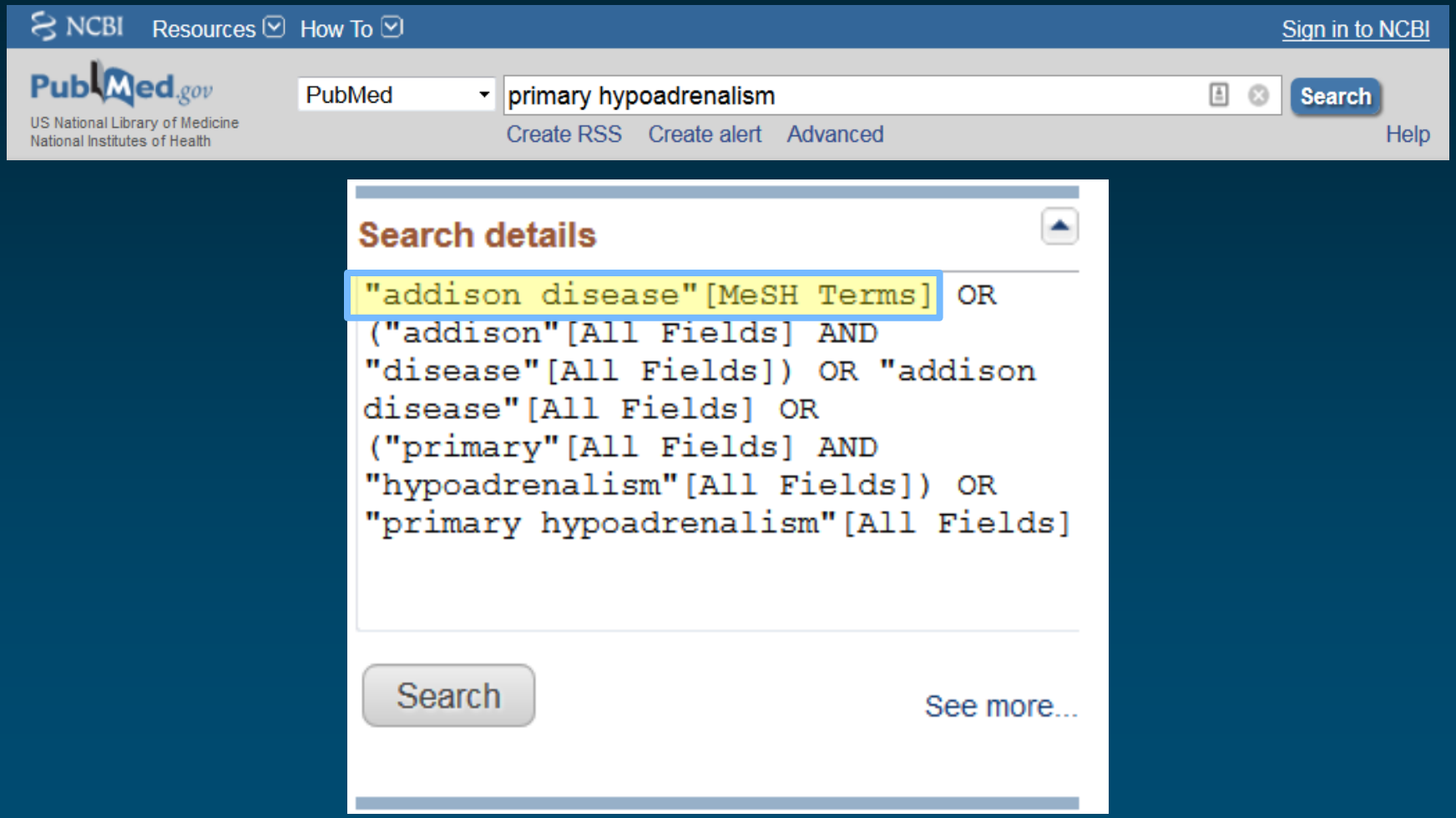
- Synonyms
- Hierarchical relations
- High-level categorization
- [Co-occurrence information]
- Translation

MeSH "synonyms" MEDLINE retrieval

- MeSH entry terms
  - Used as equivalent terms for retrieval purposes (query expansion)
  - Not always synonymous
- Increase recall without hurting precision

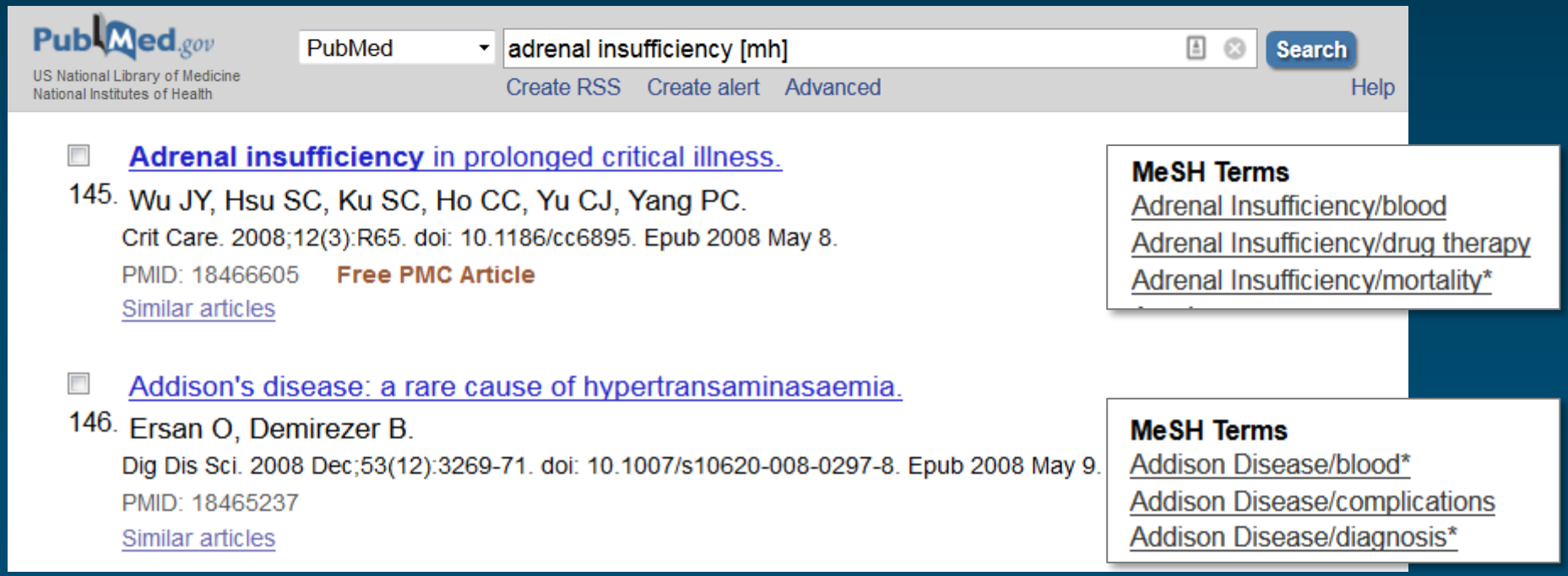
MeSH Heading	Addison Disease
Entry Term	Addison's Disease
Entry Term	Primary Adrenal Insufficiency
Entry Term	Primary Adrenocortical Insufficiency
Entry Term	Primary Hypoadrenalism

MeSH "synonyms" MEDLINE retrieval



MeSH hierarchies MEDLINE retrieval

- MeSH "explosion"
  - Search for a given MeSH term and all its descendants
  - A search on Adrenal insufficiency also retrieves articles indexed with its descendant, Addison disease



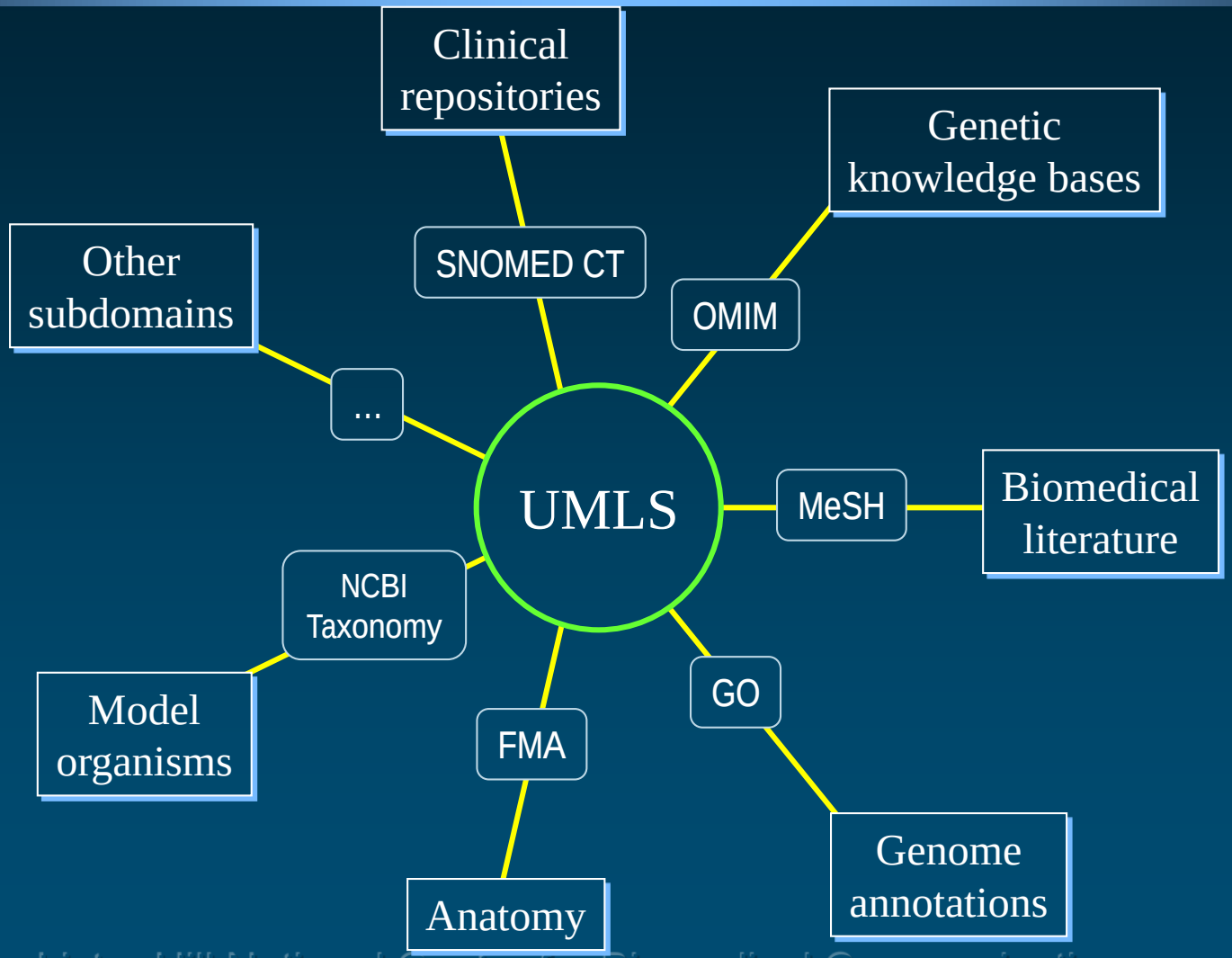
Knowledge management

Mapping across biomedical ontologies

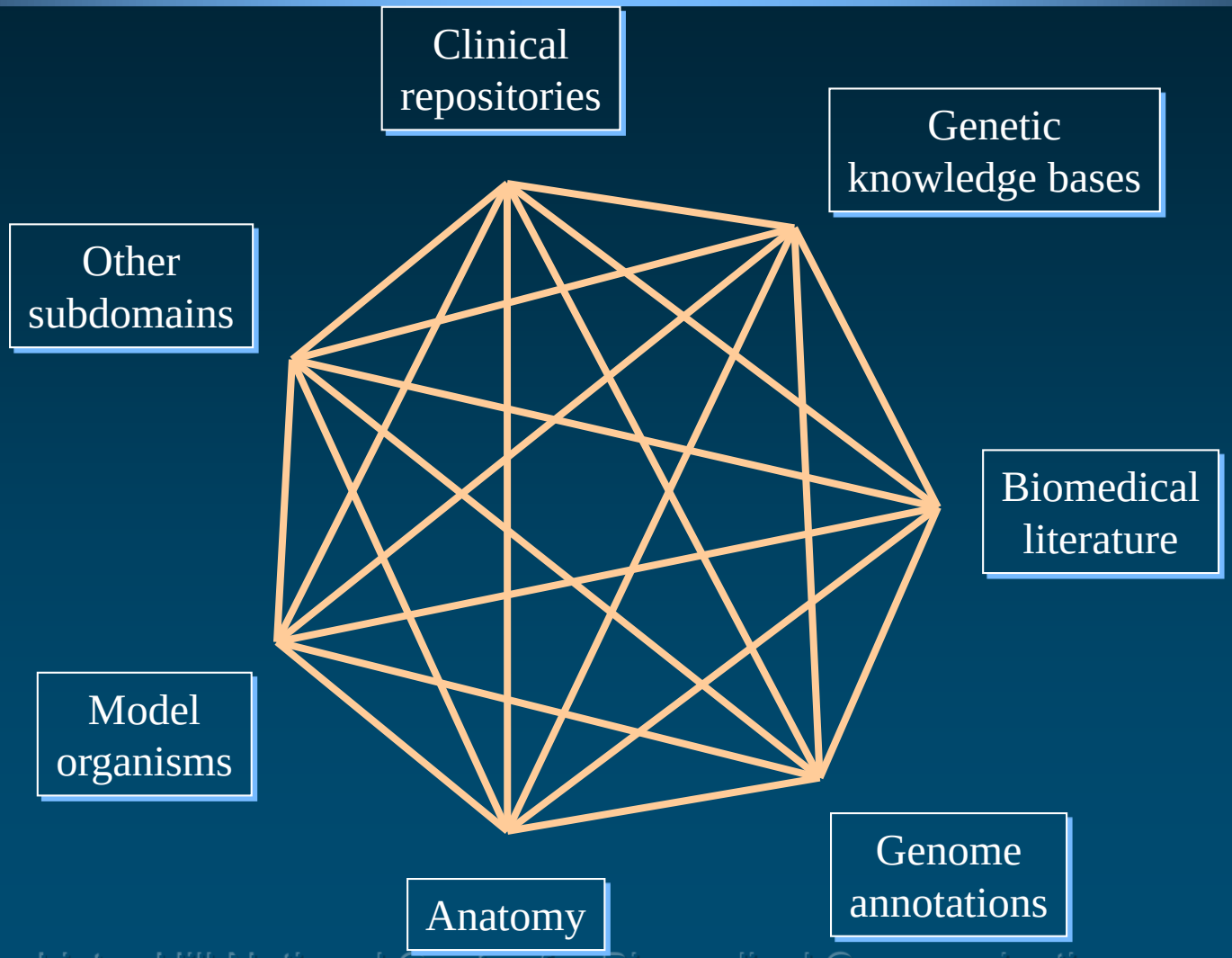
Terminology integration systems

- Terminology integration systems (UMLS, RxNorm) help bridge across vocabularies
- Uses
  - Information integration
  - Ontology alignment
  - Medication reconciliation

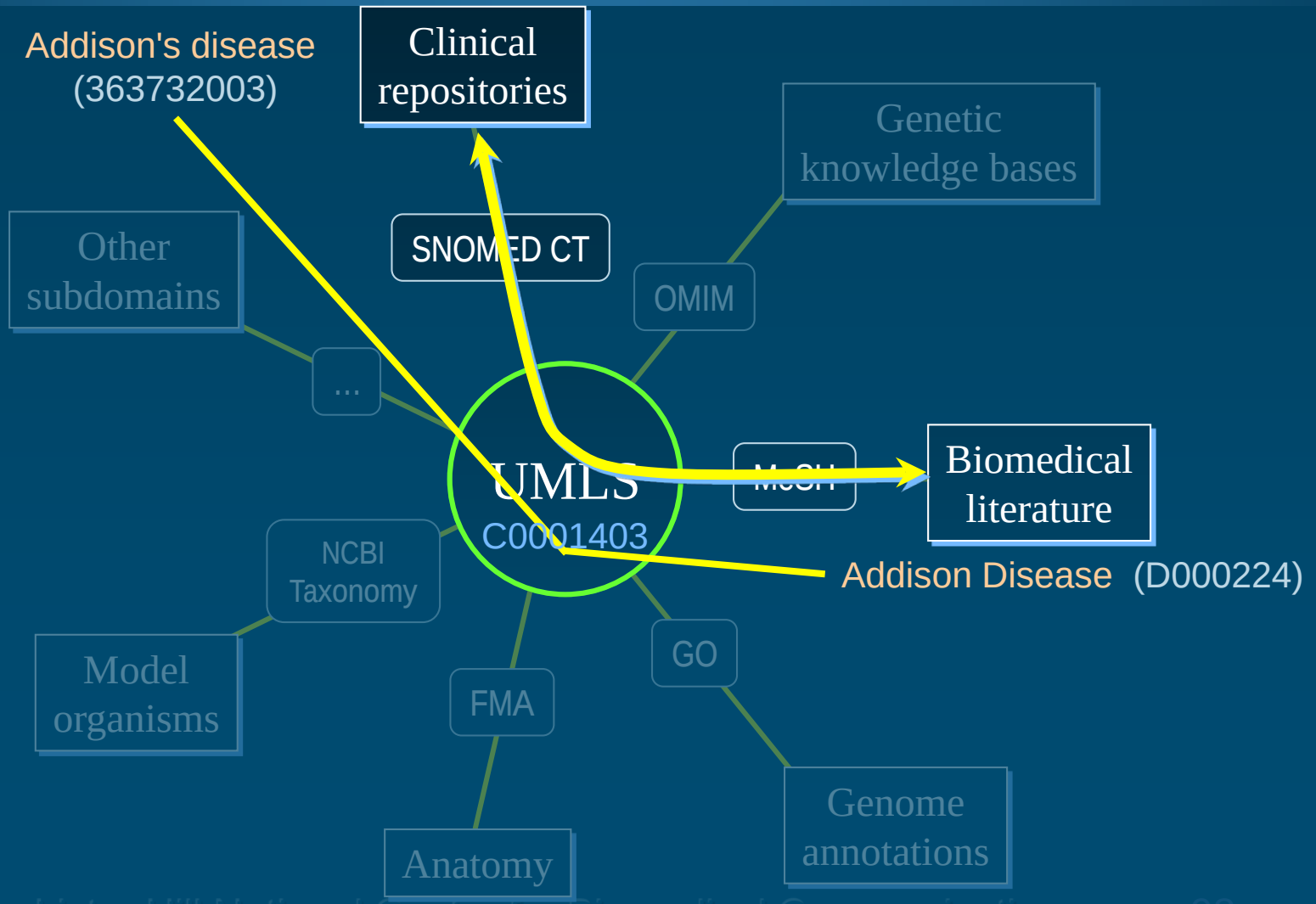
Integrating subdomains



Integrating subdomains



Trans-namespace integration



UMLS Source Vocabularies (2015AB)

- 153 families of source vocabularies
  - Not counting translations
- 25 languages
- Broad coverage of biomedicine
  - 9.8M names (normalized)
  - 3.2M concepts
  - ~13M relations among concepts
- Common presentation

Metathesaurus Basic organization

- Concepts
  - Synonymous terms are clustered into a concept
  - Properties are attached to concepts, e.g.,
    - Unique identifier
    - Definition
- Relations
  - Concepts are related to other concepts
  - Properties are attached to relations, e.g.,
    - Type of relationship
    - Source

Decision support and analytics

Data selection

- The structure of biomedical ontologies helps define groups of values from a high-level value
  - Vs. enumerating all possible values
- Useful for data selection in clinical studies
- ICD is used pervasively for this purpose
  - E.g., Study on supraventricular tachycardia (SVT), based on 2 high-level ICD codes
- Similarity with the definition of value sets for use in the information model

Data aggregation

- Ontologies help partition/aggregate data in data analysis
  - Clinical studies: Study a variable in groups of patients corresponding to the top level categories in ICD
  - Biology studies: Functional characterization of gene expression signatures with high-level concepts from the Gene Ontology
    - Recent trend: co-clustering

Decision support

- Clinical decision support
  - Ontologies help normalize the vocabulary and increase the recall of rules
  - Ontologies provide some domain knowledge and make it possible to create high-level rules (e.g., for a class of drugs rather than for each drug in the class)
- Other forms of decision support
  - Based on automatic reasoning services for OWL ontologies (e.g., grading gliomas with NCIt)

Natural language processing applications

- Ontologies provide background domain knowledge for NLP applications
  - Question answering
  - Document summarization
  - Literature-based discovery
- The UMLS is often used, but other specific resources have been developed

Knowledge discovery

- By standardizing the vocabulary in a given domain, ontologies are enabling resources for knowledge discovery through data mining
- Less frequently, the structure of the ontology is leveraged by data mining algorithms
- Example of available datasets
  - ICD-coded clinical data (in conjunction with non-clinical information, e.g., environmental data)
  - Annotation of gene products to the GO (function prediction)

Controlled Vocabularies

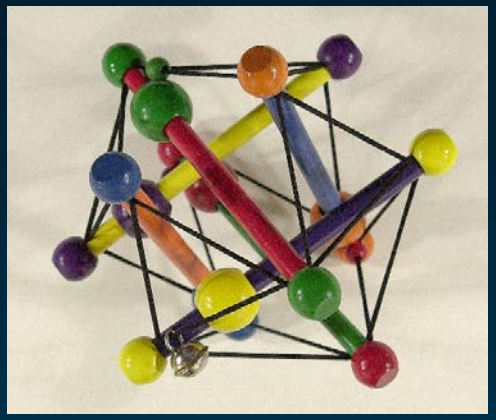
Summary

Summary

- History of biomedical ontologies
- How clinical features are reflected in disease names
- Structure of the main clinical ontologies used
  - ICD, SNOMED CT, LOINC, RxNorm
- Purpose of biomedical ontologies
  - Knowledge management, [health information exchange and semantic interoperability], and clinical decision support and analytics

Topics not discussed

- Semantic Web, URIs, Linked Data
- Ontology creation, Protege
- Accessing terminology resources (APIs)
- Ontology repositories
  - [UMLS], NCBO BioPortal, EBI Ontology Lookup Service
- NLP, named entity recognition, MetaMap
- Mapping local terms to standard terminologies
- VSAC, value sets, common data elements
- OBO ontologies, OBO Foundry
- Coordinated development of ontologies, harmonization
- Boundary between terminology and information model
- […]



Medical Ontology Research

Contact: olivier@nlm.nih.gov
Web: https://mor.nlm.nih.gov

Olivier Bodenreider

Lister Hill National Center
for Biomedical Communications
Bethesda, Maryland - USA

NIH U.S. National Library of Medicine