



NLM Georgia Biomedical Informatics Course

Fall 2017 Course

Controlled Vocabularies & Semantic Standards



Olivier Bodenreider

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



U.S. National Library of Medicine



Learning objectives

- ◆ 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
- ◆ List and describe the main biomedical semantic standards

References Review articles

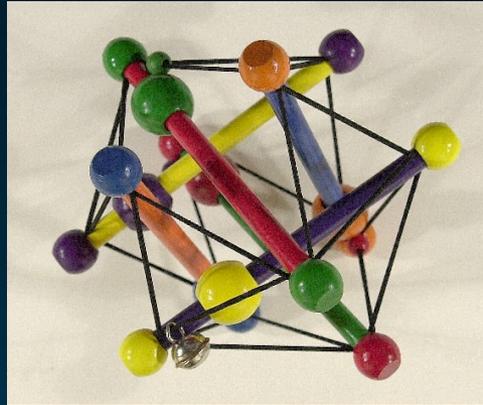
- ◆ Bodenreider O, Stevens R. *Bio-ontologies: current trends and future directions*. *Brief Bioinform*. 2006 Sep;7(3):256-74.
- ◆ Cimino JJ, Zhu X. *The practical impact of ontologies on biomedical informatics*. *Yearb Med Inform*. 2006:124-35.
- ◆ Bodenreider O. *Biomedical ontologies in action: role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



U.S. National Library of Medicine



Outline

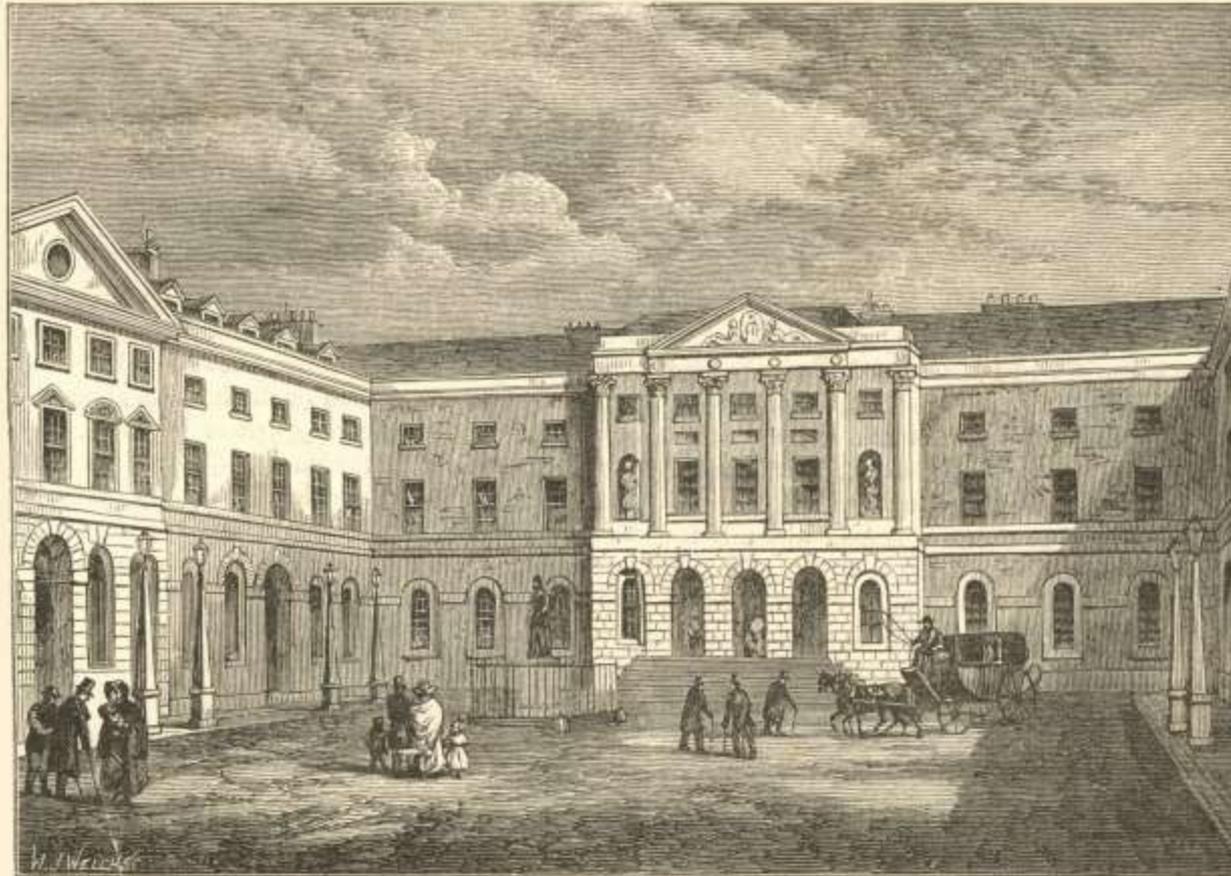
- ◆ Introduction to biomedical terminologies through an example
- ◆ “High-Impact” Biomedical Ontologies
 - Structural perspective
- ◆ Biomedical Ontologies “in Action”
 - Functional perspective



Controlled Vocabularies & Semantic Standards
Part 1

Introduction to biomedical
terminologies through an example

Guy's Hospital, London



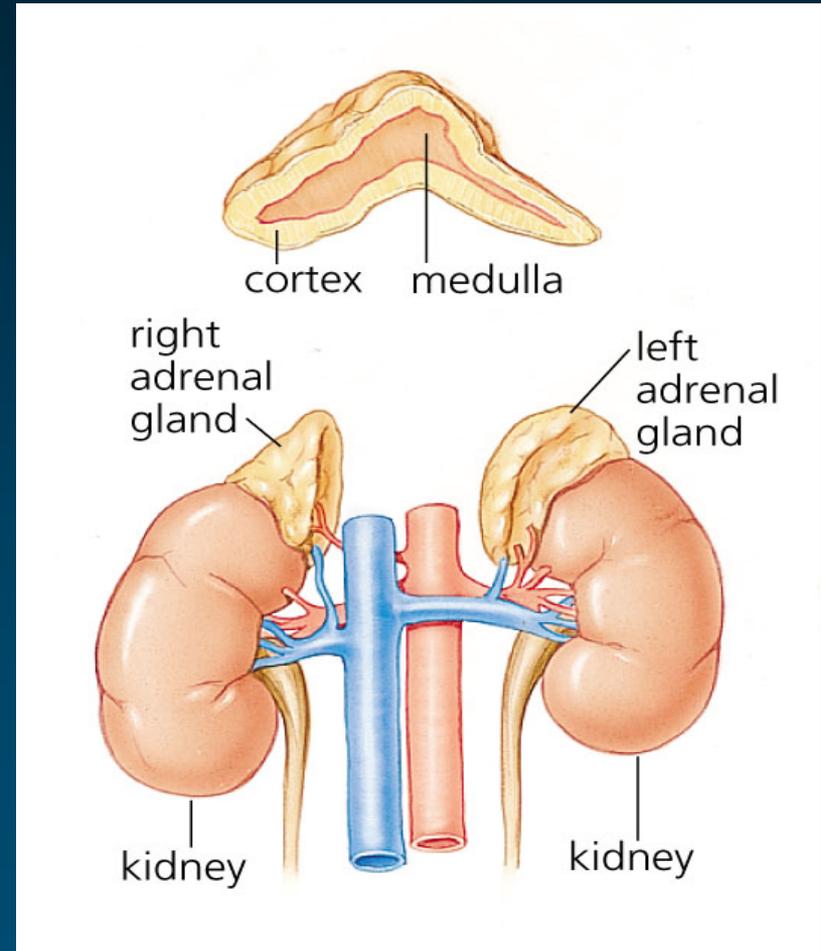
GUY'S HOSPITAL.

Thomas Addison (1795-1860)



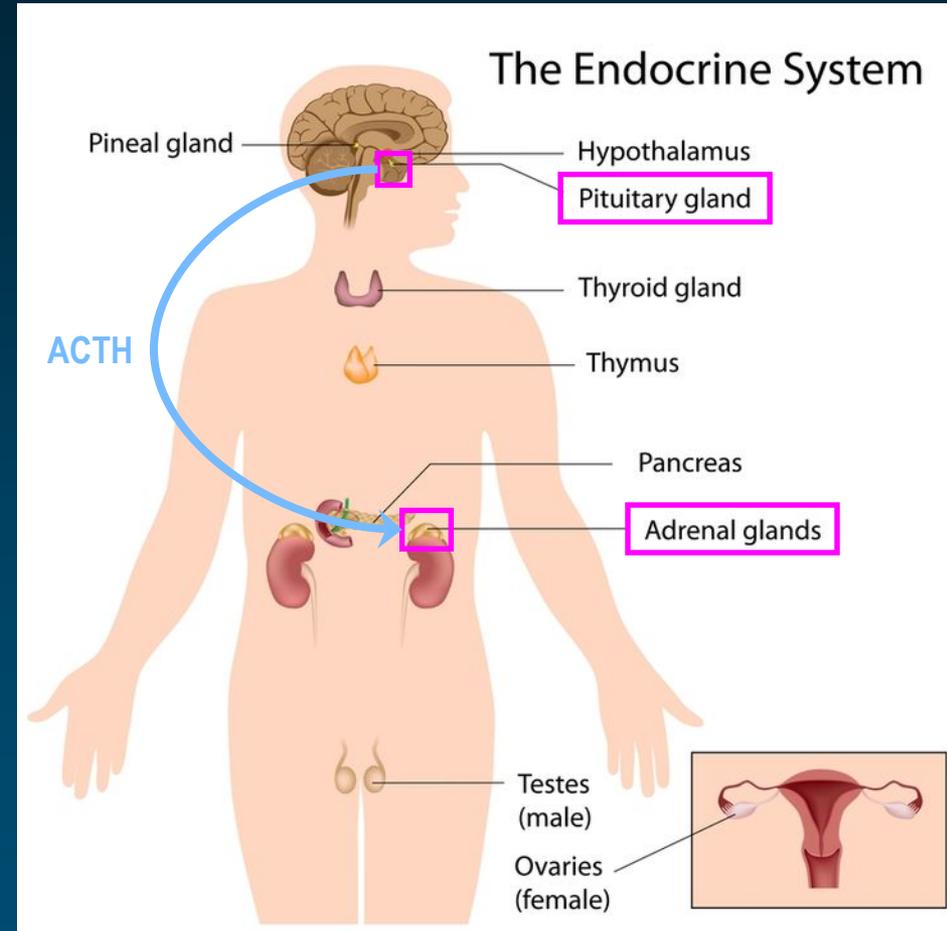
Addison's disease

- ◆ 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
 - Bronzed disease
 - Addison melanoderma
 - Asthenia pigmentosa
 - Primary adrenal deficiency
 - Primary adrenal insufficiency
 - Primary adrenocortical insufficiency
 - Chronic adrenocortical insufficiency
- } eponym
- } symptoms
- } 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
Hypoadosteronism

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 Re

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

Stated Inferred

● 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 & Semantic Standards
Part 2

**“High-Impact” Biomedical
Ontologies**

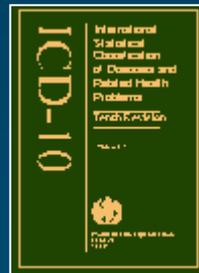
A Structural Perspective

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 (2017)
 - 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



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):	
<ul style="list-style-type: none">• brittle• juvenile-onset• ketosis-prone	
Excl.: diabetes mellitus (in):	
<ul style="list-style-type: none">• malnutrition-related (E12.-)• neonatal (P70.2)	
<ul style="list-style-type: none">• pregnancy, childbirth and the puerperium (O24.-)	
glycosuria:	
<ul style="list-style-type: none">• 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
	<i>Incl.:</i> 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*)
	<i>Excl.:</i> 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 (CDC/NCHS)
- ◆ 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 (E71.0-E71.2)
 • 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)



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 with cystinosis

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

x6

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

W58.02 Struck by alligator

W58.03 Crushed by alligator

W58.09 Other contact with alligator

W58.01A Bitten by alligator, initial encounter

W58.01D Bitten by alligator, subsequent encounter

W58.01S Bitten by alligator, sequela

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: 2.61 (June 2017)
 - 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]	

SNOMED Clinical Terms



SNOMED CT Characteristics (1)

- ◆ Current version: July 31, 2017
 - 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 Concept
 - ▶ ● Body structure (body structure)
 - ▶ ● Clinical finding (finding)
 - ▶ ● Environment or geographical location (environment / location)
 - ▶ ● Event (event)
 - ▶ ● Observable entity (observable entity)
 - ▶ ● Organism (organism)
 - ▶ ● Pharmaceutical / biologic product (product)
 - ▶ ● Physical force (physical force)
 - ▶ ● Physical object (physical object)
 - ▶ ● Procedure (procedure)
 - ▶ ● Qualifier value (qualifier value)
 - ▶ ● Record artifact (record artifact)
 - ▶ ● Situation with explicit context (situation)
 - ▶ ● SNOMED CT Model Component (metadata)
 - ▶ ● Social context (social concept)
 - ▶ ● Special concept (special concept)
 - ▶ ● Specimen (specimen)
 - ▶ ● Staging and scales (staging scale)
 - ▶ ● Substance (substance)

SNOMED CT Example

Parents

- ▶ ☰ Operation on appendix (procedure)
- ▶ ☰ Partial excision of large intestine (procedure)

☰ Appendectomy (procedure) ☆ ↗

SCTID: 80146002

80146002 | Appendectomy (procedure) |

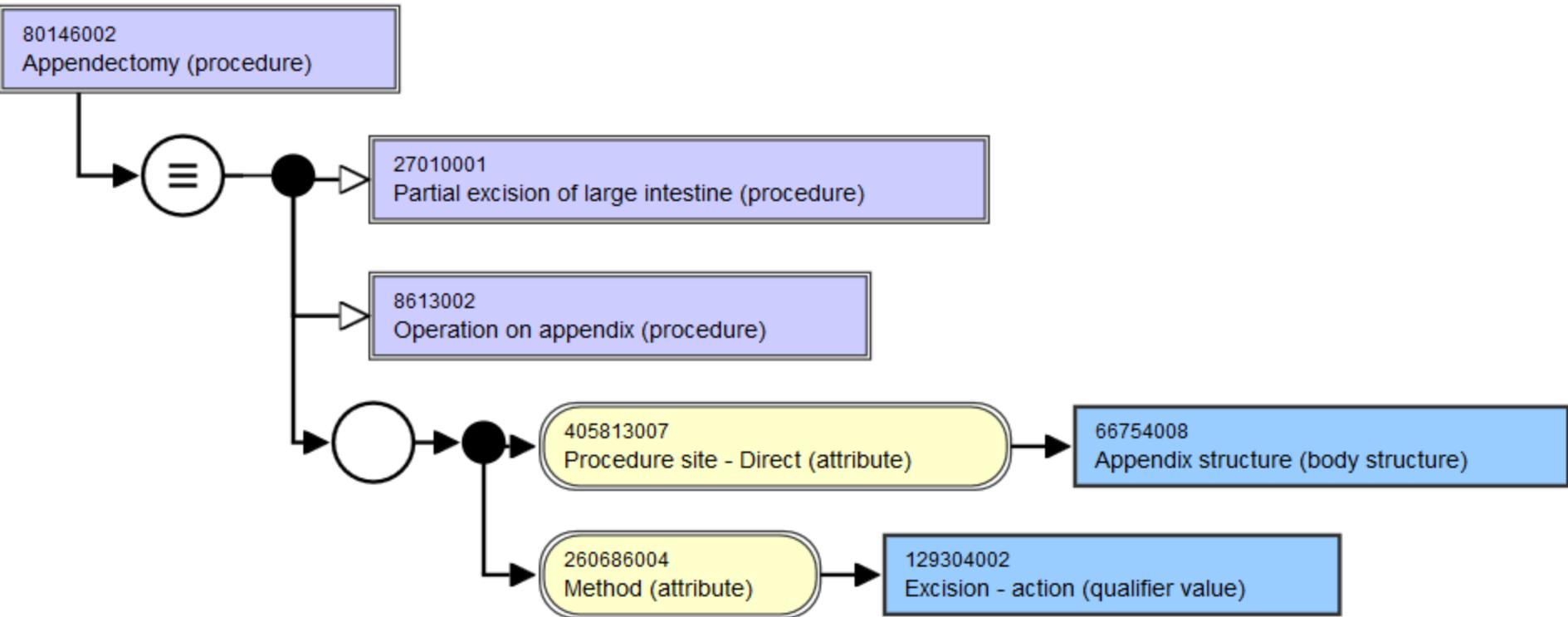
Appendectomy
Excision of appendix
Appendicectomy
Appendectomy (procedure)

Procedure site - Direct → Appendix structure
Method → Excision - action

Children (8)

- ☰ Appendectomy with drainage (procedure)
- ▶ ☰ Emergency appendectomy (procedure)
- ● Excision of appendiceal stump (procedure)
- ● Excision of ruptured appendix by open approach (procedure)
- ● Incidental appendectomy (procedure)
- ● Interval appendectomy (procedure)
- ▶ ☰ Laparoscopic appendectomy (procedure)
- ☰ Non-emergency appendectomy (procedure)

SNOMED CT Example



RxNorm

RxNorm Characteristics (1)

- ◆ Current version: August 2017
 - 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

4mg/ml

Ingredient

Fluoxetine

Dose form

Oral Solution

Strength

Semantic clinical drug component

Ingredient

Ingredient

Semantic clinical drug form

Dose form

Strength

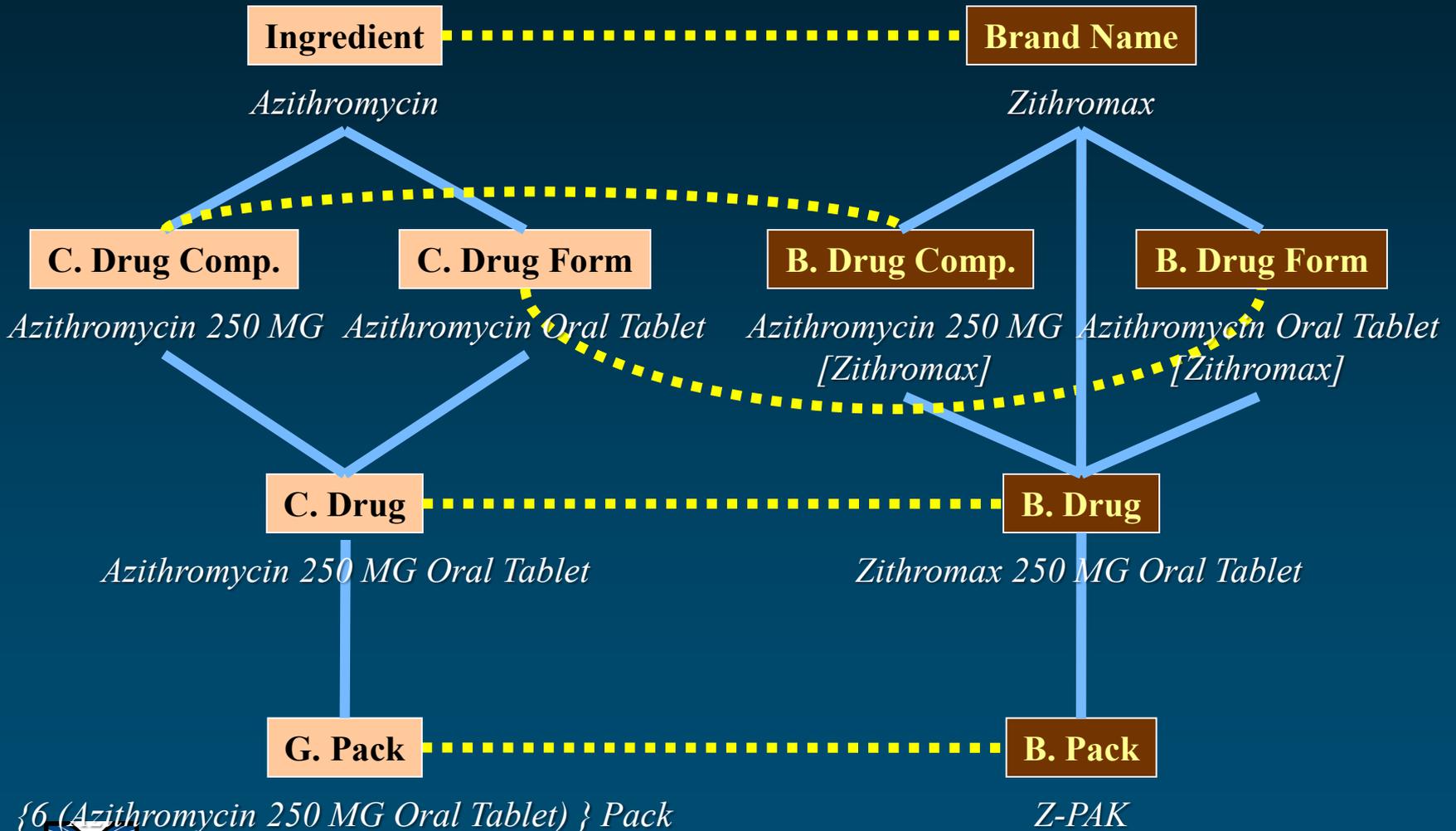
Semantic clinical drug

Ingredient

Dose form



RxNorm Example



{6 (Azithromycin 250 MG Oral Tablet) } Pack

Z-PAK





String [dropdown] warfarin [input] [search icon] [refresh icon]



Warfarin [RxCUI = 11289]

- RxNorm Graph
- RxNorm Properties
- NDC
- RxTerms
- NDF-RT
- Pill Images
- Class View
- Interaction View

- Views
- Classic
 - Simple
 - Table

- Filters
- H
 - V
 - Rx
 - S
 - Group
 - Form

- Links
- MIN
 - Pack
 - Multi
- Download

IN/MIN	Ingredient (1)
H Rx S	Warfarin

PIN	Precise Ingredient (2)
S	Warfarin Potassium
H Rx S	Warfarin Sodium

BN	Brand Name (2)
H Rx S	Coumadin
H Rx S	Jantoven

SCDC	Clinical Drug Component (11)
S	Warfarin Sodium 0.5 MG
H Rx S	Warfarin Sodium 1 MG
H Rx S	Warfarin Sodium 10 MG
H Rx S	Warfarin Sodium 2 MG



SBDC	Branded Drug Component (18)
H Rx S	Warfarin Sodium 1 MG [Coumadin]
H Rx S	Warfarin Sodium 1 MG [Jantoven]
H Rx S	Warfarin Sodium 10 MG [Coumadin]
H Rx S	Warfarin Sodium 10 MG [Jantoven]

SCD/GPCK	Clinical Drug or Pack (11)
S	Warfarin Sodium 0.5 MG Oral Tablet
H Rx S	Warfarin Sodium 1 MG Oral Tablet
H Rx S	Warfarin Sodium 10 MG Oral Tablet
H Rx S	Warfarin Sodium 2 MG Oral Tablet

SBD/BPCK	Branded Drug or Pack (18)
H Rx S	Coumadin 1 MG Oral Tablet
H Rx S	Coumadin 10 MG Oral Tablet
H Rx S	Coumadin 2 MG Oral Tablet
H Rx S	Coumadin 2.5 MG Oral Tablet

SCDG	Clinical Dose Form Group (3)
S	Warfarin Injectable Product
H Rx S	Warfarin Oral Product
H Rx S	Warfarin Pill

DFG	Dose Form Group (3)
HVRx S	Injectable Product
HVRx S	Oral Product
HVRx S	Pill

SBDG	Branded Dose Form Group (4)
H Rx S	Coumadin Oral Product
H Rx S	Coumadin Pill
H Rx S	Jantoven Oral Product
H Rx S	Jantoven Pill



Controlled Vocabularies & Semantic Standards
Part 3

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**
 - Annotating data and resources
 - Accessing biomedical information
 - Mapping across biomedical ontologies
 - **Decision support and analytics**
 - Value sets and clinical quality measures
 - **Data integration, exchange and semantic interoperability**
 - Common data models
 - Fast Healthcare Interoperability Resources (FHIR)



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

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.0b013e318172fdb.

Free cortisol in sepsis and septic shock.

Bendel S¹, 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 adrenocorticotrophic 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]

PubMed.gov

US National Library of Medicine
National Institutes of Health

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

NCBI Resources How To Sign in to NCBI

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

PubMed primary hypoadrenalism Search

Create RSS Create alert Advanced Help

Search details

```
"addison disease"[MeSH Terms] OR  
("addison"[All Fields] AND  
"disease"[All Fields]) OR "addison  
disease"[All Fields] OR  
("primary"[All Fields] AND  
"hypoadrenalism"[All Fields]) OR  
"primary hypoadrenalism"[All Fields]
```

Search See more...



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

The screenshot shows a PubMed search interface. The search bar contains 'adrenal insufficiency [mh]' and the search button is labeled 'Search'. Below the search bar are links for 'Create RSS', 'Create alert', and 'Advanced'. The search results are listed below, with two results highlighted. To the right of each result is a box containing MeSH terms.

PubMed [Help](#)

[Create RSS](#) [Create alert](#) [Advanced](#)

145. [Adrenal insufficiency in prolonged critical illness.](#)
Wu JY, Hsu SC, Ku SC, Ho CC, Yu CJ, Yang PC.
Crit Care. 2008;12(3):R65. doi: 10.1186/cc6895. Epub 2008 May 8.
PMID: 18466605 **Free PMC Article**
[Similar articles](#)

MeSH Terms
[Adrenal Insufficiency/blood](#)
[Adrenal Insufficiency/drug therapy](#)
[Adrenal Insufficiency/mortality*](#)

146. [Addison's disease: a rare cause of hypertransaminasaemia.](#)
Ersan O, Demirezer B.
Dig Dis Sci. 2008 Dec;53(12):3269-71. doi: 10.1007/s10620-008-0297-8. Epub 2008 May 9.
PMID: 18465237
[Similar articles](#)

MeSH Terms
[Addison Disease/blood*](#)
[Addison Disease/complications](#)
[Addison Disease/diagnosis*](#)

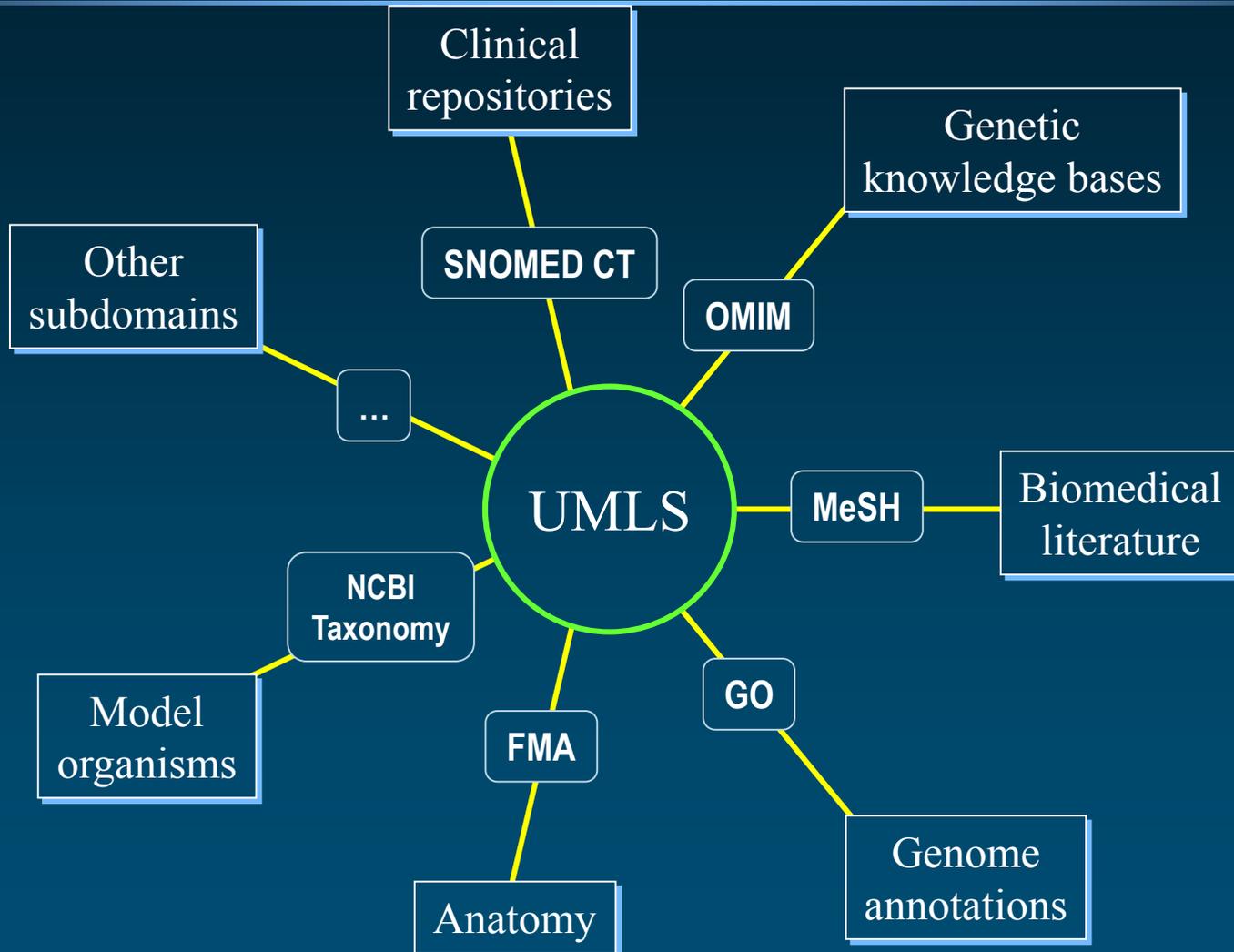
Knowledge management

Mapping across biomedical ontologies

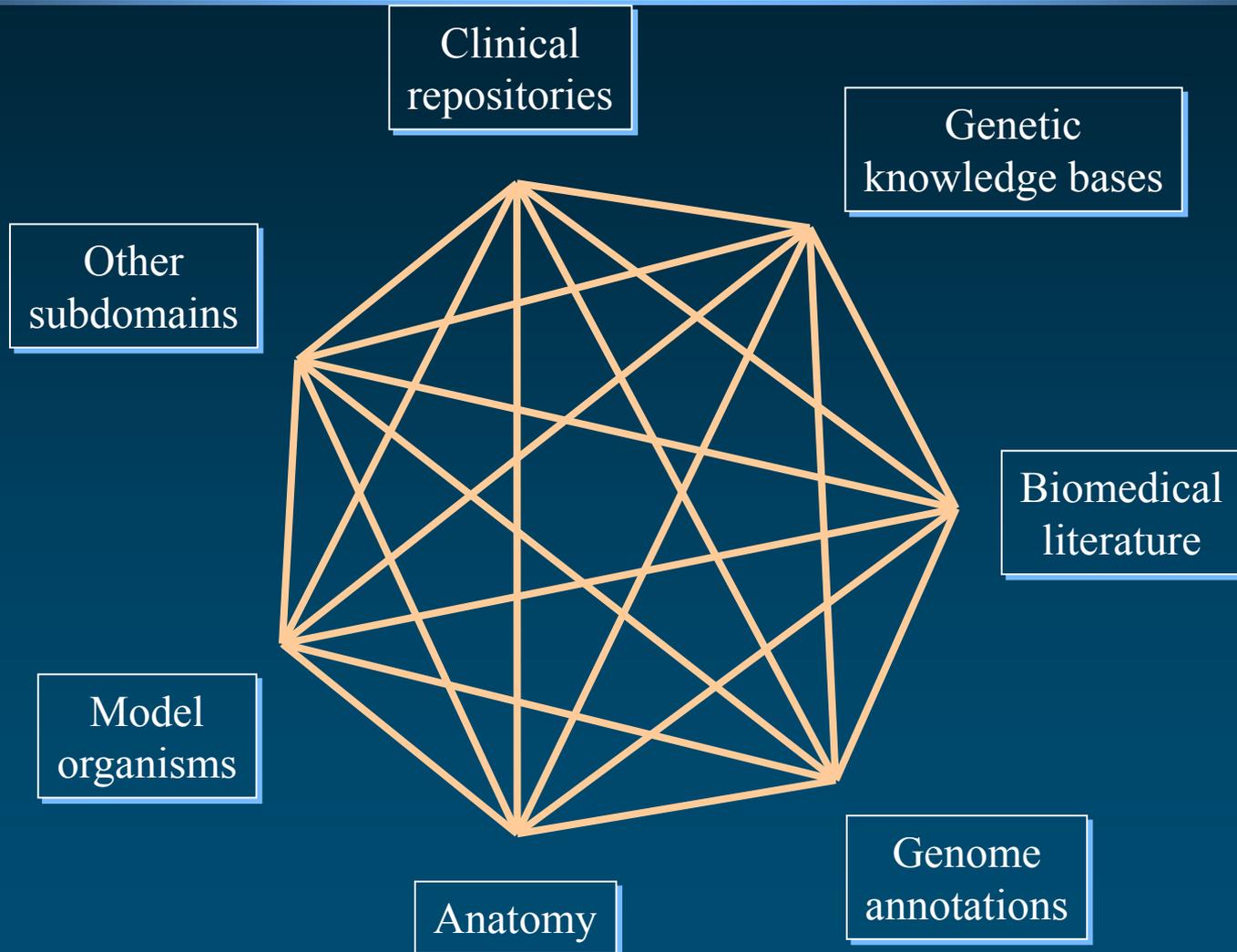
Terminology integration systems

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

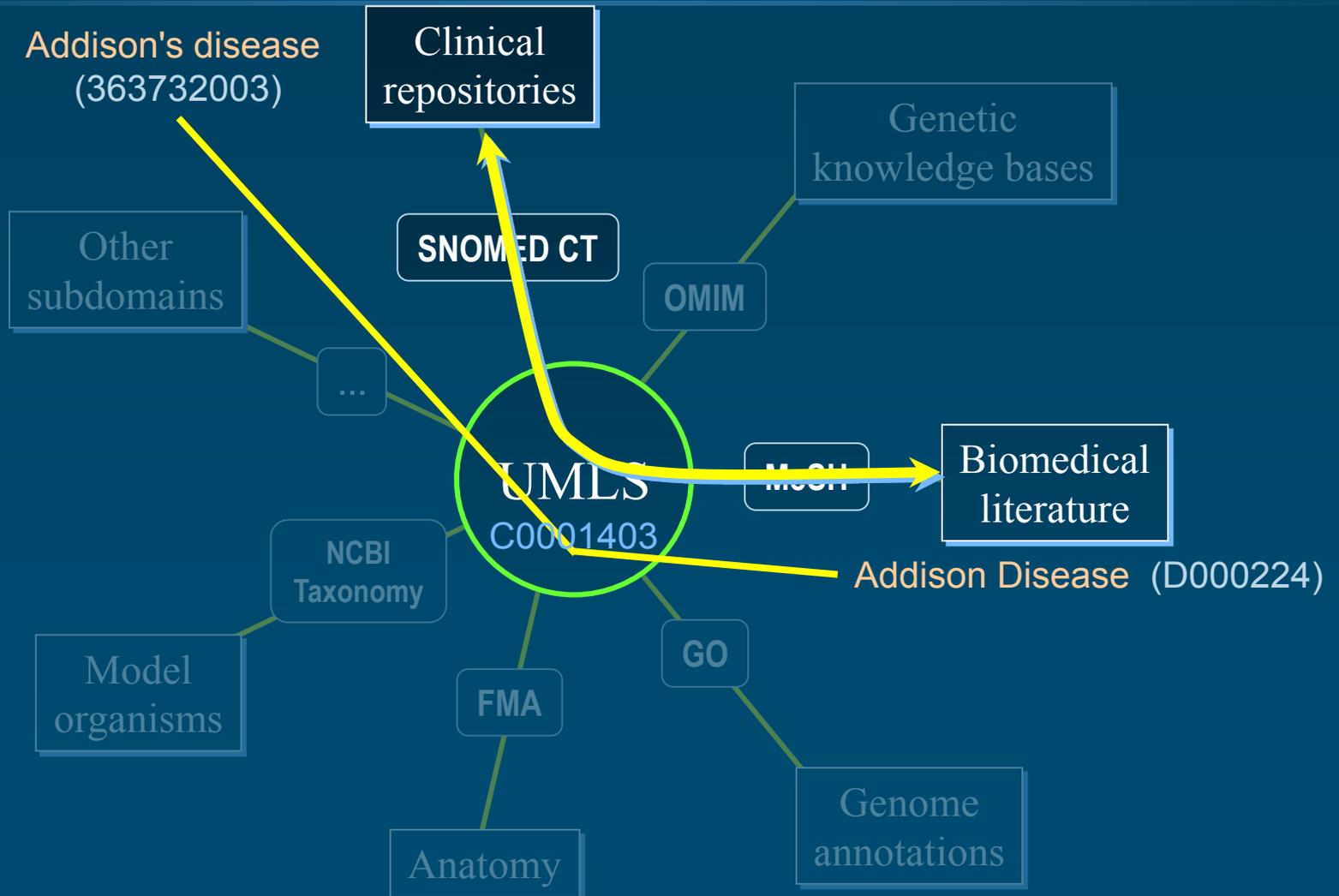
Integrating subdomains



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

Value sets and clinical quality measures

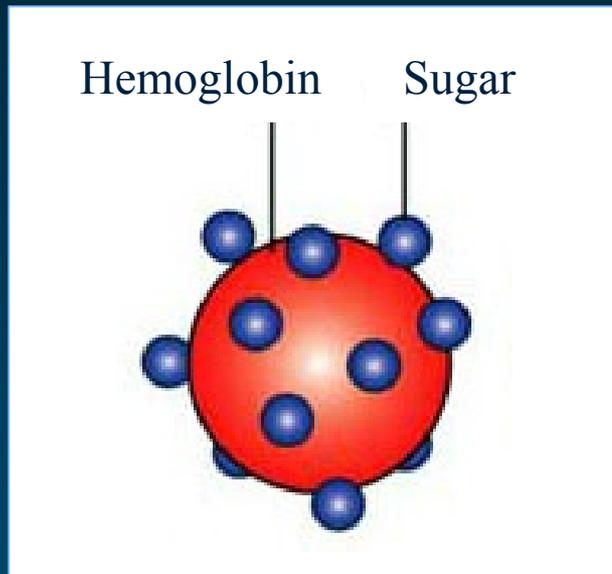
Clinical quality measures (CQMs)

- ◆ Measure and track the quality of healthcare services provided by eligible professionals, eligible hospitals and critical access hospitals within our health care system
- ◆ Measure many aspects of patient care including
 - Health outcomes
 - Clinical processes
 - Patient safety
 - Efficient use of healthcare resources
 - Care coordination
 - Patient engagement
 - Population and public health
 - Clinical guidelines

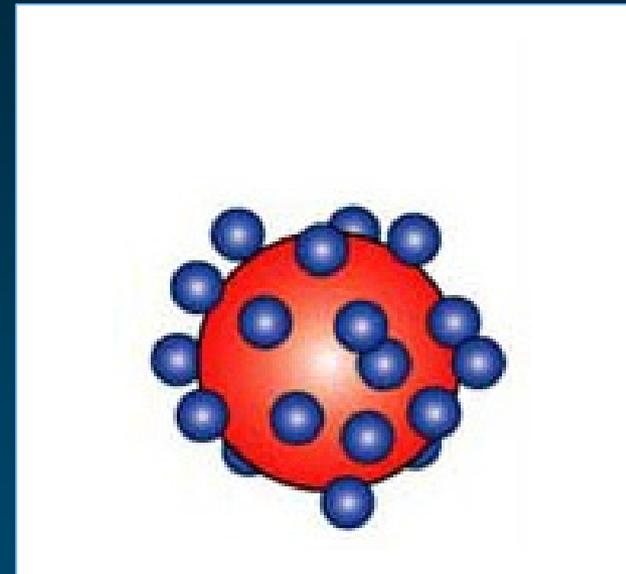


Clinical quality measures (example)

Hemoglobin A1c Test for Pediatric Patients



Normal glucose levels in blood
Low HbA1c concentration



High glucose levels in blood
High HbA1c concentration

Clinical recommendations

- 1. American Association of Clinical Endocrinologists (2002):** Recommends that *a glycosylated hemoglobin be performed during an initial assessment and during follow-up assessments*, which should occur at no longer than three-month intervals.
- 2. American Diabetes Association (2006):** Recommends *obtaining a glycosylated hemoglobin during an initial assessment and then routinely as part of continuing care*. In the absence of well-controlled studies that suggest a definite testing protocol, expert opinion recommends glycosylated hemoglobin be obtained at least twice a year in patients who are meeting treatment goals and who have stable glycemic control and more frequently (quarterly assessment) in patients whose therapy was changed or who are not meeting glycemic goals.

CQM Simplified

Hemoglobin A1c Test for Pediatric Patients

diabetic patients [age 5-17] *tested for HbA1c*

=

diabetic patients [age 5-17]



CQM Details

Hemoglobin A1c Test for Pediatric Patients

Tests for HbA1c

diabetic patients [age 5-17] *tested for HbA1c*

=

diabetic patients [age 5-17]

- Type 1 or Type 2 diabetes
- Excludes gestational diabetes

- Requires date of birth



CQM Implementation

Hemoglobin A1c Test for Pediatric Patients

List of LOINC codes

Tests for HbA1c

diabetic patients [age 5-17] *tested for HbA1c*

=

diabetic patients [age 5-17]

Data element

- Type 1 or Type 2 diabetes
- Excludes gestational diabetes

- Requires date of birth

List of SNOMED CT or
ICD 10 codes



Anatomy of a Clinical Quality Measure

Population criteria

- **Initial Patient Population =**
 - AND: "Patient Characteristic Birthdate: birth date" >= 5 year(s) starts before start of "Measurement Period"
 - AND: "Patient Characteristic Birthdate: birth date" <= 17 year(s) starts before start of "Measurement Period"
 - AND: "Diagnosis, Active: Diabetes" starts before or during (MOST RECENT : "Occurrence A of Encounter, Performed: Diabetes Visit" during "Measurement Period")
 - AND: "Encounter, Performed: Diabetes Visit" >= 12 month(s) starts before start of "Occurrence A of Encounter, Performed: Diabetes Visit"
- **Denominator =**
 - AND: "Initial Patient Population"
- **Denominator Exclusions =**
 - AND NOT: "Occurrence A of Diagnosis, Active: Gestational Diabetes" ends before start of "Measurement Period"
 - AND: "Occurrence A of Diagnosis, Active: Gestational Diabetes" starts before or during "Measurement Period"
- **Numerator =**
 - AND: "Laboratory Test, Result: HbA1c Laboratory Test (result)" during "Measurement Period"
- **Denominator Exceptions =**
 - None

Data criteria (QDM Data Elements)

- "Diagnosis, Active: Diabetes" using "Diabetes Grouping Value Set (2.16.840.1.113883.3.464.1003.103.12.1001)"
- "Diagnosis, Active: Gestational Diabetes" using "Gestational Diabetes Grouping Value Set (2.16.840.1.113883.3.464.1003.103.12.1010)"
- "Encounter, Performed: Diabetes Visit" using "Diabetes Visit Grouping Value Set (2.16.840.1.113883.3.464.1003.103.12.1012)"
- "Laboratory Test, Result: HbA1c Laboratory Test" using "HbA1c Laboratory Test Grouping Value Set (2.16.840.1.113883.3.464.1003.198.12.1013)"
- "Patient Characteristic Birthdate: birth date" using "birth date LOINC Value Set (2.16.840.1.113883.3.555.1.101.1)"

Value set = List of LOINC codes for HbA1c tests



Associated Value Set

Welcome

Search Value Sets

Download

Help

Apply Filters Clear Filters

Search the NLM Value Set Repository

Query: hba1c

Search

Narrow search results by selecting from pull-down menus below:

CMS eMeasure (NQF Number)

Select

Quality Data Model Category

Select

Steward

Select

Meaningful Use Measures

Select

Code System

Select

Search Results

Value Set Details

Value Set Information

Available Updates: Approved By Steward

Expansion Versions: MU2 Update 2017-01-06

Export Value Set Results (Excel)

Metadata

Name:

HbA1c Laboratory Test

OID:

2.16.840.1.113883.3.464.1003.198.12.1013

Description

Type:

Grouping

Definition Version: ?

20160331

Measure

Steward:

National Committee for Quality Assurance

Program:

CMS, MU2 Update 2017-01-06 using this value set

Grouping Members

Value Set Members

Expanded Code List

View Toggle Clear

Page 1 of 1 View 1 - 3 of 3

Code	Descriptor	Code System	Version	Code System OID
17856-6	Hemoglobin A1c/Hemoglobin.total in Blood by HPLC	LOINC	2.54	2.16.840.1.113883.6.1
4548-4	Hemoglobin A1c/Hemoglobin.total in Blood	LOINC	2.54	2.16.840.1.113883.6.1
4549-2	Hemoglobin A1c/Hemoglobin.total in Blood by Electrophoresis	LOINC	2.54	2.16.840.1.113883.6.1

View

Page 1 of 1 View 1 - 3 of 3

NLM Value Set Authority Center



Welcome back, bodenreider ▾

Welcome

Search Value Sets

Download

Help

Welcome to the NLM Value Set Authority Center (VSAC)

For VSAC announcements, please subscribe to the [VSAC Updates listserv](#).

The Value Set Authority Center (VSAC) is provided by the National Library of Medicine (NLM), in collaboration with the Office of the National Coordinator for Health Information Technology and the Centers for Medicare & Medicaid Services.

The VSAC provides downloadable access to all official versions of vocabulary value sets contained in the 2014 electronic Clinical Quality Measures (eQMs). Each value set consists of the numerical values (codes) and human-readable names (terms), drawn from standard vocabularies such as SNOMED CT®, RxNorm, LOINC and ICD-10-CM, which are used to define clinical concepts used in clinical quality measures (e.g., patients with diabetes, clinical visit). For information on the eQMs, visit the [eCOI Resource Center](#).

The content of the VSAC will gradually expand to incorporate value sets for other use cases, as well as for new measures and updates to existing measures.

Viewing or downloading value sets requires a free [Unified Medical Language System® Metathesaurus License](#), due to usage restrictions on some of the codes included in the value sets.

The [Data Element Catalog](#) contains the complete list of 2014 CQMs and value set names.

[Copyright](#), [Privacy](#), [Accessibility](#)

U.S. National Library of Medicine, 8600 Rockville Pike, Bethesda, MD 20894

[National Institutes of Health](#), [Health & Human Services](#)

[Freedom of Information Act](#), [Contact Us](#)



Last updated: Mar 03, 2017

First published: October 25, 2012

Version: 2.12.2



Data integration, exchange and semantic interoperability

Common data models

Clinical data models

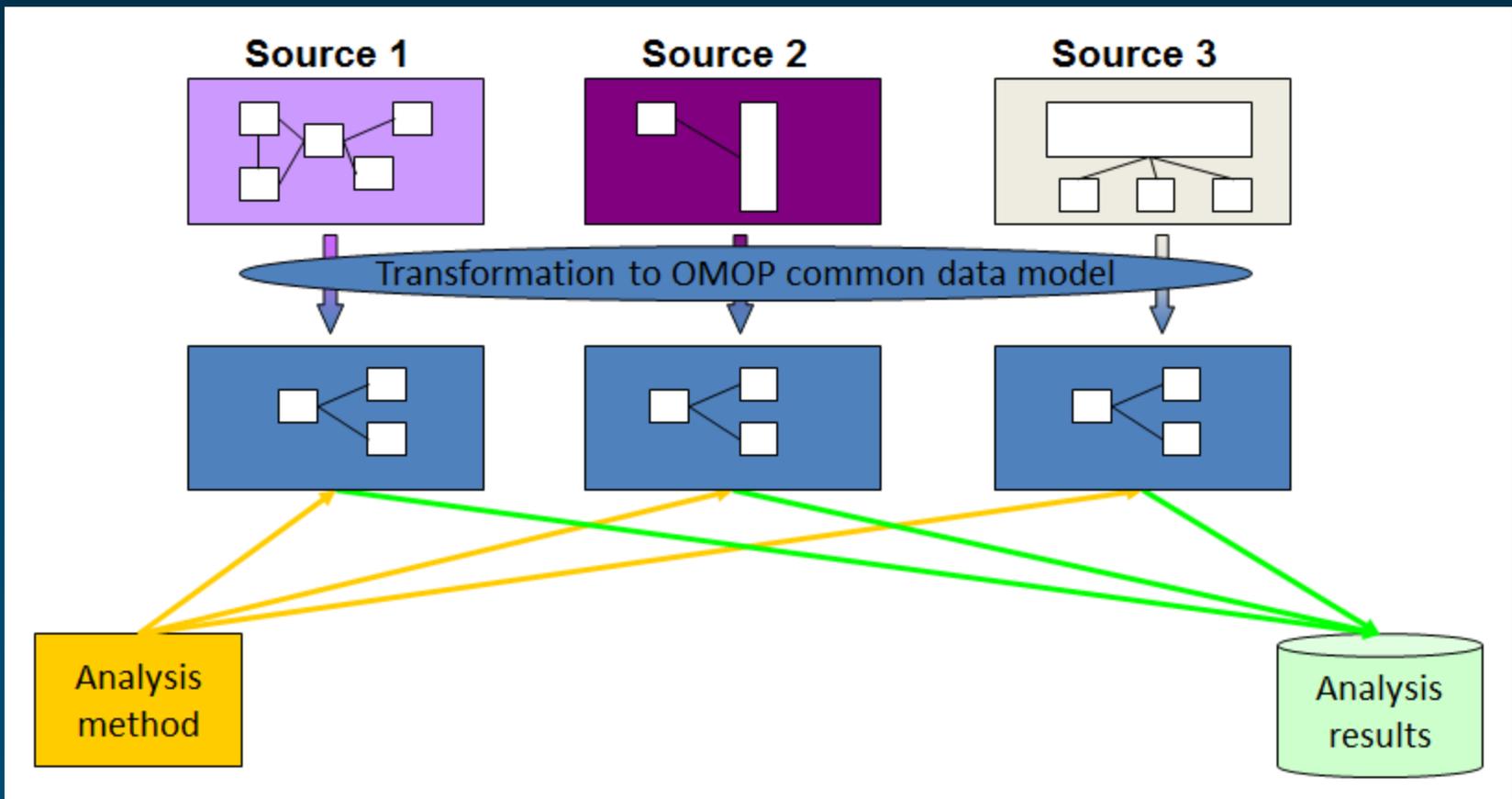
- ◆ Used in clinical data warehouses
 - Oriented towards analytics
 - Different from the transactional data models of EHR systems
 - Used to normalize data across EHR systems
- ◆ Multiple “common” data models
 - OMOP
 - i2b2
 - PCORnet
 - Sentinel
 - CDISC



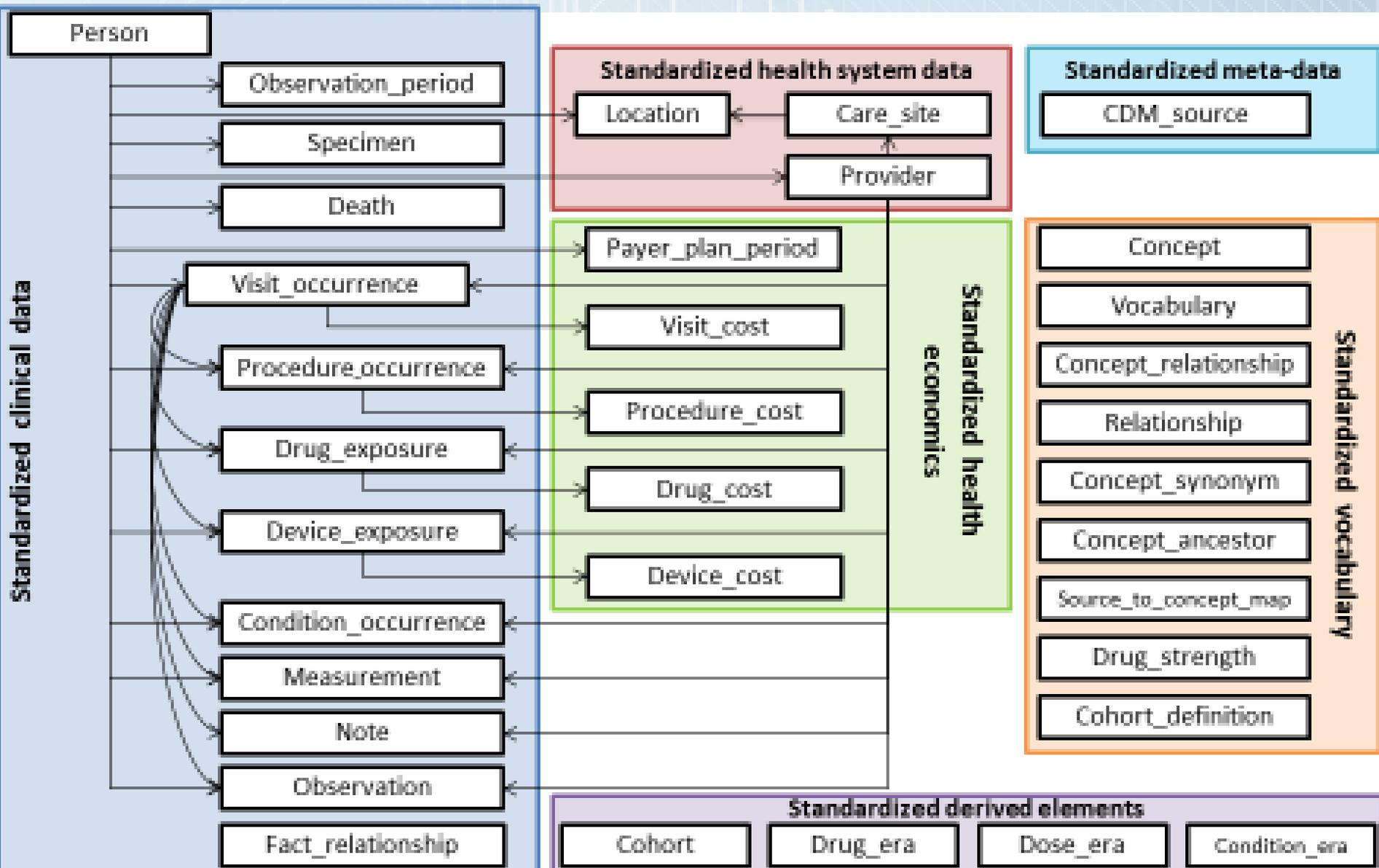


OMOP

◆ OMOP – Observational Medical Outcomes Partnership



OMOP Common Data Model



- Standardized Clinical Data Tables
 - PERSON
 - OBSERVATION_PERIOD
 - SPECIMEN
 - DEATH
 - VISIT_OCCURRENCE
 - PROCEDURE_OCCURRENCE
 - DRUG_EXPOSURE
 - DEVICE_EXPOSURE
 - CONDITION_OCCURRENCE
 - MEASUREMENT
 - NOTE
 - NOTE_NLP (V5 5.2)
 - OBSERVATION
 - FACT_RELATIONSHIP
- Standardized Health System Data Tables
 - LOCATION
 - CARE_SITE
 - PROVIDER
- Standardized Health Economics Data Tables
 - PAYER_PLAN_PERIOD
 - COST (V5.0.1)
 - VISIT_COST - removed
 - PROCEDURE_COST - removed
 - DRUG_COST - removed
 - DEVICE_COST - removed
- Standardized Derived Elements
 - COHORT
 - COHORT_ATTRIBUTE
 - DRUG_ERA
 - DOSE_ERA
 - CONDITION_ERA

- Standardized Vocabularies
 - CONCEPT
 - VOCABULARY
 - DOMAIN
 - CONCEPT_CLASS
 - CONCEPT_RELATIONSHIP
 - RELATIONSHIP
 - CONCEPT_SYNONYM
 - CONCEPT_ANCESTOR
 - SOURCE_TO_CONCEPT_MAP
 - DRUG_STRENGTH
 - COHORT_DEFINITION
 - ATTRIBUTE_DEFINITION
- Standardized meta-data
 - CDM_SOURCE



i2b2

- ◆ i2b2 – Informatics for Integrating Biology & the Bedside
- ◆ Originally developed by the i2b2 National Center for Biomedical Computing (2004-2013)
 - Now i2b2 tranSMART Foundation
- ◆ Platform to support translational research
- ◆ Widely adopted worldwide

i2b2 data model – original “star schema”

i2b2 Star Schema

visit_dimension		
PK	encounter_num	INTEGER
PK	patient_num	INTEGER
	inout_cd	VARCHAR(10)
	location_cd	VARCHAR(100)
	location_path	VARCHAR(700)
	start_date	DATETIME
	end_date	DATETIME
	visit_blob	TEXT(10)

patient_dimension		
PK	patient_num	INTEGER
	vital_status_cd	VARCHAR(10)
	birth_date	DATETIME
	death_date	DATETIME
	sex_cd	CHAR(10)
	age_in_years_num	INTEGER
	language_cd	VARCHAR(100)
	race_cd	VARCHAR(100)
	marital_status_cd	VARCHAR(100)
	religion_cd	VARCHAR(100)
	zip_cd	VARCHAR(20)
	statecityzip_path	VARCHAR(200)
	patient_blob	TEXT(10)

observation_fact		
PK	encounter_num	INTEGER
PK	concept_cd	VARCHAR(20)
PK	provider_id	VARCHAR(20)
PK	start_date	DATETIME
PK	modifier_cd	CHAR(1)
	patient_num	INTEGER
	valtype_cd	CHAR(1)
	tval_char	VARCHAR(50)
	nval_num	DECIMAL(10,2)
	valueflag_cd	CHAR(1)
	quantity_num	DECIMAL(10,2)
	units_cd	VARCHAR(100)
	end_date	DATETIME
	location_cd	TEXT(100)
	confidence_num	VARCHAR(100)
	observation_blob	TEXT(10)

concept_dimension		
PK	concept_path	VARCHAR(700)
	concept_cd	VARCHAR(20)
	name_char	VARCHAR(2000)
	concept_blob	TEXT(10)

provider_dimension		
PK	provider_path	VARCHAR(800)
	provider_id	VARCHAR(20)
	name_char	VARCHAR(2000)
	provider_blob	TEXT(10)



i2b2-OMOP convergence

◆ i2b2 on OMOP

- Supports query formulation against an OMOP-compliant data source through i2b2 tools





PCORnet

- ◆ PCORnet – National Patient-Centered Clinical Research Network
- ◆ Initiative of the Patient-Centered Outcomes Research Institute (PCORI)
 - Funded through the Patient Protection and Affordable Care Act of 2010
- ◆ “designed to make it faster, easier, and less costly to conduct clinical research”
- ◆ Made up of
 - 13 Clinical Data Research Networks (CDRNs)
 - 20 Patient-Powered Research Networks (PPRNs)



PCORnet Common Data Model v3.0

New to v3.0

DEMOGRAPHIC

PATID
BIRTH_DATE
BIRTH_TIME
SEX
HISPANIC
RACE
BIOBANK_FLAG

Fundamental basis

ENROLLMENT

PATID
ENR_START_DATE
ENR_END_DATE
CHART
ENR_BASIS

DISPENSING

DISPENSINGID
PATID
PRESCRIBINGID (optional)
DISPENSE_DATE
NDC
DISPENSE_SUP
DISPENSE_AMT

DEATH

PATID
DEATH_DATE
DEATH_DATE_IMPUTE
DEATH_SOURCE
DEATH_MATCH_CONFIDENCE

DEATH_CONDITION

PATID
DEATH_CAUSE
DEATH_CAUSE_CODE
DEATH_CAUSE_TYPE
DEATH_CAUSE_SOURCE
DEATH_CAUSE_CONFIDENCE

Data captured from processes associated with healthcare delivery

VITAL

VITALID
PATID
ENCOUNTERID (optional)
MEASURE_DATE
MEASURE_TIME
VITAL_SOURCE
HT
WT
DIASTOLIC
SYSTOLIC
ORIGINAL_BMI
BP_POSITION
SMOKING
TOBACCO
TOBACCO_TYPE

CONDITION

CONDITIONID
PATID
ENCOUNTERID (optional)
REPORT_DATE
RESOLVE_DATE
ONSET_DATE
CONDITION_STATUS
CONDITION
CONDITION_TYPE
CONDITION_SOURCE

PRO_CM

PRO_CM_ID
PATID
ENCOUNTERID (optional)
PRO_ITEM
PRO_LOINC
PRO_DATE
PRO_TIME
PRO_RESPONSE
PRO_METHOD
PRO_MODE
PRO_CAT

Data captured within multiple contexts: healthcare delivery, registry activity, or directly from patients

ENCOUNTER

ENCOUNTERID
PATID
ADMIT_DATE
ADMIT_TIME
DISCHARGE_DATE
DISCHARGE_TIME
PROVIDERID
FACILITY_LOCATION
ENC_TYPE
FACILITYID
DISCHARGE_DISPOSITION
DISCHARGE_STATUS
DRG
DRG_TYPE
ADMITTING_SOURCE

DIAGNOSIS

DIAGNOSISID
PATID
ENCOUNTERID
ENC_TYPE (replicated)
ADMIT_DATE (replicated)
PROVIDERID (replicated)
DX
DX_TYPE
DX_SOURCE
PDX

PROCEDURES

PROCEDURESID
PATID
ENCOUNTERID
ENC_TYPE (replicated)
ADMIT_DATE (replicated)
PROVIDERID (replicated)
PX_DATE
PX
PX_TYPE
PX_SOURCE

Data captured from healthcare delivery, direct encounter basis

LAB_RESULT_CM

LAB_RESULT_CM_ID
PATID
ENCOUNTERID (optional)
LAB_NAME
SPECIMEN_SOURCE
LAB_LOINC
PRIORITY
RESULT_LOC
LAB_PX
LAB_PX_TYPE
LAB_ORDER_DATE
SPECIMEN_DATE
SPECIMEN_TIME
RESULT_DATE
RESULT_TIME
RESULT_QUAL
RESULT_NUM
RESULT_MODIFIER
RESULT_UNIT
NORM_RANGE_LOW
NORM_MODIFIER_LOW
NORM_RANGE_HIGH
NORM_MODIFIER_HIGH
ABN_IND

PRESCRIBING

PRESCRIBINGID
PATID
ENCOUNTERID (optional)
RX_PROVIDERID
RX_ORDER_DATE
RX_ORDER_TIME
RX_START_DATE
RX_END_DATE
RX_QUANTITY
RX_REFILLS
RX_DAYS_SUPPLY
RX_FREQUENCY
RX_BASIS
RXNORM_CUI

PCORNET_TRIAL

PATID
TRIALID
PARTICIPANTID
TRIAL_SITEID
TRIAL_ENROLL_DATE
TRIAL_END_DATE
TRIAL_WITHDRAW_DATE
TRIAL_INVITE_CODE

Associations with PCORnet clinical trials

HARVEST

NETWORKID
NETWORK_NAME
DATAMARTID
DATAMART_NAME
DATAMART_PLATFORM
CDM_VERSION
DATAMART_CLAIMS
DATAMART_EHR
BIRTH_DATE_MGMT
ENR_START_DATE_MGMT
ENR_END_DATE_MGMT
ADMIT_DATE_MGMT
DISCHARGE_DATE_MGMT
PX_DATE_MGMT
RX_ORDER_DATE_MGMT
RX_START_DATE_MGMT
RX_END_DATE_MGMT
DISPENSE_DATE_MGMT
LAB_ORDER_DATE_MGMT
SPECIMEN_DATE_MGMT
RESULT_DATE_MGMT
MEASURE_DATE_MGMT
ONSET_DATE_MGMT
REPORT_DATE_MGMT
RESOLVE_DATE_MGMT
PRO_DATE_MGMT
REFRESH_DEMOGRAPHIC_DATE
REFRESH_ENROLLMENT_DATE
REFRESH_ENCOUNTER_DATE
REFRESH_DIAGNOSIS_DATE
REFRESH_PROCEDURES_DATE
REFRESH_VITAL_DATE
REFRESH_DISPENSING_DATE
REFRESH_LAB_RESULT_CM_DATE
REFRESH_CONDITION_DATE
REFRESH_PRO_CM_DATE
REFRESH_PRESCRIBING_DATE
REFRESH_PCORNET_TRIAL_DATE
REFRESH_DEATH_DATE
REFRESH_DEATH_CAUSE_DATE

Process-related data

Bold font indicates fields that cannot be null due to primary key definitions or record-level constraints.



Sentinel

- ◆ Initiative of the Food and Drug Administration (FDA)
- ◆ Effort to create a national electronic system for monitoring the performance of FDA-regulated medical products (drugs, vaccines, and other biologics)
- ◆ Develop a system to obtain information from existing electronic health care data from multiple sources to assess the safety of approved medical products
- ◆ Distributed dataset reached 100 lives in 2011



Sentinel Common Data Model

List of Tables



Table Name	Source	Description
1. Enrollment	Created by Data Partners using Data Partner data.	The SCDM Enrollment Table has a start/stop structure that contains one record per continuous enrollment period. Members with medical coverage, drug coverage, or both should be included. A unique combination of PatID, Enr_Start, Enr_End, MedCov, DrugCov, and Chart identifies a unique record. A break in enrollment (of at least one day) or a change in either the medical or drug coverage variables should generate a new record.
2. Demographic	Created by Data Partners using Data	The SCDM Demographic Table contains one record per PatID with the most recent information on Birth_Date, Sex, Race/Ethnicity, and Zip Code.
3. Dispensing	Created by Data Partners using Data Partner data.	The SCDM Outpatient Pharmacy Dispensing Table contains one record per unique combination of PatID, NDC, and RxDate. Each record represents an outpatient pharmacy dispensing. Rollback transactions and other adjustments should be processed before populating this table.
4.1 Encounter	Created by Data Partners using Data Partner data.	The SCDM Encounter Table contains one record per PatID and EncounterID. Each encounter should have a single record in the SCDM Encounter Table. Each diagnosis and procedure recorded during the encounter should have a separate record in the Diagnosis or Procedure Tables. Multiple visits to the same provider on the same day should be considered one encounter and should include all diagnoses and procedures that were recorded during those visits. Visits to different providers on the same day, such as a physician appointment that leads to a hospitalization, should be considered multiple encounters. Rollback transactions and other adjustments should be processed before populating this table.
4.2 Diagnosis	Created by Data Partners using Data Partner data.	The SCDM Diagnosis Table contains one record per unique combination of PatID, EncounterID, DX, and DX_CodeType. This table should capture all uniquely recorded diagnoses for all encounters.
4.3 Procedure	Created by Data Partners using Data	The SCDM Procedure Table contains one record per unique combination of PatID, EncounterID, F PX_CodeType. This table should capture all uniquely recorded procedures for all encounters.
5.1 Death	Created by Data Partners using Data Partner data.	The SCDM Death Table contains one record per PatID. When legacy data have conflicting report make a local determination as to which to use. There is typically a 1-2 year lag in death registry
5.2 Cause of Death	Created by Data Partners using Data Partner data.	The SCDM Cause of Death Table contains one record per unique combination of PatID and COD. legacy data have conflicting reports, please make a local determination as to which to use. Ther typically a 1-2 year lag in death registry data.
6.1 Laboratory Result	Created by Data Partners using Data Partner data.	The SCDM Laboratory Result Table contains one record per result/entry. Only include resulted la Data Partners are strongly encouraged to review the comprehensive Sentinel Common Data Mod Laboratory Result Table Documentation for details on how to populate each variable.

List of Tables (cont.)

Table Name

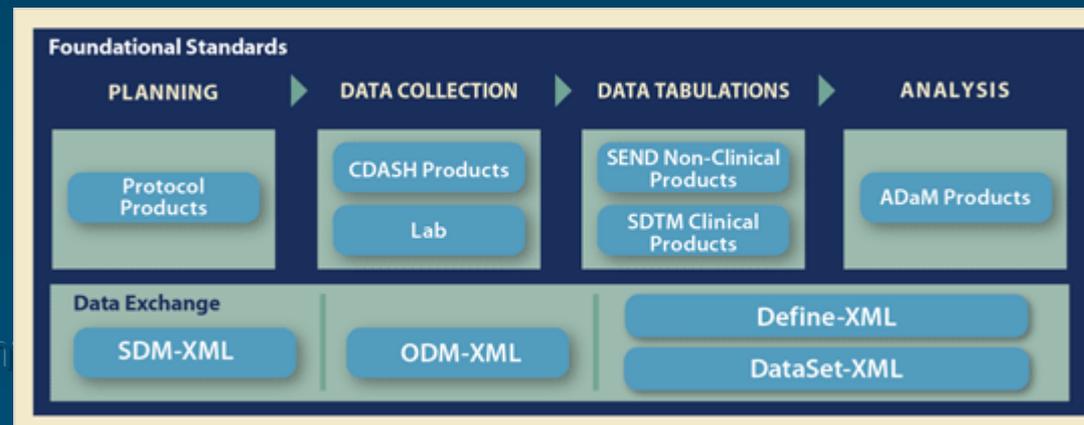
[6.2 Vital Signs](#)

[7. Inpatient Pharmacy](#)

[8. Inpatient Transfusion](#)

CDISC

- ◆ CDISC – Clinical Data Interchange Standards Consortium
- ◆ “develop and support global, platform-independent data standards that enable information system interoperability to improve medical research and related areas of healthcare”
- ◆ Set of standards required by FDA for regulatory submissions (clinical research)



Common data models in action



COLLOQUIUM
PAPER

Characterizing treatment pathways at scale using the OHDSI network

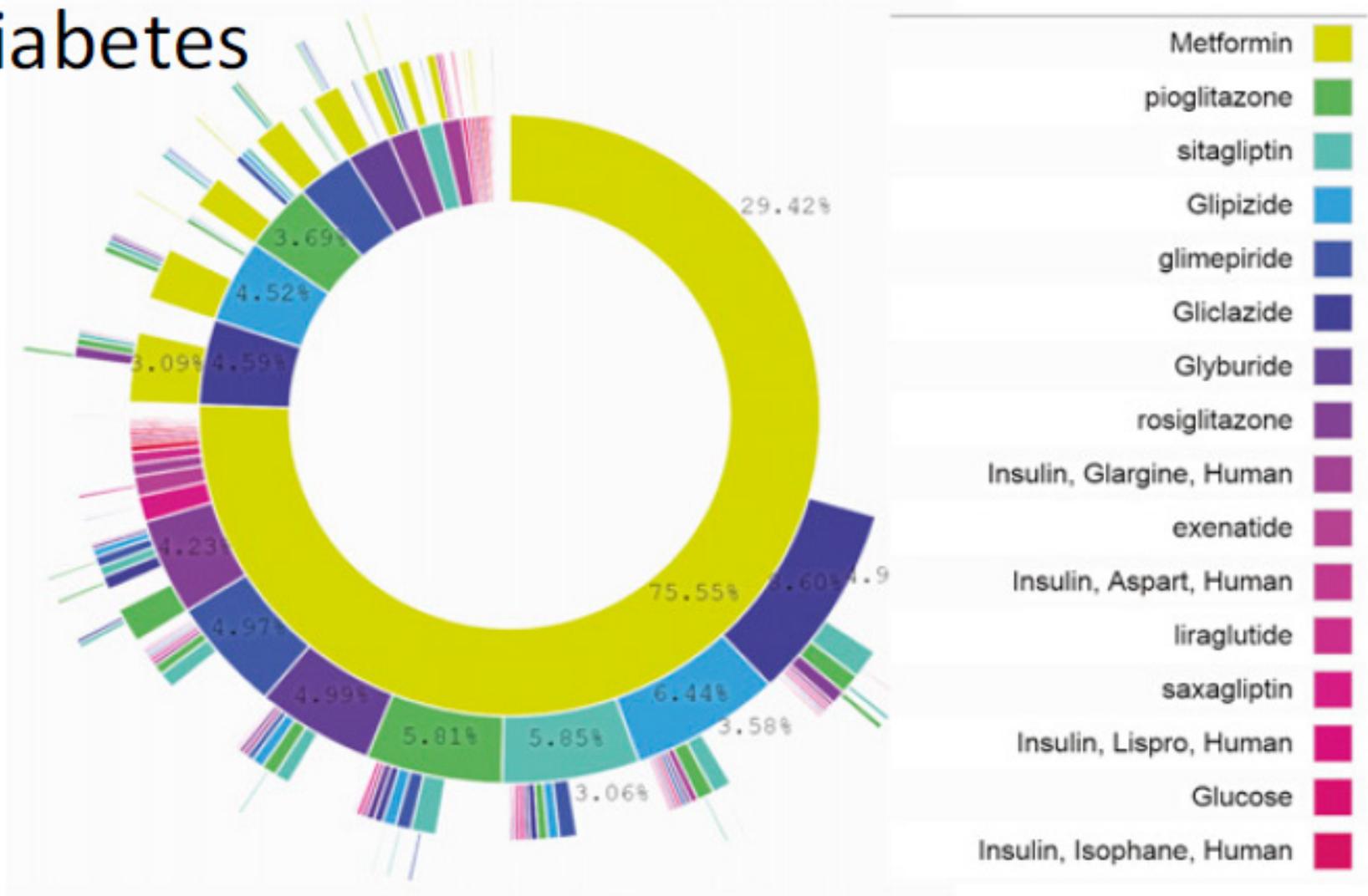
George Hripcsak^{a,b,c,1}, Patrick B. Ryan^{c,d}, Jon D. Duke^{c,e}, Nigam H. Shah^{c,f}, Rae Woong Park^{c,g}, Vojtech Huser^{c,h}, Marc A. Suchard^{c,i,j,k}, Martijn J. Schuemie^{c,d}, Frank J. DeFalco^{c,d}, Adler Perotte^{a,c}, Juan M. Banda^{c,f}, Christian G. Reich^{c,l}, Lisa M. Schilling^{c,m}, Michael E. Matheny^{c,n,o}, Daniella Meeker^{c,p,q}, Nicole Pratt^{c,r}, and David Madigan^{c,s}

www.pnas.org/cgi/doi/10.1073/pnas.1510502113

PNAS | July 5, 2016 | vol. 113 | no. 27 | 7329–7336



A Diabetes



Data integration, exchange and semantic interoperability

*Fast Healthcare Interoperability Resources
(FHIR)*

Fast Healthcare Interoperability Resources

- ◆ New standard for exchanging healthcare information electronically
- ◆ Developed by HL7 FHIR foundation
- ◆ Based on resources
 - Basic building blocks of information (patient, condition, procedure, practitioner)
 - Can be extended as needed
- ◆ Supports 4 different paradigms for exchange: the RESTful API, Messaging, Documents, and Services

Clinical Terminology

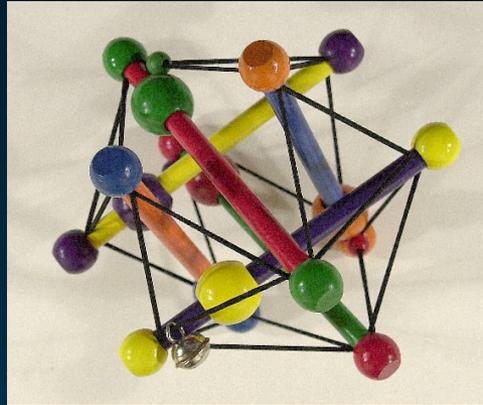
Summary

Summary

- ◆ How clinical features are reflected in disease names
- ◆ Structure of the main clinical ontologies used
 - ICD, SNOMED CT, LOINC, RxNorm
- ◆ Use cases for biomedical ontologies
 - Knowledge management
 - Decision support and analytics
 - Data integration, exchange and semantic interoperability

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
- ◆ 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



U.S. National Library of Medicine

