Clinical Terminology
(aka Biomedical Terminologies/Ontologies)

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


Additional references


Outline

◆ Historical perspective
◆ Introduction to biomedical terminologies through an example

◆ “High-Impact” Biomedical Ontologies
  ● Structural perspective

◆ Biomedical Ontologies “in Action”
  ● Functional perspective
Clinical Terminology
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)
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...
… 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
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)
“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.*

From “bad air” to “bad water” (John Snow)
History of Medical Ontologies

Synopsis
Nosologiae Methodicae

1603
1700
1785
1855
1900
1975

ICD
ICD9

ICD

OPCS

EmTree

MeSH

OPCS

SNOP

CPT

SNOMED-2

SNOMED

International

SNOMED-RT

SNOMED-CT

FMA

GALEN

DM&D

OPCS3

OPCS4

OPCS4.3

UMLS

READ

CTV3

1975
1985
1995
2005

1900

1985

1900

[1603-1700]

[1700-1785]

[1785-1855]

[1855-1900]

[1900-1975]

[1975-2005]

[Lister Hill National Center for Biomedical Communications]

[Bodenreider, BIB 2006]

[courtesy of J. Rogers]
Clinical Terminology
Part 2

Introduction to biomedical terminologies through an example
Thomas Addison (1795-1860)
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

◆ Contexts: different hierarchies
  - eponym
  - symptoms
  - clinical
  - variants
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
E27.6 Other specified disorders of adrenal gland
E27.7 Endocrine gland, unspecified

E28 Ovarian dysfunction
E29 Testicular dysfunction
E30 Disorders of puberty, not elsewhere classified
E31 Polyglandular dysfunction
E32 Diseases of thymus
E33 Other endocrine disorders
E34 Disorders of endocrine glands in diseases classified elsewhere

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

E27.6 Other specified disorders of adrenal gland

E27.7 Endocrine gland, unspecified

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

E27.6 Other specified disorders of adrenal gland

E27.7 Endocrine 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
Clinical Terminology
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 (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/](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
Idiosyncratic inclusion/exclusion criteria

Type 1 diabetes mellitus

**Incl.:**
- 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)
postoperative hypoinsulinaemia (E89.1)
ICD Example

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

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>E84</td>
<td>Cystic fibrosis</td>
</tr>
<tr>
<td>E84.0</td>
<td>Cystic fibrosis with pulmonary manifestations</td>
</tr>
<tr>
<td>E84.1</td>
<td>Cystic fibrosis with intestinal manifestations</td>
</tr>
<tr>
<td>E84.8</td>
<td>Cystic fibrosis with other manifestations</td>
</tr>
<tr>
<td>E84.9</td>
<td>Cystic fibrosis, unspecified</td>
</tr>
</tbody>
</table>

*Incl.*: mucoviscidosis

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>E84</td>
<td>Cystic fibrosis</td>
</tr>
</tbody>
</table>

*Excl.*: meconium obstruction (ileus) in cases where cystic fibrosis is known not to be present (P76.0)

*†* P75*:*
**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](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 (R75)

Excludes:
- aromatic amino-acid metabolism (E70.1)
- branched-chain amino-acid metabolism (E71.3)
- fatty-acid metabolism (E71.3)
- purine and pyrimidine metabolism (E79.3)
- gout (M10.0)

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 (E70.5)
ICD-10 vs. ICD-10-CM

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

W58.1 Contact with crocodile
- W58.11 Bitten by crocodile
- W58.12 Struck by crocodile
- W58.13 Crushed by crocodile
- W58.19 Other contact with crocodile

W58.01A Bitten by alligator, initial encounter
W58.01D Bitten by alligator, subsequent encounter
W58.01S Bitten by alligator, sequela
Logical Observation Identifiers, Names and Codes (LOINC)
LOINC Characteristics (1)

- Current version: 2.59 (Feb. 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]

<table>
<thead>
<tr>
<th>Axis</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component</td>
<td>Sodium</td>
</tr>
<tr>
<td>Property</td>
<td>SCnc – Substance Concentration (per volume)</td>
</tr>
<tr>
<td>Timing</td>
<td>Pt – Point in time (Random)</td>
</tr>
<tr>
<td>System</td>
<td>Ser/Plas – Serum or Plasma</td>
</tr>
<tr>
<td>Scale</td>
<td>Qn – Quantitative</td>
</tr>
<tr>
<td>Method</td>
<td>--</td>
</tr>
</tbody>
</table>
## NAME

<table>
<thead>
<tr>
<th>Fully-Specified Name:</th>
<th>Component</th>
<th>Property</th>
<th>Time</th>
<th>System</th>
<th>Scale</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sodium</td>
<td></td>
<td>SCnc</td>
<td>Pt</td>
<td>Ser/Plas</td>
<td>Qn</td>
<td></td>
</tr>
</tbody>
</table>

## 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/](http://creativecommons.org/licenses/by-sa/3.0/) for details.

**Source:** Wikipedia, URL: [Sodium (Wikipedia)](https://en.wikipedia.org/wiki/Sodium)

## 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

<table>
<thead>
<tr>
<th>Unit</th>
<th>Source Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>mmol/L</td>
<td>EXAMPLE UCUM UNITS</td>
</tr>
<tr>
<td>mmol/L</td>
<td>REGENSTRIEF</td>
</tr>
<tr>
<td>mmol/L</td>
<td>eCHN</td>
</tr>
</tbody>
</table>

## UNITS AND RANGE

<table>
<thead>
<tr>
<th>Range</th>
<th>Units Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>mmol/L [136,145]</td>
<td></td>
</tr>
</tbody>
</table>

*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, 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
- Appendectomy
- 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: March 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
<table>
<thead>
<tr>
<th>Strength</th>
<th>Ingredient</th>
<th>Dose form</th>
</tr>
</thead>
<tbody>
<tr>
<td>4mg/ml</td>
<td>Fluoxetine</td>
<td>Oral Solution</td>
</tr>
</tbody>
</table>

Semantic clinical drug component

Semantic clinical drug form

Semantic clinical drug
RxNorm Example

- **Ingredient**: Azithromycin
- **C. Drug Comp.**: Azithromycin 250 MG
- **C. Drug Form**: Azithromycin Oral Tablet
- **B. Drug Comp.**: Azithromycin 250 MG
- **B. Drug Form**: Azithromycin Oral Tablet
- **B. Drug**: Zithromax
- **B. Pack**: Zithromax 250 MG Oral Tablet
- **G. Pack**: {6 (Azithromycin 250 MG Oral Tablet)} Pack
- **B. Pack**: Z-PAK
Clinical Terminology
Part 4

Biomedical Ontologies “in Action”

A Functional Perspective
Overview

◆ Functional perspective
  ◆ 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)

[Bodenreider, YBMI 2008]
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)
## ALDH2  aldehyde dehydrogenase 2 family (mitochondrial) [ *Homo sapiens* (human) ]

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

<table>
<thead>
<tr>
<th>Function</th>
<th>Evidence Code</th>
<th>Pubs</th>
</tr>
</thead>
<tbody>
<tr>
<td>aldehyde dehydrogenase (NAD) activity</td>
<td>EXP</td>
<td></td>
</tr>
<tr>
<td>aldehyde dehydrogenase (NAD) activity</td>
<td>IDA</td>
<td>PubMed</td>
</tr>
<tr>
<td>aldehyde dehydrogenase [NAD(P)+] activity</td>
<td>TAS</td>
<td>PubMed</td>
</tr>
<tr>
<td>electron carrier activity</td>
<td>TAS</td>
<td>PubMed</td>
</tr>
</tbody>
</table>

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

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

<table>
<thead>
<tr>
<th>Function</th>
<th>Evidence Code</th>
<th>Pubs</th>
</tr>
</thead>
<tbody>
<tr>
<td>NADH binding</td>
<td>ISO</td>
<td></td>
</tr>
<tr>
<td>aldehyde dehydrogenase (NAD) activity</td>
<td>IDA</td>
<td></td>
</tr>
<tr>
<td>aldehyde dehydrogenase (NAD) activity</td>
<td>ISO</td>
<td></td>
</tr>
<tr>
<td>identical protein binding</td>
<td>ISO</td>
<td></td>
</tr>
<tr>
<td>oxidoreductase activity</td>
<td>IEA</td>
<td></td>
</tr>
<tr>
<td>oxidoreductase activity, acting on the aldehyde or o xo group of donors. NAD or NADP as acceptor</td>
<td>IEA</td>
<td></td>
</tr>
<tr>
<td>protein binding</td>
<td>PI</td>
<td>PubMed</td>
</tr>
</tbody>
</table>
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
Free cortisol in sepsis and septic shock.

Bendel S¹, Karlsson S, Pettilä V, Loisa P, Varpula M, Ruokonen E; Finnsepsis Study Group.

Abstract

BACKGROUND: Severe sepsis activates the hypothalamic-pituitary 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.
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

<table>
<thead>
<tr>
<th>MeSH Heading</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Addison Disease</td>
<td></td>
</tr>
<tr>
<td>Entry Term</td>
<td>Addison's Disease</td>
</tr>
<tr>
<td>Entry Term</td>
<td>Primary Adrenal Insufficiency</td>
</tr>
<tr>
<td>Entry Term</td>
<td>Primary Adrenocortical Insufficiency</td>
</tr>
<tr>
<td>Entry Term</td>
<td>Primary Hypoadrenalism</td>
</tr>
</tbody>
</table>
MeSH “synonyms” MEDLINE retrieval

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]
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

**PubMed**

Adrenal insufficiency in prolonged critical illness.
Wu JY, Hsu SC, Ku SC, Ho CC, Yu CJ, Yang PC.
PMID: 18466605  Free PMC Article

Addison's disease: a rare cause of hypertransaminasaemia.
Ersan O, Demirezer B.
PMID: 18465237

MeSH Terms
Adrenal Insufficiency/blood
Adrenal Insufficiency/drug therapy
Adrenal Insufficiency/mortality*

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

- Clinical repositories
- Genetic knowledge bases
- SNOMED CT
- OMIM
- MeSH
- Biomedical literature
- GO
- Genome annotations
- FMA
- Anatomy
- NCBI Taxonomy
- Model organisms
- Other subdomains
Integrating subdomains

- Clinical repositories
- Genetic knowledge bases
- Biomedical literature
- Genome annotations
- Anatomy
- Model organisms
- Other subdomains
Trans-namespace integration

Addison's disease (363732003)

Other subdomains

Clinical repositories

SNOMED CT

Genetic knowledge bases

OMIM

Biomedical literature

Addison Disease (D000224)

NCBI Taxonomy

Model organisms

FMA

Anatomy

Genome annotations

UMLS C0001403

MeSH

Clinical repositories

Biomedical literature
UMLS Source Vocabularies

- 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

Hemoglobin | Sugar
---|---

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.
Hemoglobin A1c Test for Pediatric Patients

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

# diabetic patients [age 5-17]
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
Hemoglobin A1c Test for Pediatric Patients

- Type 1 or Type 2 diabetes
- Excludes gestational diabetes

Tests for HbA1c

List of LOINC codes

List of SNOMED CT or ICD 10 codes

Data element

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

# diabetic patients [age 5-17]

- Requires date of birth
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 (2.16.840.1.113883.3.557)"

Value set = List of LOINC codes for HbA1c tests
## Associated Value Set

- **Name:** HbA1c Laboratory Test
- **Type:** Grouping
- **Steward:** National Committee for Quality Assurance

### Value Set Members

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>17856-6</td>
<td>Hemoglobin A1c/Hemoglobin total in Blood by HPLC</td>
<td>LOINC</td>
<td>2.54</td>
<td>2.16.840.1.113883.3.464.1003.198.12.1013</td>
</tr>
<tr>
<td>4548-4</td>
<td>Hemoglobin A1c/Hemoglobin total in Blood</td>
<td>LOINC</td>
<td>2.54</td>
<td>2.16.840.1.113883.3.464.1003.198.12.1013</td>
</tr>
<tr>
<td>4549-2</td>
<td>Hemoglobin A1c/Hemoglobin total in Blood by Electrophoresis</td>
<td>LOINC</td>
<td>2.54</td>
<td>2.16.840.1.113883.3.464.1003.198.12.1013</td>
</tr>
</tbody>
</table>
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 (eCQMs). 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 eCQMs, visit the eCQI 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.
Clinical Terminology

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
- OBO ontologies, OBO Foundry
- Coordinated development of ontologies, harmonization
- Boundary between terminology and information model
- [...]

Lister Hill National Center for Biomedical Communications

118