The Unified Medical Language System

*What is it and how to use it?*

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

◆ What is the UMLS?
  ● Introduction
  ● Overview through an example
  ● The three UMLS Knowledge Sources

◆ How to use the UMLS?
  ● Obtaining a license
  ● Remote access
  ● Local installation and customization
  ● A UMLS-based algorithm
  ● Benefits and limitations
Part I

What is the UMLS?
Outline

◆ Part I: What is the UMLS?
  ● Introduction
  ● Overview through an example
  ● The three UMLS Knowledge Sources
    ■ UMLS Metathesaurus
    ■ UMLS Semantic Network
    ■ SPECIALIST Lexicon and lexical tools
Part I
What is the UMLS?

(1) Introduction
What does UMLS stand for?

- Unified
- Medical
- Language
- System

UMLS®
Unified Medical Language System®
UMLS Metathesaurus®
Motivation

- Started in 1986
- National Library of Medicine
- “Long-term R&D project”
- Complementary to IAIMS

«[...] the UMLS project is an effort to overcome two significant barriers to effective retrieval of machine-readable information.

- The first is the variety of ways the same concepts are expressed in different machine-readable sources and by different people.
- The second is the distribution of useful information among many disparate databases and systems.»
The UMLS in practice

- **Database**
  - Series of relational files

- **Interfaces**
  - Web interface: Knowledge Source Server (UMLSKS)
  - Application programming interfaces
    (Java and XML-based)

- **Applications**
  - lvgl (lexical programs)
  - MetamorphoSys (installation and customization)
  - RRF browser (browsing subsets)

The UMLS is *not* an end-user application
Part I
What is the UMLS?

(2) Overview through an example
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
  - Melasma addisonii
  - Asthenia pigmentosa
  - Primary adrenal deficiency
  - Primary adrenal insufficiency
  - Primary adrenocortical insufficiency
  - Chronic adrenocortical insufficiency

◆ Contexts: different hierarchies
  - eponym
  - symptoms
  - clinical
  - variants
Organize terms

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

<table>
<thead>
<tr>
<th>Term</th>
<th>MeSH</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Addison Disease</td>
<td>D000224</td>
<td></td>
</tr>
<tr>
<td>Primary hypoadrenalism</td>
<td>MedDRA 10036696</td>
<td></td>
</tr>
<tr>
<td>Primary adrenocortical insufficiency</td>
<td>ICD-10 E27.1</td>
<td></td>
</tr>
<tr>
<td>Addison's disease (disorder)</td>
<td>SNOMED CT 363732003</td>
<td></td>
</tr>
<tr>
<td>C0001403</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Addison's disease
Endocrine disorders

Adrenal gland disorders

Adrenal cortical hypofunctions

Addison’s Disease
Diseases of the endocrine system

Diseases of the adrenal glands

Addison’s Disease
Disorder of endocrine system

Disorder of adrenal gland

Hypoadrenalism

Adrenal hypofunction

Disorder of adrenal cortex

Adrenal cortical hypofunction

Addison’s Disease
Organize concepts

- Inter-concept relationships: hierarchies from the source vocabularies
- Redundancy: multiple paths
- One graph instead of multiple trees (multiple inheritance)
organize concepts

Disease

Endocrine system diseases

Adrenal gland diseases

Adrenal gland hypofunction

Adrenal cortex diseases

Adrenal cortical hypofunction

Addison’s Disease
Relate to other concepts

- Additional hierarchical relationships
  - link to other trees
  - make relationships explicit
- Non-hierarchical relationships
- Co-occurring concepts
- Mapping relationships
Categorize concepts

- High-level categories (semantic types)
- Assigned by the Metathesaurus editors
- Independently of the hierarchies in which these concepts are located
How do they do that?

- Lexical knowledge
- Semantic pre-processing
- UMLS editors
Lexical knowledge

- Adrenal gland diseases
- Adrenal disorder
- Disorder of adrenal gland
- Diseases of the adrenal glands
- C0001621
Semantic pre-processing

- Metadata in the source vocabularies
- Tentative categorization
- Positive (or negative) evidence for tentative synonymy relations based on lexical features
Additional knowledge: UMLS editors

- Adrenal gland diseases
  - Adrenal gland hypofunction
  - Adrenal cortical hypofunction
    - Addison’s Disease
  - Adrenal cortex diseases
  - Other disorders of adrenal gland
UMLS Summary

- Synonymous terms clustered into concepts
- Unique identifier
- Finer granularity
- Broader scope
- Additional hierarchical relationships
- Semantic categorization
Part I
What is the UMLS?

(3) UMLS Knowledge Sources
Unified Medical Language System

- **SPECIALIST Lexicon**
  - 360,000 lexical items
  - Part of speech and variant information

- **Metathesaurus**
  - 6M names from over 100 terminologies
  - 1.5M concepts
  - 8M relations

- **Semantic Network**
  - 135 high-level categories
  - 7000 relations among them
UMLS Metathesaurus
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
Source Vocabularies

- 143 source vocabularies
  - 17 languages
- Broad coverage of biomedicine
  - 5.9M names
  - 1.4M concepts
  - 8M relations
- Common presentation
Biomedical terminologies

◆ General vocabularies
  ● anatomy (UWDA, Neuronames)
  ● drugs (RxNorm, First DataBank, Micromedex)
  ● medical devices (UMD, SPN)

◆ Several perspectives
  ● clinical terms (SNOMED CT)
  ● information sciences (MeSH, CRISP)
  ● administrative terminologies (ICD-9-CM, CPT-4)
  ● data exchange terminologies (HL7, LOINC)
Biomedical terminologies (cont’d)

- Specialized vocabularies
  - nursing (NIC, NOC, NANDA, Omaha, PCDS)
  - dentistry (CDT)
  - oncology (PDQ)
  - psychiatry (DSM, APA)
  - adverse reactions (COSTART, WHO ART)
  - primary care (ICPC)

- Terminology of knowledge bases (AI/Rheum, DXplain, QMR)

The UMLS serves as a vehicle for the regulatory standards (HIPAA, CHI)
Integrating subdomains

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

Addison's disease (363732003)

Clinical repositories

Genetic knowledge bases

Biomedical literature

Addison Disease (D000224)

Symptoms

Other subdomains

Model organisms

NCBI Taxonomy

FMA

GO

Anatomy

Genome annotations

UMLS C0001403

SNOMED CT
Addison’s Disease: Concept

An adrenal disease characterized by the progressive destruction of the adrenal cortex, resulting in insufficient production of aldosterone and hydrocortisone. Clinical symptoms include anorexia; nausea; weight loss; muscle weakness; and hyperpigmentation of the skin due to increase in circulating levels of ACTH precursor hormone which stimulates melanocytes.
**Metathesaurus Concepts**

- **Concept (~ 1.4M) CUI**
  - Set of synonymous concept names

- **Term (~ 5.3 M) LUI**
  - Set of normalized names

- **String (~ 5.9M) SUI**
  - Distinct concept name

- **Atom (~ 7.2M) AUI**
  - Concept name in a given source

<table>
<thead>
<tr>
<th>Concept ID</th>
<th>Concept Name</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>A0066000</td>
<td>Headache</td>
<td>MeSH</td>
</tr>
<tr>
<td>A0066007</td>
<td>Headaches</td>
<td>MedDRA</td>
</tr>
<tr>
<td>A0066000</td>
<td>Headache</td>
<td>ICD-10</td>
</tr>
<tr>
<td>A12003304</td>
<td>Headaches</td>
<td>OMIM</td>
</tr>
<tr>
<td>A0540936</td>
<td>Cephalodynia</td>
<td>MeSH</td>
</tr>
</tbody>
</table>

---

*MedDRA* refers to Medical Dictionary for Regulatory Activities, *OMIM* refers to Online Mendelian Inheritance in Man, and *MeSH* refers to Medical Subject Headings.
<table>
<thead>
<tr>
<th>Term</th>
<th>Concept</th>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>L0001403</td>
<td></td>
<td></td>
<td>[Addison's disease, Addison's Disease, Addison Disease, Addisons Disease, Disease, Addison]</td>
</tr>
<tr>
<td>L0494940</td>
<td></td>
<td>S5907336</td>
<td>Primary Adrenocortical Insufficiency, Insufficiencies, Primary Adrenocortical</td>
</tr>
<tr>
<td>L0494851</td>
<td></td>
<td>S5907334</td>
<td>Primary Adrenal Insufficiency, Adrenal Insufficiency, Primary</td>
</tr>
<tr>
<td>L0585243</td>
<td></td>
<td>S5907343</td>
<td>Primary Hypoadrenalism, Primary hypoadrenalism</td>
</tr>
<tr>
<td>L3541031</td>
<td></td>
<td>S4115514</td>
<td>primary; hypoadrenocorticism, hypoadrenocorticism; primary</td>
</tr>
<tr>
<td>L1229627</td>
<td></td>
<td>S1471573</td>
<td>Addison-Krankheit, GER</td>
</tr>
<tr>
<td>L5345155</td>
<td></td>
<td>S6107160</td>
<td>Maladie d'Addison, FRE</td>
</tr>
</tbody>
</table>
Beyond concepts  Descriptor level

- In some vocabularies, the unit of information is an aggregate of concepts
  - e.g., descriptors in MeSH, for indexing purposes

<table>
<thead>
<tr>
<th>MeSH Heading</th>
<th>Hydrocortisone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tree Number</td>
<td>D04.808.745.745.654.600</td>
</tr>
<tr>
<td>Tree Number</td>
<td>D06.472.040.585.353.476</td>
</tr>
<tr>
<td>Tree Number</td>
<td>D06.472.040.585.478.392</td>
</tr>
<tr>
<td>Scope Note</td>
<td>The main glucocorticoid secreted by the ADRENAL CORTEX. Its synthetic counterpart is used, either as an injection or tablet, for the treatment of inflammation, allergies, collagen diseases, asthma, adrenocortical deficiency, shock, and some neoplastic conditions.</td>
</tr>
<tr>
<td>Entry Term</td>
<td>11-Epicortisol</td>
</tr>
<tr>
<td></td>
<td>Cortifair</td>
</tr>
<tr>
<td></td>
<td>Cortisol</td>
</tr>
<tr>
<td></td>
<td>Cortril</td>
</tr>
<tr>
<td></td>
<td>Epicortisol</td>
</tr>
<tr>
<td></td>
<td>Hydrocortisone, (11 alpha)-Isomer</td>
</tr>
<tr>
<td></td>
<td>Hydrocortisone, (9 beta, 10 alpha, 11 alpha)-Isomer</td>
</tr>
</tbody>
</table>

**MeSH SUBJECT HEADINGS**
### Beyond concepts  Descriptor level

<table>
<thead>
<tr>
<th>Code</th>
<th>Name</th>
<th>MeSH</th>
</tr>
</thead>
<tbody>
<tr>
<td>A0070119</td>
<td>Hydrocortisone</td>
<td>(MeSH)</td>
</tr>
<tr>
<td>A0043102</td>
<td>Cortisol</td>
<td>(MeSH)</td>
</tr>
<tr>
<td>A0066000</td>
<td>Epicortisol</td>
<td>(MeSH)</td>
</tr>
<tr>
<td>A0016625</td>
<td>11-Epicortisol</td>
<td>(MeSH)</td>
</tr>
<tr>
<td>A0055118</td>
<td>Hydrocortisone, (11 alpha)-Isomer</td>
<td>(MeSH)</td>
</tr>
<tr>
<td>A7801724</td>
<td>Hydrocortisone, (9 beta,10 alpha,11 alpha)-Isomer</td>
<td>(MeSH)</td>
</tr>
<tr>
<td>A7757592</td>
<td>Cortifair</td>
<td>(MeSH)</td>
</tr>
<tr>
<td>A7757595</td>
<td>Cortril</td>
<td>(MeSH)</td>
</tr>
<tr>
<td></td>
<td>D006854</td>
<td></td>
</tr>
</tbody>
</table>
Metathesaurus  Evolution over time

- Concepts never die (in principle)
  - CUIs are permanent identifiers
- What happens when they do die (in reality)?
  - Concepts can merge or split
  - Resulting in new concepts and deletions
Metathesaurus Relationships

- Symbolic relations: ~8 M pairs of concepts
- Statistical relations: ~6 M pairs of concepts (co-occurring concepts)
- Mapping relations: ~150,000

- Categorization: Relationships between concepts and semantic types from the Semantic Network
Symbolic relations

◆ Relation
  ● Pair of “atom” identifiers
  ● Type
  ● Attribute (if any)
  ● List of sources (for type and attribute)

◆ Semantics of the relationship:
  defined by its type [and attribute]

Source transparency: the information is recorded at the “atom” level
Symbolic relationships

<table>
<thead>
<tr>
<th>Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hierarchical</td>
<td></td>
</tr>
<tr>
<td>- Parent / Child</td>
<td>PAR/CHD</td>
</tr>
<tr>
<td>- Broader / Narrower than</td>
<td>RB/RN</td>
</tr>
<tr>
<td>Derived from hierarchies</td>
<td></td>
</tr>
<tr>
<td>- Siblings (children of parents)</td>
<td>SIB</td>
</tr>
<tr>
<td>Associative</td>
<td></td>
</tr>
<tr>
<td>- Other</td>
<td>RO</td>
</tr>
<tr>
<td>Various flavors of near-synonymy</td>
<td></td>
</tr>
<tr>
<td>- Similar</td>
<td>RL</td>
</tr>
<tr>
<td>- Source asserted synonymy</td>
<td>SY</td>
</tr>
<tr>
<td>- Possible synonymy</td>
<td>RQ</td>
</tr>
</tbody>
</table>
Symbolic relationships

- **Hierarchical**
  - isa (is-a-kind-of)
  - part-of

- **Associative**
  - location-of
  - caused-by
  - treats
  - ...

- **Cross-references (mapping)**
Mapping relations

◆ Simple mappings
  ● `<atom 1> mapped_to <atom 2>`
  ● e.g.,
    ■ SNOMED CT to ICD-9-CM

◆ Complex mappings
  ● `<atom 1> mapped_to <boolean expression>`
  ● e.g.,
    ■ ICD-9-CM to MeSH (search strategies)

NB: partially redundant with relations in MRREL
Everything else

- **Co-occurrence information (MRCOC)**
  - Co-occurrence of MeSH descriptors in MEDLINE for the most part

- **Source-specific attributes (MRSAT)**
  - Legacy identifiers, external cross-references
    - SNOMED International legacy codes (SNOMED CT)
    - RxNorm to NDC
  - Concept status in a particular source (SNOMED CT)
  - Frequency of occurrence in MEDLINE (MeSH)
  - MedlinePlus URL (MeSH)
  - ...
UMLS Semantic Network
Semantic Network

- Semantic types (135)
  - tree structure
  - 2 major hierarchies
    - Entity
      - Physical Object
      - Conceptual Entity
    - Event
      - Activity
      - Phenomenon or Process
Semantic Network

● Semantic network relationships (54)
  ● hierarchical (isa = is a kind of)
    ■ among types
      – Animal isa Organism
      – Enzyme isa Biologically Active Substance
    ■ among relations
      – treats isa affects
  ● non-hierarchical
    ■ Sign or Symptom diagnoses Pathologic Function
    ■ Pharmacologic Substance treats Pathologic Function
“Biologic Function” hierarchy (isa)

- Biologic Function
  - Physiologic Function
    - Organism Function
    - Organ or Tissue Function
    - Cell Function
    - Molecular Function
      - Genetic Function
  - Pathologic Function
    - Cell or Molecular Dysfunction
    - Disease or Syndrome
      - Mental or Behavioral Dysfunction
      - Neoplastic Process
    - Experimental Model of Disease
Associative (non-isa) relationships

- Organism
- Biologic Function
- Finding
- Laboratory or Test Result
- Sign or Symptom
- Physiologic Function
- Pathologic Function
- Injury or Poisoning
- Body Location or Region
- Body Space or Junction
- Gene or Genome
- Cell Component
- Cell
- Tissue
- Body Part, Organ or Organ Component
- Body System
- Congenital Abnormality
- Acquired Abnormality
- Anatomical Attribute
- Anatomical Structure
- Embryonic Structure
- Fully Formed Anatomical Structure
- concept of
- property of
- evaluation of
- process of
- disrupts
- co-occurs with
- adjacent to
- location of
- contains, produces

NLM
Why a semantic network?

- Semantic Types serve as high level categories assigned to Metathesaurus concepts, independently of their position in a hierarchy.

- A relationship between 2 Semantic Types (ST) is a possible link between 2 concepts that have been assigned to those STs.
  - The relationship may or may not hold at the concept level.
  - Other relationships may apply at the concept level.
Relationships can inherit semantics
SPECIALIST Lexicon and lexical tools
SPECIALIST Lexicon

◆ Content
  ● English lexicon
  ● Many words from the biomedical domain

◆ 360,000 lexical items

◆ Word properties
  ● morphology
  ● orthography
  ● syntax

◆ Used by the lexical tools
Morphology

◆ Inflection
  - noun  nucleus, nuclei
  - verb  cauterize, cauterizes, cauterized, cauterizing
  - adjective  red, redder, reddest

◆ Derivation
  - verb ⇄ noun  cauterize -- cauterization
  - adjective ⇄ noun  red -- redness
Orthography

◆ Spelling variants

- oe/e  oesophagus - esophagus
- ae/e  anaemia - anemia
- ise/ize cauterise - cauterize
- genitive mark Addison's disease
  Addison disease
  Addisons disease
Syntax

◆ Complementation
  ● verbs
    ▪ intransitive  I'll treat.
    ▪ transitive  He treated the patient.
    ▪ ditransitive  He treated the patient with a drug.
  ● nouns
    ▪ prepositional phrase

Valve of coronary sinus

◆ Position for adjectives
Lexical tools

- To manage lexical variation in biomedical terminologies

- Major tools
  - Normalization
  - Indexes
  - Lexical Variant Generation program (lvgl)

- Based on the SPECIALIST Lexicon

- Used by noun phrase extractors, search engines
Normalization

- Remove genitive: Hodgkin’s diseases, NOS
- Remove stop words: Hodgkin diseases, NOS
- Lowercase: Hodgkin diseases,
- Strip punctuation: hodgkin diseases,
- Uninflect: hodgkin diseases
- Sort words: hodgkin disease

Hodgkin diseases, NOS

Hodgkin diseases,

hodgkin diseases,

hodgkin diseases

hodgkin disease

disease hodgkin
Normalization: Example

Hodgkin Disease
HODGKINS DISEASE
Hodgkin's Disease
Disease, Hodgkin's
Hodgkin's, disease
HODGKIN'S DISEASE
Hodgkin's disease
Hodgkins Disease
Hodgkin's disease NOS
Hodgkin's disease, NOS
Disease, Hodgkins
Diseases, Hodgkins
Hodgkins Diseases
Hodgkins disease
hodgkin's disease
Disease, Hodgkin

normalize
disease hodgkin
Normalization Applications

- Model for lexical resemblance
- Help find lexical variants for a term
  - Terms that normalize the same usually share the same LUI
- Help find candidates to synonymy among terms
- Help map input terms to UMLS concepts
Indexes

- **Word index**
  - word to Metathesaurus strings
  - one word index per language

- **Normalized word index**
  - normalized word to Metathesaurus strings
  - English only

- **Normalized string index**
  - normalized term to Metathesaurus strings
  - English only
Lexical Variant Generation program

- Tool for specialists (linguists)
- Performs atomic lexical transformations
  - generating inflectional variants
  - lowercase
  - ...
- Performs sequences of atomic transformations
  - a specialized sequence of transformations provides the normalized form of a term (the norm program)
Part II

How to use the UMLS?
Outline

◆ Part II: *How to use the UMLS?*

- Obtaining a license
- Remote access
  - Knowledge Source Server (UMLSKS)
  - UMLSKS Application programming interface (API)
- Local installation and customization (MetamorphoSys)
- A UMLS-based algorithm: *Restrict to MeSH*
- Benefits and limitations
Part II
How to use the UMLS?

(1) Obtaining a license
First step  License agreement

◆ Online Web-based license:


- Read license
- Read appendix 1 and 2
- Print a copy for your records
- Complete the Web form

- Verify:
  - receive e-mail from NLM; go to Web site within 72 hours and enter first and last name
- NLM official will countersign (turn-around time of a few days)
- Receive 2nd e-mail from NLM with new license number
This license contains the SNOMED CT® Affiliate Licence Agreement as Appendix 2. Users must accept this page, the contents of Appendix 1, and the contents of Appendix 2 to receive a UMLS license code.

License Agreement for Use of the UMLS® Metathesaurus®

This Agreement is made by and between the National Library of Medicine, Department of Health and Human Services (hereinafter referred to as "NLM") and the LICENSEE.

WHEREAS, the NLM was established by statute in order to assist the advancement of medical and related sciences, and to aid the dissemination and exchange of scientific and other information important to the progress of medicine and to the public health, (section 465 of the Public Health Service Act, as amended (42 U.S.C. section 286) and to carry out this purpose has been authorized to develop the Unified Medical Language System® (UMLS) to facilitate the retrieval and integration of machine-readable biomedical information from disparate sources;

WHEREAS, the NLM's UMLS project has produced the UMLS Metathesaurus, a machine-readable vocabulary knowledge source, that is useful in a variety of settings;

WHEREAS, the LICENSEE is willing to use the UMLS Metathesaurus at its sole risk and at no expense to NLM, which will result in information useful to NLM, may provide immediate improvements in biomedical information transfer to segments of the biomedical community, and is consistent with NLM's statutory functions,

NOW THEREFORE, it is mutually agreed as follows:
1. The NLM hereby grants a nonexclusive, non-transferable right to LICENSEE to use the UMLS Metathesaurus and incorporate its content in any computer applications or systems designed to improve access to biomedical information of any type subject to the restrictions in other provisions of this Agreement. The names and addresses of licensees authorized to use the UMLS products are public information.

2. No charges, usage fees or royalties will be paid to NLM.

3. LICENSEE is prohibited from distributing the UMLS Metathesaurus or subsets of it, including individual vocabulary sources within the Metathesaurus, except (a) as an integral part of computer applications developed by LICENSEE for a purpose other than redistribution of vocabulary sources contained in the UMLS Metathesaurus and (b) if permitted by paragraph 12 of this agreement.

4. LICENSEE agrees to inform NLM prior to distributing any application(s) in which it is using the UMLS Metathesaurus and is encouraged to inform NLM of any difficulties encountered in using the UMLS Metathesaurus, and changes or enhancements to the UMLS Metathesaurus that would make it more useful to LICENSEE and its user groups.

5. Within 30 days of the end of any calendar year in which LICENSEE makes use of the UMLS Metathesaurus, LICENSEE agrees to provide NLM with a brief report on the usefulness of the UMLS Metathesaurus in general and, if applicable, on the usefulness of CPT format in particular. LICENSEE is strongly encouraged to submit to NLM locally developed extensions to the UMLS Metathesaurus potentially useful to other UMLS users for consideration for potential inclusion in the UMLS Metathesaurus.

6. NLM represents that the data provided under this Agreement were formatted with a reasonable standard of care, but makes no warranties express or implied, including no warranty of merchantability or fitness for particular purpose, regarding the accuracy or completeness of the data or that the machine-readable copy is error free. Therefore, LICENSEE agrees to hold NLM, the Government, and any organization contributing a vocabulary source to the UMLS Metathesaurus free from any liability resulting from errors in terminology or other data or on the machine-readable copy. NLM and such other organizations disclaim any liability for any consequences due to use, misuse, or interpretation of information contained or not contained in the UMLS Metathesaurus.

7. NLM represents that its ability to continue to include certain vocabulary sources within the UMLS Metathesaurus is dependent on continuing contractual relations or agreements with the copyright holders for these vocabulary sources. Therefore, LICENSEE agrees to hold NLM and the individual copyright holder free from any liability resulting from the removal of any vocabulary source from future editions of the UMLS Metathesaurus.

8. NLM reserves the right to change the type and format of its machine-readable data. NLM agrees to inform LICENSEE of any changes to the format of the UMLS Metathesaurus, EXCEPT the addition of entirely new data elements to the Metathesaurus, at least 90 days before the data are distributed.

9. The presence in the UMLS Metathesaurus of vocabulary or data produced by organizations other than NLM does not imply any endorsement of the UMLS Metathesaurus by these organizations.

10. LICENSEE shall acknowledge NLM as its source of the UMLS Metathesaurus, citing the year and version number, in a suitable and customary manner but may not in any way indicate or imply that NLM or any of the organizations whose vocabulary sources are included in the UMLS has endorsed LICENSEE or its products.
11. Some of the Material in the UMLS Metathesaurus is from copyrighted sources. If LICENSEE uses any material from copyrighted sources from the UMLS Metathesaurus:

   a) the LICENSEE is required to display in full, prior to providing user access to the Metathesaurus or any of the vocabulary sources within the UMLS, the following wording in order that its users be made aware of these copyright constraints:

   "Some material in the UMLS Metathesaurus is from copyrighted sources of the respective copyright holders. Users of the UMLS Metathesaurus are solely responsible for compliance with any copyright, patent or trademark restrictions and are referred to the copyright, patent or trademark notices appearing in the original sources, all of which are hereby incorporated by reference."

   to display a list of all of the vocabularies contained within the UMLS Metathesaurus that are used in the LICENSEE's application, and to indicate for each vocabulary any appropriate copyright notice and whether the entire contents is present or only a portion of it.

   b) the LICENSEE is prohibited from altering UMLS and other vocabulary source content contained within the UMLS Metathesaurus, but may include content from other sources in applications that also contain content from the UMLS Metathesaurus. The LICENSEE may not imply in any way that data from other sources is part of the UMLS Metathesaurus or of any

   c) the LICENSEE is required to include in its applications identifiers from the UMLS Metathesaurus such that the original source vocabularies for any data obtained from the UMLS Metathesaurus can be determined by reference to a complete version of the UMLS Metathesaurus.

12. For material in the UMLS Metathesaurus obtained from some sources additional restrictions on LICENSEE's use may apply. The categories of additional restrictions are described below. The list of UMLS Metathesaurus Vocabulary Sources, which is part of this Agreement and is updated when each version of the Metathesaurus is released, indicates the category of additional restrictions, if any, that apply to each vocabulary source.

   LICENSEE should contact the copyright holder directly to discuss uses of a source vocabulary beyond those allowed under this license agreement. If LICENSEE or LICENSEE's end user has a separate agreement with the copyright holder for use of a UMLS Metathesaurus source vocabulary, LICENSEE or LICENSEE's end user may use vocabulary source content obtained from the UMLS Metathesaurus in accordance with the terms of the separate agreement.
12. 1. **Category 1:**

LICENSEE is prohibited from translating the vocabulary source into another language or from producing other derivative works based on this single vocabulary source.

12. 2. **Category 2:**

All category 1 restrictions AND

LICENSEE is prohibited from using the vocabulary source in operational applications that create records or information containing data from the vocabulary source. Use for data creation research or product development is allowed.

12. 3. **Category 3:**

LICENSEE's right to use material from the source vocabulary is restricted to internal use at the LICENSEE's site(s) for research, product development, and statistical analysis only. Internal use includes use by employees, faculty, and students of a single institution at multiple sites. Notwithstanding the foregoing, use by students is limited to doing research under the direct supervision of faculty. Internal research, product development, and statistical analysis use expressly excludes: use of material from these copyrighted sources in routine patient data creation; incorporation of material from these copyrighted sources in any publicly accessible computer-based information system or public electronic bulletin board including the Internet, publishing or translating or creating derivative works from material from these copyrighted sources; selling, leasing, licensing, or otherwise making available material from these copyrighted works to any unauthorized party, and copying for any purpose except for back up or archival purposes.

LICENSEE may be required to display special copyright, patent and/or trademark notices before displaying content from the vocabulary source. Applicable notices are included in the list of UMLS Metathesaurus Vocabulary sources, that is part of this Agreement.

12. 4. **Category 4:**
12.4. Category 4:

12.4.1. LICENSEE is prohibited from translating the vocabulary source into another language or from altering the vocabulary source content.

12.4.2. LICENSEE's right to use the vocabulary source is restricted to use in the U.S. by LICENSEE's employees, contractors, faculty, students, clients, patients, or constituents within electronic systems or devices built, purchased, licensed, or used by LICENSEE for U.S. governmental purposes or for any health care, public health, research, educational, or statistical use in the U.S. Use by students is limited to research or educational activities under the direct supervision of faculty.

12.4.3. LICENSEE has the right to distribute the vocabulary source in the U.S., but only in combination with other UMLS Metathesaurus content. Further, LICENSEE's right to distribute is restricted to:

a. Electronic distribution to LICENSEE's direct U.S. affiliates, or to other U.S. entities that have signed the UMLS license, in order to facilitate use of the vocabulary for health care, public health, research, educational or statistical purposes in the U.S. only.

   i. LICENSEE must take reasonable precautions to prevent distribution of the vocabulary source to non-US entities.

   ii. LICENSEE must include in its annual report a list of all U.S. affiliates or other U.S. entities to whom it has distributed content from the vocabulary source.

b. Distribution of encoded patient level data sets or knowledge encoded in the vocabulary source by LICENSEE to any U.S. entity for use in the U.S. only.

c. Inclusion of encoded records or content from the vocabulary source in: (1) free publicly accessible retrieval systems or (2) fee-based retrieval systems that are accessible within the U.S. only, provided that these systems do not permit users to copy or extract any significant portion of the vocabulary source.

12.4.4. DEFINITIONS

a. U.S. is defined as all U.S. states, territories, and the District of Columbia; any U.S. government facility or office, whether permanent or temporary, wherever located; and access to a system in any of these locations by U.S. government employees, designated representatives or contractors, wherever located, for U.S. government purposes.

b. U.S. entity is defined as (i) for government entities, an agency or department of the U.S. Government, (ii) for corporations, as a corporation incorporated and operating in the U.S.; and (iii) for other entities as an entity organized under the laws of the U.S.
13. LICENSEE shall take reasonable steps to ensure that anyone who has authorized access to data or vocabulary sources from the UMLS Metathesaurus under this Agreement complies with its provisions.

14. LICENSEE and/or its end users shall be solely responsible for compliance with any copyright or other restrictions on vocabulary sources in the UMLS Metathesaurus; NLM assumes no responsibility or liability associated with the LICENSEE’s (or any of the LICENSEE’s users) use and/or reproduction of copyrighted material, patent or trademark violations. Anyone contemplating reproduction of all or any portion of the UMLS Metathesaurus or any of its vocabulary sources should consult legal counsel.

The holder of a copyright in any vocabulary source shall be a third party beneficiary to this agreement and shall have a right to enforce the agreement against any LICENSEE that violates any provision pertaining to that copyright holder.

15. This Agreement shall be effective until terminated by one of the parties upon 30 days written notice to the other party. LICENSEE’s failure to abide by the terms of the Agreement shall be grounds for its termination. Neither the Government, its employees, or any vocabulary sources contained in the UMLS Metathesaurus shall be liable or responsible to LICENSEE in any manner whatsoever for damages of any nature whatsoever arising from the termination of this Agreement.

16. In the event that any provision of this Agreement is determined to violate any law or is unenforceable, the remainder of the Agreement shall remain in full force and effect.
APPENDIX A.1

Appendix to the License Agreement for Use of the UMLS® Metathesaurus

UMLS METATHESAURUS® SOURCE VOCABULARIES - 2007AB Edition

Sources are listed in order according to the abbreviations used in the UMLS Metathesaurus files. If additional restrictions and notices apply, the category of restrictions and the special notices appear under the name of the source. See the license agreement for an explanation of the categories of restrictions. Many sources publish printed editions and/or other explanatory information that may be essential to understanding the purpose and application of particular sources in data creation and retrieval. Contact information is provided for each source. Please address questions about permissions or license agreements for additional uses not covered by this Agreement, or other inquiries about individual sources, to the appropriate contacts.

NLM is working toward inclusion in the UMLS Metathesaurus of the complete, current edition of most of these vocabulary sources.


Contact: May Cheh, Lister Hill National Center for Biomedical Communications, National Library of Medicine, Building 38A, Room 9E902, 8600 Rockville Pike, Bethesda, MD 20894; e-mail: cheh@nlm.nih.gov.


CATEGORY 3 RESTRICTIONS APPLY

Contact: Alternative Link LLC, 6121 Indian School Road NE, Suite 131, Albuquerque, NM 87110; Phone: (505) 875-0001; Toll Free: (877)621-5465; Fax: (505) 875-0002; e-mail: ail@alternativelink.com

CATEGORY 2 RESTRICTIONS APPLY

Contact: WHO Collaborating Centre for International Drug Monitoring, Stora Target 3, S-753 20 Uppsala, Sweden; fax: +46-18-656080


CATEGORY 2 RESTRICTIONS APPLY

Contact: WHO Collaborating Centre for International Drug Monitoring, Stora Target 3, S-753 20 Uppsala, Sweden; fax: +46-18-656080


CATEGORY 2 RESTRICTIONS APPLY

Contact: WHO Collaborating Centre for International Drug Monitoring, Stora Target 3, S-753 20 Uppsala, Sweden; fax: +46-18-656080


CATEGORY 2 RESTRICTIONS APPLY

Contact: WHO Collaborating Centre for International Drug Monitoring, Stora Target 3, S-753 20 Uppsala, Sweden; fax: +46-18-656080
APPENDIX 2

SNOMED CT® AFFILIATE LICENCE AGREEMENT

IMPORTANT NOTICE – PLEASE READ THE FOLLOWING CAREFULLY

This is a Licence Agreement between (1) The International Health Terminology Standards Development Organisation (forening med begrænset ansvar), an association (foreningen) established under the laws of Denmark, whose principal place of business is at Rued Langgaards Vej 7, 5te A22, DK-2300 Copenhagen S, Denmark (the “Licensor”) and (2) the person or organisation to which the International Release of SNOMED CT (whether on its own or as part of a Member’s National Release of SNOMED CT) is distributed or otherwise made available (the “Licensee”).

By downloading, accessing or using any part of the International Release of SNOMED CT or a Member’s National Release of SNOMED CT, or exercising any rights granted under this Licence Agreement, the Licensee agrees to be bound by the terms of this Licence Agreement.
License Restriction Levels 0-4

◆ Level 0  (79.3%)
  ● unrestricted

◆ Level 1  (5.0%)
  ● negotiate to translate

◆ Level 2  (0.5%)
  ● negotiate to use in health data creation

◆ Level 3  (25.4%)
  ● negotiate to use in production
  ● explicitly prohibited to provide Internet access

◆ Level 4  (21.8%)
  ● SNOMED CT (unrestricted in member countries)

There may be additional restrictions, or separate license fees, associated with usage of specific vocabularies. Read the UMLS License, including the Appendix!
Part II
How to use the UMLS?

(2) Remote access
Remote Access


- Web search interface
- Application Programming Interface (API)

- Coming soon: web services
Knowledge Source Server

*Web search interface*
UMLSKS Web search interface

- Logging in
- Basic searching
- Advanced searching
UMLS Knowledge Source Server Home Page

- Tabs across top access basic searching of 3 Knowledge Sources
- Advanced searching options on right-hand side
Metathesaurus Basic Search

Addison’s disease

- UMLS Release
- Search Term
- UMLS Knowledge Source
Concept Report

Addison’s disease

Concept Name /CUI

Semantic Type(s)

Definition(s)

Synonyms

Concept: Addison’s disease
CUI: C0001405
Semantic Type: Disease or Syndrome

Definition:
disease characterized by hypotension, weight loss, anorexia, weakness, and sometimes a bronze-like melanotic hyperpigmentation of the skin; due to tuberculosis or autoimmune induced disease (hypofunction) of the adrenal glands that results in deficiency of aldosterone and cortisol. (CRISP Thesaurus)

An adrenal disease characterized by the progressive destruction of the ADRENAL CORTEX, resulting in insufficient production of ALDOSTERONE and HYDROCORTISONE. Clinical symptoms include ANOREXIA; NAUSEA; WEIGHT LOSS; MUSCLE WEAKNESS; and HYPERPIGMENTATION of the SKIN due to increase in circulating levels of ACTH precursor hormone which stimulates MELANOCYTES. (MeSH)

Synonyms:
Addison’s disease
Addison, disease or syndrome
Addison’s disease (disorder)
ADDISON DIS
d adrenal cortex; deficiency, primary
Display All

“Display” shows results for selected options

“Display All” shows results for all available options
Metathesaurus Basic Search

Adrenal gland insufficiency

- Specify:
  - UMLS Release
  - Search term

- Algorithm:
  - Search
  - Normalized String
  - Search
  - Normalized Word
  - Suggest Spelling
Basic Concept Report

Adrenal gland insufficiency

Concept: Adrenal gland hypofunction
CUI: C0001623

Semantic Type: Disease or Syndrome

Definition:
Adrenocortical hypofunction includes all conditions in which hormone secretion falls below the requirements of the body. It may be divided into two general categories: (1) those associated with primary failure of the adrenal, and (2) those associated with a secondary failure due to a primary adrenocorticotropin. (Harrison’s Principles of Internal Medicine (MeSH))

Synonyms:
Adrenal gland hypofunction
Adrenal Gland Insufficiency

- **Concept Name/CUI**
- **Semantic Type(s)**
- **Definition(s)**
- **Synonyms, including foreign languages**
- **Relations (broader, narrower, etc.)**
- **Co-occurrence data**
Hierarchies

Concept Report Display All (continued)
About Concept Display All (continued)

Narrower Concepts:

- Addison's disease (National Drug File – Reference Terminology) [Relation: isa]
- Addison's disease (Metathesaurus Names) [Relation: ]
- Hypoaldosteronism (National Drug File – Reference Terminology) [Relation: isa]
- Other adrenal hypofunction NOS (Metathesaurus Names)
- Adrenoleukodystrophy (National Drug File – Reference Terminology) [Relation: isa]

Narrower Concepts:

- Addison's disease (National Drug File – Reference Terminology) [Relation: isa]
- Addison's disease (Metathesaurus Names) [Relation: ]
- Hypoaldosteronism (National Drug File – Reference Terminology) [Relation: isa]

- Adrenal Gland Diseases (Metathesaurus Names) [Relation: ]
- Endocrine Diseases (Metathesaurus Names) [Relation: ]
- Endocrine Diseases (WHO Adverse Reaction Terms) [Relation: ]
- Adrenal cortical hypofunction (SNOMED Clinical Terms)
Concept Report Display All (continued)

Co-occurrence data

Co-occurring MeSH Terms:

- 64 Corticotropin
- 57 Hydrocortisone
- 38 Glucocorticoids
- 31 Esophageal achalasia
- 30 Dog Diseases
- 28 Adrenal Glands

Associated Expressions: None found.
Metathesaurus Advanced Search Options

- Focused Search
- Raw Relational Records
Metathesaurus Advanced Search Feature

Focused Search

- UMLS Release
- Search Term
- Source Vocabularies
- String Criteria
  - Exact Match
  - Normalized string & word
  - Word
  - Truncation (left/right)
  - Approximate Match
- Language
Restricted Source Concept Report

Addison’s Disease

- **UMLS Release:** 2004AB
- **Search Term:** addison’s disease
- **Source Vocabulary:** SNOMED CT
- **String Criteria:** Normalized string
- **Language:** English
Addison’s disease in SNOMED CT
Preferred Term and Code

TTY: Term Type
ID: Source Code Descriptor
Metathesaurus Advanced Search Feature
Relational Record Request

- **UMLS Release**

- **Search Term**

- **UMLS Relational Table**

---

**UMLS Knowledge Source Server (UMLSKS)**

**About the UMLSKS**
- Home
- Overview
- Frequently Asked Questions
- Edit Views/Profile

**Downloads**
- UMLS Knowledge Sources
- Developer's API

**Documentation**
- User's Guide
- Developer’s Guide
- Developer's API Javadoc
- UMLS Documentation Set

**Resources**
- NLP & Lexical Resources
- Semantic Network Resources
- Metathesaurus Resources

---

**Metathesaurus Relational Records Request:**

**Select UMLS Release:**

1) Enter a term or concept unique identifier (CUI):

2) Select the UMLS table whose raw records are to be returned:

Raw record return allows users to search for a concept using either a user entered concept unique identifier (CUI) or a term name which is entered into the text entry box below. Any of the UMLS tables listed in the drop down menu may be selected for return.

Pressing the **Perform Concept Search** button will update the display to show the raw records from the selected table based on the concept information found.
| CUI | LAT | TS | LUI | STT | SUI | ISPREF | AU1 | SAUI | SCUI | SDUI | SAB | TTY | CODE | STR | SUPPRESS | CVF | CZE | P | L3180742 | PF | S3708232 | Y | A3910108 | D000224 | MSHCZE | MH | D000224 | ADDISONOVA | NEMOC | 
| C0001403 | DUT | P | L2048638 | PF | S2386860 | N | A6566810 | 10001130 | MDRDUT | LT | 10001130 | Addison, ziekte | 
| C0001403 | DUT | P | L2048638 | PF | S2386860 | Y | A3931189 | D000224 | MSHDUT | MH | D000224 | Addison, ziekte van | 
| C0001403 | DUT | S | L2048637 | PF | S2386859 | Y | A3931188 | D000224 | MSHDUT | SY | D000224 | Addison, syndroom | 
| C0001403 | DUT | S | L2528364 | PF | S2985131 | Y | A3966882 | D000224 | MSHDUT | SY | D000224 | Ziekte van Addison | 
| C0001403 | DUT | S | L3205108 | PF | S3732602 | N | A5146733 | E27.1 | ICD10DUT | PT | E27.1 | Primaire bijnierschade | 
| C0001403 | DUT | S | L3205108 | PF | S3732602 | Y | A3970882 | D000224 | MSHDUT | SY | D000224 | Primaire bijnierschade | 
| C0001403 | DUT | S | L4999233 | PF | S5686738 | Y | A6627443 | 10036696 | MDRDUT | LT | 10036696 | primair hypoadrenalisme | 
| C0001403 | DUT | S | L4999270 | PF | S5686775 | N | A6627493 | 10052381 | MDRDUT | LT | 10052381 | primaire bijnierschade | 
| C0001403 | DUT | S | L4999270 | PF | S5686775 | Y | A6627494 | 10052381 | MDRDUT | PT | 10052381 | primaire bijnierschade | 
| C0001403 | DUT | S | L5012413 | PF | S5699917 | Y | A6645695 | 10013096 | MDRDUT | LT | 10013096 | ziekte van Addison | 
| C0001403 | ENG | P | L0001403 | PF | S0354372 | N | A0388276 | 10000006 | ADDISON ´S DISEASE | 0 | 
| C0001403 | ENG | P | L0001403 | PF | S0354372 | N | A0388277 | 0060-3321 | CSP | PT | 0060-3321 | Addison ´s disease | 
| C0001403 | ENG | P | L0001403 | PF | S0354372 | N | A0388279 | 10000061 | ADDISON ´S DISEASE | 0 | 
| C0001403 | ENG | P | L0001403 | PF | S0354372 | N | A0388280 | 10001390 | MDR | LT | 10001130 | Addison ´s disease | 
| C0001403 | ENG | P | L0001403 | PF | S0354372 | N | A0388281 | 10000061 | ADDISON ´S DISEASE | 3 |
Semantic Network Searching

- Select Tab along top
- Quick search
- Advanced Search on right-hand side
Semantic Network Search

- Enter search string
- Select semantic type
- Select semantic relation
Semantic Type Clinical Drug

- Browse ST hierarchy
- View Concepts with ST
- View Relations valid for the ST
- View Raw Relational Records
Show Relations Between Types

- Validates whether a selected Semantic Relationship (SR) holds between two selected Semantic Types (ST)
SPECIALIST Lexicon Searching

- Select Tab along top
- Quick search
SPECIALIST Lexicon Search

About the UMLSKS
- Home
- Overview
- Frequently Asked Questions
- Edit Views/Profile

Downloads
- UMLSKS Knowledge Sources
- Developer's API

Documentation
- User's Guide
- Developer's Guide
- Developer's API Javadocs

SPECIALIST Lexicon
The SPECIALIST Lexicon is an English language lexicon containing many biomedical terms. The lexicon entry for each word or term records syntactic, morphological, and orthographic information.

Lexical entries may be single or multi-word terms.
SPECIALIST Lexical Record

Specialist Lexical Record

```plaintext
;base=Addison's disease
entry=E0000160
cat=noun
variants=uncount
variants=reg
```

View "Addison's disease" in relational format.
UMLS Resources

- NLP & Lexical Resources
  - MetaMap Transfer (MMTx)
  - Word Sense Disambiguation (WSD) Test Collection
- Semantic Network
  - Semantic Navigator
  - Semantic Groups
- Metathesaurus
  - String Properties
The “new” UMLSKS (coming soon)
UMLS Semantic Navigator

- Web-based
- Concept- and relation-centric
- Displays contexts graphically
- Displays all relations simultaneously
- Excludes hierarchical cycles in the UMLS graph
- Search
  - By CUI
  - By word
RRF Browser

- Distributed with the UMLS
  - Along with MetamorphoSys
- Standalone
- Can browse particular subsets of the Metathesaurus
- Search
  - By code
  - By CUI
  - By word
**Concept:** [CUI C0001403] Addison Disease

**Semantic Type**
Disease or Syndrome

**Definition**
An adrenal disease characterized by the progressive destruction of the ADRENAL CORTEX, resulting in insufficient production of ALDOSTERONE and HYDROCORTISONE. Clinical symptoms include ANOREXIA; NAUSEA; WEIGHT LOSS; MUSCLE WEAKNESS; and HYPERPIGMENTATION of the SKIN due to increase in circulating levels of ACTH precursor hormone which stimulates MELANOCYTES.

**Atoms (17):** (Sorted by Source, String)
- ADDISON DIS [A12078968/MSH/DEV/D000224] Y
- ADDISONS DIS [A12075312/MSH/DEV/D000224] Y
- Addison Disease [A6954527/MSH/MH/D000224]
- Addison's Disease [A6954528/MSH/EN/D000224]
- Addisons Disease [A0019742/MSH/PM/D000224]
- Adrenal Insufficiency, Primary [A6993206/MSH/PM/D000224]
- Adrenocortical Insufficiencies, Primary [A6993209/MSH/PM/D000224]
- Disease, Addison [A0049628/MSH/PM/D000224]
Knowledge Source Server

Application Programming Interface
UMLSKS API basics

- Remote server at NLM
- Local application connected through TCP/IP socket

### Java RMI
- Java-based applications
- Developer’s Guide: Chapter 3
- Set of Java classes (part of the UMLSKS API download)
- Detailed Javadoc documentation online and with API download

### TCP/IP socket
- XML-based queries
- Developer’s Guide: Chapter 5
- XML schema
- Socket server
  - Host: umlsks.nlm.nih.gov
  - Port: 8042
Developer’s Guide

This guide describes the installation and usage of the programmatic interface for the UMLS Knowledge Source Server (UMLSKS).

**Audience**
The audience for this guide is developers of UMLSKS applications using the UMLSKS API.

**Release Notes**
Please refer to the [Release Bulletin](#) for a detailed list of features, bug fixes, and known problems with this version of the UMLSKS.

**How to Use This Guide**
This manual contains the following chapters:

- **Chapter 1 - Introduction** describes the basic features and architecture of the UMLSKS.
- **Chapter 2 - Installing the UMLSKS** provides administrators instructions on installing and tailoring a UMLSKS installation.
- **Chapter 3 - Building UMLSKS Software Applications** describes the functions available to developers wanting to interface to the UMLSKS through another Java program.
- **Chapter 4 - Using the XML Query Facility** describes how to use the querying facility of the UMLSKS wherein users build XML queries to be executed.
- **Chapter 5 - Using the UMLSKS Socket Server** describes how to use the socket server to pass XML formatted commands or command-line type queries (e.g. ks-meta-c aids) that are to be executed by the server, with the results passed back to a client listening to the socket.
Documentation  Java API

UMLSKS Knowledge Source Server (UMLSKS)

About the UMLSKS
- Home
- Overview
- Frequently Asked Questions
- Edit Views/Profile

Downloads
- UMLSKS Knowledge Sources
- Developer's API

Documentation
- User's Guide
- Developer's Guide
  1. Introduction
  2. Installing the UMLSKS
  3. Building UMLSKS Software Applications
  4. Using the XML Query Facility
  5. Using the UMLSKS Socket Server
- UMLSKS Documentation Set

Developer's Guide

Print/Text Version  Table of Contents  About This Guide

UMLSKS API Download

The following instructions describe the procedures for downloading and installing the UMLSKS API. The sections include:

- Downloading the UMLSKS API
- Building the Example .java Files
- Running the Client
- Running the ExpertClient
- Running the SocketClient
- Running the StandardQueryClient
- Available Documentation
- Sample Output and XML Query Examples

Downloading the UMLSKS API

The complete set of files for the UMLSKS API may be downloaded by clicking on the link below. Clicking the
## Packages

<table>
<thead>
<tr>
<th>Packages</th>
</tr>
</thead>
<tbody>
<tr>
<td>gov.nih.nlm.kss.api</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.model</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.models</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.models.lex</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.models.meta</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.models.meta.assocExp</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.models.meta.attribute</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.models.meta.concept</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.models.meta.context</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.models.meta.cooccurrence</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.models.meta.deltas</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.models.meta.locator</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.models.meta.meshentry</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.models.meta.relation</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.models.meta.source</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.models.sem</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.models.sem.rel</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.models.sem.units</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.queries</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.queries.lex</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.queries.meta</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.queries.sem</td>
</tr>
<tr>
<td>gov.nih.nlm.kss.util</td>
</tr>
<tr>
<td>gov.nih.nlm.nls.lexCheckLib</td>
</tr>
</tbody>
</table>
Sample XML query (1) Current version

```xml
<?xml version="1.0"?>
<getCurrentUMLSVVersion version="1.0"/>

<?xml version="1.0"?>
<CurrentUMLSYear version="1.0">
  2007AB
</CurrentUMLSYear>
```
Sample XML query  (2) Concepts by string

```xml
<?xml version="1.0"?>
<findCUI version="1.0">
<conceptName>appendicectomy</conceptName>
<language>ENG</language>
<exact/>
<noSuppressibles/>
</findCUI>

<?xml version="1.0"?>
<ConceptIdCollection version="1.0">
<release>2004AB</release>
<conceptId>
<cui>C0003611</cui>
<cn>Appendectomy</cn>
</conceptId>
</ConceptIdCollection>
```
Sample XML query (3) Concepts properties

```xml
<?xml version="1.0"?>
<getSemanticType version="1.0">
  <cui>C0033572</cui>
</getSemanticType>

<?xml version="1.0"?>
<SemanticTypeCollection version="1.0">
  <release>2004AB</release>
  <cui>C0033572</cui>
  <cn>Prostate</cn>
  <semanticType>
    <tui>T023</tui>
    <sty>Body Part, Organ, or Organ Component</sty>
  </semanticType>
</SemanticTypeCollection>
```
Sample XML query (4) Relationships

```xml
<?xml version="1.0"?>
<getRelations version="1.0">
  <cui>C0033572</cui>
  <rel>RO</rel>
</getRelations>

<?xml version="1.0"?>
<RelationCollection version="1.0">
  [...]  
  <relation>
    <aui>A3188910</aui>
    <sab>SNOMEDCT</sab>
    <relSource>
      <cui>C0007112</cui>
      <cn>Adenocarcinoma of prostate</cn>
      <aui>A3318222</aui>
      <rel>RO</rel>
      <rui>R54806623</rui>
      <rela>has_finding_site</rela>
    </relSource>
  </relation>
  [...]  
</RelationCollection>
```
Sample XML query (5) All semantic type IDs

```xml
<?xml version="1.0"?>
<listSemTypeIds version="1.0">
</listSemTypeIds>
```

```xml
<?xml version="1.0"?>
<SemNetIdCollection version="1.0">
    <release>2004AB</release>
    <semnetId>
        <name>Acquired Abnormality</name>
        <ui>T020</ui>
        <semtype/>
    </semnetId>
    <semnetId>
        <name>Activity</name>
        <ui>T052</ui>
        <semtype/>
    </semnetId>
    [...]
</SemNetIdCollection>
```
Performing XML queries from UMLSKS

Metathesaurus Advanced Search Options

Perform Focused Search

Focused searching allows users to search for a concept that contains either a user entered term name or a concept unique identifier (CUI). The advanced searching option allows user to restrict the results to a set of source vocabularies and may specify the method of matching the entered term name. When searching for a term name, the user may also specify the criteria to be used in matching the entered string to the UMLSKS contents.

Perform XML Query

The XML query facility allows users to specify a query using the extensible Markup Language (XML) dialect created for the UMLSKS. Users of this searching feature must be familiar with the internal database structures and how tables data are related to each other.

Request ASCII Relational Records

The data tables used to populate the backend Oracle database can be returned to the user that match a user's input term or concept unique identifier.
Performing XML queries from UMLSKS

```xml
<?xml version="1.0"?>
<getRelations>
  <cui>C0033572</cui>
  <rel>RO</rel>
</getRelations>
```
Part II
How to use the UMLS?

(3) Installing the UMLS locally and Customizing the Metathesaurus using MetamorphoSys
What is MetamorphoSys?

- Tool distributed with the UMLS
- Multi-platform Java software

- The UMLS installation and customization wizard
  - Installs Knowledge Sources to local storage
  - Subsets and customizes a local Metathesaurus
Using MetamorphoSys

- Simple to use
- Screens and tabs lead you through process
- Installs NLM data format files to local storage
Why use MetamorphoSys?

Customize the Metathesaurus

- To remove terminology that is unhelpful, or even harmful, to your needs and purposes

- To comply with terms of license agreement
Why use MetamorphoSys?

Changing Default Settings

- To alter the preferred name
- To alter suppressibility of specific source term types
Customization is Critical

- Requires a clear understanding of:
  - Characteristics of source vocabularies
  - License arrangements
  - User’s functional requirements
  - User’s purpose and perspective

- Technical expertise

… and requires a multidisciplinary technical team
Machine Requirements

- A fast CPU – 1 GHz or higher
- 1 GB RAM recommended (512 MB min.)
- 6x (or better) DVD drive
- 22 GB minimum free disk space

- Runs on Sun Solaris 8 & 9, Windows XP, NT, and 2000, Linux, and Mac
- 1-10 hours run time on platforms tested
Download from UMLSKS ...

- High speed Internet connection required
- Read the README file for the release
...or DVD?

◆ Order at: umls_support@nlm.nih.gov
◆ Include your license number

◆ Run MetamorphoSys from DVD
  ● Windows
    ◦ Autorun; or go to root directory and click on “windows_mmsys.bat”
  ● Linux, Solaris, Macintosh
    ◦ open a terminal window, change to the root directory and type appropriate command: ./linux_mmsys.sh, ./solaris_mmsys.sh, ./macintosh_mmsys.sh
Welcome Screen

METAMORPHOSYS
THE UMLS INSTALLATION TOOL

Install UMLS
Browse my subset

Unified Medical Language System®
Install UMLS

INSTALL UMLS

Source: E:\ This directory contains UMLS files.
Destination: D:\ Select where to create the top-level UMLS directory. The directory name will correspond to the UMLS release, e.g., 2007AB.

Install UMLS Knowledge Sources:
- Metathesaurus
- Semantic Network
- SPECIALIST Lexicon & Lexical Tools

Select one or all of these options

OK Cancel
UMLS License Notice

Notice:
The Metathesaurus contains source vocabularies produced by many different copyright holders. The majority of the content of the Metathesaurus is available for use under the basic (and quite open) terms described in the Metathesaurus license http://www.nlm.nih.gov/research/umls/license.html.

However, some vocabulary producers place ADDITIONAL RESTRICTIONS ON THE USE OF THEIR CONTENT AS DISTRIBUTED WITHIN THE METATHESAURUS.

The various levels of additional restrictions are described in Section 12 of the license. The level that applies to individual vocabularies is recorded in the Appendix to the license. If a UMLS user already has a separate license for use of one of the source vocabularies, the user's existing license also applies to that source as distributed within the Metathesaurus. In some cases, UMLS users may have to request permission or negotiate a separate license with a vocabulary producer in order to use that vocabulary in a production system. There may be a charge associated with these separate permissions or license agreements.

Please select 'Accept' or 'Do Not Accept' below after reviewing the license agreement at the URL above.
Installation progress monitor

Install UMLS
The following operations will be performed:
- [x] Configure Metathesaurus Subset
- [ ] Installing Specialist Lexicon & Tools
- [ ] Installing Semantic Network
- [ ] Installing Metathesaurus
  - [x] Initializing CUI List
  - [ ] Subsetting Content
  - [ ] Subsetting Indexes
  - [ ] Final Processes
- [ ] Done

Operation is 12% Complete

150
Select a default subset

Level 0 → no separate additional license agreements

Level 0 + SNOMEDCT → Users from non-IHTSDO member countries must have separate license agreements
Input Options Tab

Select the format in which subset data was saved and indicate the location of the source folder that contains the existing subset configuration. See Help for more information.

Users may proceed to the other Options tabs in any order by clicking on the tabs across the top of the screen or by clicking "Done" on the File bar and selecting "Begin Subset".

Input Format Options

Select Input Format

NLM Data File Format

Browse...

NLM Data File Format Configuration

Source Folder - Location of Metathesaurus Files

E:

Browse...
Output Options Tab

Select data output options to customize the subset and create additional files if desired. See Help for more information.

Users may proceed to the other Options tabs in any order by clicking on the tabs across the top of the screen or by clicking "Done" on the File bar and selecting "Begin Subset".

Output Format Options

Select Output Format:

- Rich Release Format

Rich Release Format Configuration

Subset Folder - Location of Subset Files:

D:\2007AB\META
Output Options Tab

Write Database Load Scripts

- Write Oracle load script.
- Write MySQL load script

Source Abbreviation Format

- Output versioned source abbreviations rather than versionless source abbreviations.

Maximum Field Length

- Truncate long fields to 4000 characters.

Eliminate Extended Unicode Characters

- Remove records containing extended UTF-8 characters.

Remove MTH only concepts

- Remove concepts containing only MTH atoms.

Calculate MD5 values for output files

- Calculate MD5s for output files - writes mmsys.md5 file.
Highlighted rows are excluded from the subset.
Precedence Tab

- Ranks names by types of terms within sources
- Highest ranking name determines the Preferred Name
Suppressibility Tab

Highlighted source term types will be marked as suppressible.
File menu

UMLS Metathesaurus Configuration 2007AB

File menu options:
- Enable/Disable Filter
- Import Filter
- New Configuration
- Open Configuration
- Save Configuration
- Exit

Input Format Options:
- Select Input Format
  - NLM Data File Format
    - Browse...

NLM Data File Format Configuration:
- Source Folder - Location of Metathesaurus Files
  - E:\
Edit menu

Undo Suppressibility action  Ctrl+Z
Options menu

Users may proceed to the other Options tabs in any order by clicking on the tabs across the top of the screen or by clicking "Done" on the File bar and selecting "Begin Subset".

Input Format Options

Select Input Format

NLM Data File Format

NLM Data File Format Configuration

Source Folder - Location of Metathesaurus Files

E:\
Reset menu

- Returns all filters to default selections
- Default selections in “mmsys.prop.default file” in config folder
- mmsys.prop.default contains properties in last run
Done – Begin Subset

Select the format in which subset data was saved and indicate the location of the source folder that contains the existing subset configuration. See Help for more information.

Users may proceed to the other Options tabs in any order by clicking on the tabs across the top of the screen or by clicking "Done" on the File bar and selecting "Begin Subset".

Select Input Format

NLM Data File Format

Source Folder - Location of Metathesaurus Files

E:\
Save configuration for next installation

You have made changes to this configuration. Would you like to save the changes?

Yes  No
How MetamorphoSys Works

◆ Removes all information from relational files in excluded vocabularies
  - atoms, strings, relationships, attributes, mappings, etc.

◆ Applies additional options selected by user
  - such as adding source term suppressibility or altering precedence

◆ Produces a full set of Metathesaurus files
  - relational files with customized data
  - reflecting other user criteria
MetamorphoSys log

MetamorphoSys Version: MMSYS-2007AB-20070621
MetamorphoSys Build Date: 2007_06_21_18_35_25
UMLS Build Date: 2007_06_29_11_50_02
Release Version: 2007AB
Release Date: 20070616
Release Description: Base Release for Spring 2007
Metathesaurus Source paths: E:\
Subsetted Metathesaurus folder: D:\2007AB\META
Configuration file used: D:\2007AB\ob-mesh.properties
Initialize CUI List completed: Fri Jul 27 16:33:00 EDT 2007
Subset Index Files completed: Fri Jul 27 17:36:34 EDT 2007
Subset Release Metadata completed: Fri Jul 27 17:36:56 EDT 2007
Finished at: Fri Jul 27 17:39:02 EDT 2007
Concepts in source: 1436586
Concepts in subset: 278431
Time elapsed: 01:10:26

Details about MetamorphoSys Input Handler
MetamorphoSys log

MetamorphoSys Version: MMSYS-2007AB-20070621
MetamorphoSys Build Date: 2007_06_21_18_35_25
UMLS Build Date: 2007_06_29_11_50_02
Release Version: 2007AB
Release Date: 20070616
Release Description: Base Release for Spring 2007
Metathesaurus Source paths: E:
Subsetted Metathesaurus folder: D:\2007AB\META
Configuration file used: D:\2007AB\ob-mesh.properties
Initialize CUI List completed: Fri Jul 27 16:33:00 EDT 2007
Subset Index Files completed: Fri Jul 27 17:36:34 EDT 2007
Subset Release Metadata completed: Fri Jul 27 17:36:56 EDT 2007
Finished at: Fri Jul 27 17:39:02 EDT 2007
Concepts in source: 1436586
Concepts in subset: 278431
Time elapsed: 01:10:26

Details about MetamorphoSys Input Handler
------------------------------------------
MetamorphoSys Input: NLM Data File Format

Details about MetamorphoSys Output Handler
------------------------------------------
Metathesaurus Output: Rich Release Format
Long fields were not truncated.
Source Abbreviations were written out with a versionless (root) representation.
Fields containing UTF-8 characters were not removed.
UTF-8 BOM characters were not added to output files.
Md5s were not calculated.
Concepts containing only MTH atoms were not removed.

Excluded Sources
AI/RHEUM, 1993
Output directory contents
Part II

How to use the UMLS?

(4) A UMLS-based algorithm
Indexing Initiative

For noun phrases extracted from medical texts, map to UMLS concepts

Then, select from the MeSH vocabulary the concepts that are the most closely related to the original concepts

[Aronson & al., AMIA, 2000]
Restrict to MeSH

- Based on the principle of **semantic locality**
- Use different components of the UMLS
- 4 techniques of increasing aggressiveness
  - Use Synonymy: MRCONSO
  - Use Associated expressions (ATXs): MRATX + MRREL
  - Explore the Ancestors: MRREL + SN
  - Explore the Other related concepts: MRREL + SN

[Bodenreider & al., AMIA, 1998]
Restrict to MeSH Synonymy

- Term mapped to Source concept
- For this concept, is there a synonym term that comes from MeSH? (MRCONSO)
Restrict to MeSH Assoc. expressions

- If not,
- Is there an associated expression (ATX) that describes this concept using a combination of MeSH descriptors? *(MRATX/MRMAP + MRREL)*

Endoscopic removal of intraluminal foreign body from oesophagus without incision

AND

MH/SH

Esophagus surgery Foreign Bodies
Restrict to MeSH Ancestors

- If not, let us build the graph of the ancestors of this concept
  - using parents and broader concepts (MRREL)
  - all the way to the top
  - excluding ancestors whose semantic types are not compatible with those of the source concept (MRSTY)

- From the graph, select the concepts that come from MeSH (MRCONSO)

- Remove those that are ancestors of another concept coming from MeSH
Restrict to MeSH  Other related concepts

- If not, explore the other related concepts (MRREL) whose semantic types are compatible with those of the source concept (MRSTY)
- From those, select the concepts that come from MeSH (MRCONSO)
Restrict to MeSH Example

Vein of neck, NOS

There is a MeSH term in the synonyms of SC

SC is described by a combination of MeSH terms (ATX)

The ancestors of SC contain MeSH terms

MeSH terms from non-hierarchically related concepts

Vein + Neck
Restrict to MeSH Quantitative results

- 86% of UMLS concepts mapped to MeSH (2007)

Other related concepts

Synonymy

12%

23%

54%

11%

Graph of ancestors

Built-in mappings
Restrict to MeSH Qualitative results

◆ Qualitative evaluation
  ● 1,036 concepts extracted from 200 MEDLINE citations
  ● manual review of every mapping or failure

◆ 61% Relevant
  ● Subtotal Gastrectomy ➔ Gastrectomy
  ● Encephalopathy, NOS ➔ Brain Diseases

◆ 28% More or less relevant
  ● Vitamin A measurement ➔ Laboratory Procedure
  ● Swelling, NOS ➔ Symptoms

◆ 11% Non relevant
Part II
How to use the UMLS?

(5) Benefits and Limitations
Benefits
UMLS compared to individual vocabularies

- Broader scope
- Extended coverage
- Finer granularity
- Unique identifier
- Synonymous terms clustered into concepts
- Additional synonyms
- Additional hierarchical relationships
- Semantic categorization
Direct benefits

- Concept categorization
- Information retrieval
  - Synonyms
  - Cross-language features
- Information extraction
  - MetaMap
  - Normalization
- Information visualization
  - Knowledge Source Server
  - Semantic Navigator
  - RRF browser
UMLS as an enabling resource

◆ Examples
  ● Mapping across vocabularies
  ● Semantics of statistical associations
  ● Redundancy in hierarchical relations
Limitations
Limitations

- **Structural inconsistency**
  - Cycles in the graph of hierarchical relations
- **Semantic inconsistency**
  - Between Metathesaurus and Semantic Network
- **Underspecified relationships**
- **Missing relations**
  - Synonymy
  - Hierarchical relations (missing or underspecified)
Structural inconsistency  From trees to graph

- Multiple tree structures combined into a graph structure
- Expected: Directed acyclic graph (DAG)
Structural inconsistency  

Cycles in the UMLS graph

- Anti-infective Agents
- Disinfectants and Cleansers
- Disinfectants
- Disinfectant soap
- Germicidal soap
Structural inconsistency  Issues

◆ Theoretical
  ● Violate the antisymmetry property of partial ordering relations

◆ Practical
  ● Loops in graph traversal
  ● Impossible to perform transitive reduction

[Bodenreider, AMIA 2001]
Acyclicity

“back edge” from a child concept to a parent concept

- Reflexive: 13,000
- Direct: 1800
- Indirect: 120
Semantic inconsistency  A two-level structure
Semantic inconsistency  A limited study

6894 interconcept relationships

- among the 3764 concepts in the semantic neighborhood of “Heart”

ICR and SNR not compatible

ICR = SNR or ICR descendant of SNR

ICR not specified and SNR compatible and multiple

ICR not specified and SNR compatible and unique

Semantic inconsistency Issues

- The UMLS integrates what terminologies represent
- Hierarchies in source vocabularies
  - Often task-driven rather than based on principles
  - Usually suitable for information retrieval
  - Not necessarily suitable for reasoning
- No automatic correction possible
  - Wrong categorization
  - Wrong inter-concept relationship
  - [Wrong semantic network relationship]
Underspecified relationships

- Relationship “attribute” not always present
- Relations used to create hierarchies vs. hierarchical relations
Missing relations  Example

- eczema
- acute eczema
- infantile eczema
- acute infantile eczema
- diseases of the skin and subcutaneous tissues
Missing relations  

Example

diseases of the skin and subcutaneous tissues

eczema

acute eczema

infantile eczema

acute infantile eczema
Missing relations  A limited study

- 28,851 pairs of terms
  - Original SNOMED term
  - Demodified term (found in UMLS)
- Corresponding relationship in the Metathesaurus
  - Hierarchical  in 50% of the cases
  - « Sibling »  in 25% of the cases
  - Missing  in 25% of the cases

[Bodenreider & al., 77A, 2001]
Compensation mechanisms

◆ Examples
  ● Removing cycles from hierarchical relations
    ▪ Using redundancy (number of sources asserting the relation)
    ▪ Using terminological knowledge (e.g., NEC)
  ● Lexically-suggested hyponymic relations
    ▪ Properties of adjectival modification
More limitations

- Semantics of hierarchical relations
- Some missing / wrong relations are hard to detect
- Some relations are present but hard to find
Disease

Endocrine system diseases

Autoimmune diseases

Non-neoplastic endocrine disorder

Non-neoplastic adrenal gland disorder

Adrenal gland hypofunction

Adrenal cortical hypofunction

Adrenal cortex diseases

Tuberculous Addison's disease

Addison's disease due to autoimmunity

Endocrine / nutritional / metabolic disorder

Disorders of other endocrine glands

Other disorders of adrenal gland

Addison's Disease

Addison's diseases due to autoimmunity
Autoimmune Diseases

Addison’s disease

Tuberculcus Addison’s disease

Addison’s disease due to autoimmunity

is generally a
Relations  Missing and difficult to detect

chronic uremia  chronic renal failure
hypertensive renal failure
chronic hypertensive uremia
**Relations**  Existing but difficult to find

- **UMLS**
  - carrier protein
  - iron-binding protein
  - ferritin
  - iron

- **Gene Ontology**
  - ferritin
  - iron ion transport
  - iron

- ferritin *isa* iron transporter
- ferritin *transports* iron

reified “transport” relationship

“transport” relationship

**NLM**
How to address these limitations?

- Description logics
- Natural Language Processing (semantic interpretation of the terms)
- Comparing knowledge sources (alignment, inference)
Summary
UMLS Summary

◆ UMLS = 3 Knowledge Sources
  - Metathesaurus
  - Semantic Network
  - SPECIALIST Lexicon and Lexical Tools

◆ MetamorphoSys
  - installs
  - customizes

◆ UMLSKS
  - remote access
  - resources and documentation
Medical Ontology Research

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

Olivier Bodenreider
Lister Hill National Center for Biomedical Communications
Bethesda, Maryland - USA
Bibliography
References: UMLS home page

- UMLS home page

- UMLS documentation
  - Formerly known as the “Green Book”
  - Now online documentation
References

◆ Short presentation

◆ UMLS as a research project
References

◆ Technical papers

◆ Comprehensive bibliography 1986-96
  
Documentation and Support
UMLS documentation and support

◆ UMLS homepage
  ● links to various UMLS resources
  ● http://www.nlm.nih.gov/research/umls/

◆ UMLSKS homepage
  ● links to the User’s and Developer’s guides
  ● http://umlsks.nlm.nih.gov/

◆ UMLS mailing list
  ● UMLSUSERS-L@LIST.NIH.GOV

◆ Email address for support
  ● custserv@nlm.nih.gov
Appendix

UMLS files in Rich Release Format
## MRCONSO (sample rows 1..5) (2004AB)

<table>
<thead>
<tr>
<th>CUI</th>
<th>LAT</th>
<th>IS</th>
<th>LUI</th>
<th>SIT</th>
<th>SUI</th>
<th>ISREF</th>
<th>AUI</th>
<th>SAUI</th>
<th>SCUI</th>
<th>SDUI</th>
</tr>
</thead>
<tbody>
<tr>
<td>C0001403</td>
<td>ENG</td>
<td>P</td>
<td>L0001403</td>
<td>PF</td>
<td>S0354372</td>
<td>Y</td>
<td>A4367951</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C0001403</td>
<td>ENG</td>
<td>P</td>
<td>L0001403</td>
<td>PF</td>
<td>S0354372</td>
<td>N</td>
<td>A2922421</td>
<td>485624014</td>
<td>363732003</td>
<td></td>
</tr>
<tr>
<td>C0001403</td>
<td>ENG</td>
<td>P</td>
<td>L0001403</td>
<td>VC</td>
<td>S0010794</td>
<td>Y</td>
<td>A0019740</td>
<td></td>
<td>M0000346</td>
<td>D000224</td>
</tr>
<tr>
<td>C0001403</td>
<td>ENG</td>
<td>S</td>
<td>L0494851</td>
<td>PF</td>
<td>S2164152</td>
<td>N</td>
<td>A2018589</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C0001403</td>
<td>FRE</td>
<td>P</td>
<td>L3246333</td>
<td>PF</td>
<td>S3773545</td>
<td>Y</td>
<td>A3996251</td>
<td></td>
<td></td>
<td>D000224</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SAB</th>
<th>ITY</th>
<th>CODE</th>
<th>STR</th>
<th>SRL</th>
<th>SUPPRESS</th>
<th>CVF</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTH</td>
<td>PN</td>
<td>NOCODE</td>
<td>Addison's disease</td>
<td>0</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>SNOMEDCT</td>
<td>PT</td>
<td>363732003</td>
<td>Addison's disease</td>
<td>4</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>MSH</td>
<td>MH</td>
<td>D000224</td>
<td>Addison’s Disease</td>
<td>0</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>MDR</td>
<td>LT</td>
<td>10052381</td>
<td>Primary adrenal insufficiency</td>
<td>3</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>MSHFRE</td>
<td>MH</td>
<td>D000224</td>
<td>Addison, maladie</td>
<td>3</td>
<td>N</td>
<td></td>
</tr>
</tbody>
</table>
### MRCONSO (sample rows 6..10)

(2004AB)

<table>
<thead>
<tr>
<th></th>
<th>CUI</th>
<th>LAT</th>
<th>IS</th>
<th>LUI</th>
<th>SRT</th>
<th>SUI</th>
<th>ISREF</th>
<th>AUI</th>
<th>SAUI</th>
<th>SCUI</th>
<th>SDUI</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>C0001403</td>
<td>FRE</td>
<td>S</td>
<td>L1272481</td>
<td>PF</td>
<td>S1514427</td>
<td>Y</td>
<td>A1464383</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>C0001403</td>
<td>GER</td>
<td>P</td>
<td>L1229627</td>
<td>PF</td>
<td>S1471573</td>
<td>Y</td>
<td>A4030156</td>
<td>D000224</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>C0001403</td>
<td>GER</td>
<td>S</td>
<td>L1239271</td>
<td>PF</td>
<td>S1481217</td>
<td>Y</td>
<td>A4034094</td>
<td>D000224</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>C0001403</td>
<td>JPN</td>
<td>P</td>
<td>L3437833</td>
<td>PF</td>
<td>S3965327</td>
<td>Y</td>
<td>A4264008</td>
<td>D000224</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>C0001403</td>
<td>JPN</td>
<td>S</td>
<td>L3465347</td>
<td>PF</td>
<td>S3992841</td>
<td>Y</td>
<td>A4291522</td>
<td>D000224</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### SAB, SY, CODE, STR, SRL, SUPPRESS, CVF

<table>
<thead>
<tr>
<th></th>
<th>SAB</th>
<th>SY</th>
<th>CODE</th>
<th>STR</th>
<th>SRL</th>
<th>SUPPRESS</th>
<th>CVF</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>WHOFRE</td>
<td>IT</td>
<td>0410</td>
<td>MALADIE D'ADDISON</td>
<td>2</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>MSHGER</td>
<td>MH</td>
<td>D000224</td>
<td>Addison-Krankheit</td>
<td>3</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>MSHGER</td>
<td>SY</td>
<td>D000224</td>
<td>Bronzechautkrankheit</td>
<td>3</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>MSHJPN</td>
<td>MH</td>
<td>D000224</td>
<td>Addison病</td>
<td>3</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>MSHPJN</td>
<td>SY</td>
<td>D000224</td>
<td>副腎性黒皮症</td>
<td>3</td>
<td>N</td>
<td></td>
</tr>
</tbody>
</table>

Appendix - Metathesaurus relational files (RRF)
## MRCONSO (sample rows 11-13) (2004AB)

<table>
<thead>
<tr>
<th>CUI</th>
<th>LAT</th>
<th>ISO</th>
<th>LUI</th>
<th>STF</th>
<th>SUI</th>
<th>ISPREF</th>
<th>AUI</th>
<th>SAUI</th>
<th>SClUI</th>
<th>SDUI</th>
</tr>
</thead>
<tbody>
<tr>
<td>C0001403</td>
<td>POR</td>
<td>P</td>
<td>L3302998</td>
<td>PF</td>
<td>S3831123</td>
<td>N</td>
<td>A6382080</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C0001403</td>
<td>RUS</td>
<td>P</td>
<td>L3336992</td>
<td>PF</td>
<td>S3864473</td>
<td>Y</td>
<td>A4157629</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C0001403</td>
<td>SPA</td>
<td>P</td>
<td>L1226877</td>
<td>PF</td>
<td>S1468823</td>
<td>Y</td>
<td>A1419475</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SAB</th>
<th>LTV</th>
<th>CODE</th>
<th>STR</th>
<th>SRL</th>
<th>SUPPRESS</th>
<th>CVF</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDRPOR</td>
<td>LT</td>
<td>1001130</td>
<td>Doença de Addison</td>
<td>3</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>MSHRUS</td>
<td>MH</td>
<td>D000224</td>
<td>АДДИСОНОВА БОЛЕЗНЬ</td>
<td>3</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>WHOSPA</td>
<td>IT</td>
<td>0410</td>
<td>ADDISON, ENFERMEDAD</td>
<td>3</td>
<td>N</td>
<td></td>
</tr>
</tbody>
</table>

Appendix - Metathesaurus relational files (RRF)
### MRHIER (sample rows) (2004AB)

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CUI</td>
<td>AUI</td>
<td>CXN</td>
<td>PAUI</td>
<td>SAB</td>
</tr>
<tr>
<td>1</td>
<td>C0001403</td>
<td>A0019740</td>
<td>1</td>
<td>A0020270</td>
<td>MSH</td>
</tr>
<tr>
<td>2</td>
<td>C0001403</td>
<td>A0019740</td>
<td>2</td>
<td>A0028022</td>
<td>MSH</td>
</tr>
<tr>
<td>3</td>
<td>C0001403</td>
<td>A0019743</td>
<td>3</td>
<td>A1988358</td>
<td>PSY</td>
</tr>
<tr>
<td>4</td>
<td>C0001403</td>
<td>A2922421</td>
<td>1</td>
<td>A3307650</td>
<td>SNOMEDCT</td>
</tr>
<tr>
<td>5</td>
<td>C0001403</td>
<td>A2922421</td>
<td>2</td>
<td>A3307650</td>
<td>SNOMEDCT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Appendix - Metathesaurus relational files (RRF)

<table>
<thead>
<tr>
<th>PTR</th>
<th>HCD</th>
<th>CVF</th>
</tr>
</thead>
<tbody>
<tr>
<td>A0449751.A1988279.A1988358</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## MRREL (sample rows)

(2004AB)

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CUI1</td>
<td>AUI1</td>
<td>STYPE1</td>
<td>REL</td>
<td>CUI2</td>
<td>AUI2</td>
<td>STYPE2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C0001403</td>
<td>C0001403</td>
<td>CUI</td>
<td>RB</td>
<td>C001621</td>
<td>CUI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C0001403</td>
<td>A0019738</td>
<td>AUI</td>
<td>SY</td>
<td>C0001403</td>
<td>A0049628</td>
<td>AUI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C0001403</td>
<td>A2922421</td>
<td>SCUI</td>
<td>CHD</td>
<td>C0085859</td>
<td>A2977940</td>
<td>SCUI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C0001403</td>
<td>A6326321</td>
<td>SCUI</td>
<td>RO</td>
<td>C0688490</td>
<td>A6339383</td>
<td>SCUI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C0001403</td>
<td>A0019743</td>
<td>AUI</td>
<td>PAR</td>
<td>C0935495</td>
<td>A1988358</td>
<td>AUI</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RELA</td>
<td>RUI</td>
<td>SRUI</td>
<td>SAB</td>
<td>SL</td>
<td>RG</td>
<td>DIR</td>
<td>SUPPRESS</td>
<td>CVF</td>
<td></td>
</tr>
<tr>
<td>R02837989</td>
<td>MTH</td>
<td>N</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R18849683</td>
<td>MSH</td>
<td>MSH</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R19859511</td>
<td>1658795027</td>
<td>SNOMEDCT</td>
<td>SNOMEDCT</td>
<td>0</td>
<td>Y</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R27600039</td>
<td>NDFRT</td>
<td>NDFRT</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R08110401</td>
<td>PSY</td>
<td>PSY</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A disease characterized by hypotension, weight loss, anorexia, weakness, and sometimes a bronze-like melanotic hyperpigmentation of the skin. It is due to tuberculosis- or autoimmune-induced disease (hypofunction) of the adrenal glands that results in deficiency of aldosterone and cortisol. In the absence of replacement therapy, it is usually fatal.

<table>
<thead>
<tr>
<th>CUI</th>
<th>AUI</th>
<th>ATUI</th>
<th>SATUI</th>
<th>SAB</th>
<th>DEF</th>
</tr>
</thead>
<tbody>
<tr>
<td>C0001403</td>
<td>A0019740</td>
<td>AT15061584</td>
<td>MSH</td>
<td></td>
<td>A disease characterized by hypotension, weight loss, anorexia, weakness, and sometimes a bronze-like melanotic hyperpigmentation of the skin. It is due to tuberculosis- or autoimmune-induced disease (hypofunction) of the adrenal glands that results in deficiency of aldosterone and cortisol. In the absence of replacement therapy, it is usually fatal.</td>
</tr>
</tbody>
</table>
## MRSAT (sample rows)

(2004AB)

<table>
<thead>
<tr>
<th>CUI</th>
<th>LUI</th>
<th>SUI</th>
<th>METAUI</th>
<th>STYPE</th>
<th>CODE</th>
</tr>
</thead>
<tbody>
<tr>
<td>C0001403</td>
<td>L0001403</td>
<td>S0010792</td>
<td>A0019738</td>
<td>AUI</td>
<td>D000224</td>
</tr>
<tr>
<td>C0001403</td>
<td>L0001403</td>
<td>S0010794</td>
<td>A6326321</td>
<td>SCUI</td>
<td>C712</td>
</tr>
<tr>
<td>C0001403</td>
<td>L0001403</td>
<td>S0354372</td>
<td>A2922421</td>
<td>SAUI</td>
<td>363732003</td>
</tr>
<tr>
<td>C0001403</td>
<td></td>
<td></td>
<td>R15742591</td>
<td>SRUI</td>
<td></td>
</tr>
<tr>
<td>C0001403</td>
<td></td>
<td></td>
<td>CUI</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ATUI</th>
<th>SATUI</th>
<th>ATN</th>
<th>SAB</th>
<th>ATV</th>
<th>SUPPRESS</th>
<th>CUIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT15321482</td>
<td></td>
<td>DID</td>
<td>MSH</td>
<td>D000224</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>AT33411754</td>
<td></td>
<td>MESH_UI</td>
<td>NDFRT</td>
<td>D000224</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>AT24166602</td>
<td></td>
<td>DESCRIPTION STATUS</td>
<td>SNOMEDCT</td>
<td>0</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>AT27438950</td>
<td></td>
<td>REFINABILITY</td>
<td>SNOMEDCT</td>
<td>0</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>AT02925340</td>
<td></td>
<td>ST</td>
<td>MTH</td>
<td>R</td>
<td>N</td>
<td></td>
</tr>
</tbody>
</table>

**Appendix - Metathesaurus relational files (RRF)**
### Appendix - Metathesaurus relational files (RRF)

<table>
<thead>
<tr>
<th>CUI</th>
<th>TUI</th>
<th>STN</th>
<th>STY</th>
<th>ATUI</th>
<th>CVF</th>
</tr>
</thead>
<tbody>
<tr>
<td>C0001403</td>
<td>T047</td>
<td>B2.2.1.2.1</td>
<td>Disease or Syndrome</td>
<td>AT17683850</td>
<td></td>
</tr>
</tbody>
</table>
### MRHIST (sample rows)

#### (2004AB)

<table>
<thead>
<tr>
<th>CUI</th>
<th>SOURCECUI</th>
<th>SAB</th>
<th>SVER</th>
<th>CHANGETYPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>C0001403</td>
<td>1198962018</td>
<td>SNOMEDCT</td>
<td>20020731</td>
<td>0</td>
</tr>
<tr>
<td>C0001403</td>
<td>1212124016</td>
<td>SNOMEDCT</td>
<td>20020731</td>
<td>0</td>
</tr>
<tr>
<td>C0001403</td>
<td>1490869013</td>
<td>SNOMEDCT</td>
<td>20030131</td>
<td>0</td>
</tr>
<tr>
<td>C0001403</td>
<td>363732003</td>
<td>SNOMEDCT</td>
<td>20020129</td>
<td>0</td>
</tr>
<tr>
<td>C0001403</td>
<td>373662000</td>
<td>SNOMEDCT</td>
<td>20020731</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHANGEKEY</th>
<th>CHANGEVAL</th>
<th>REASON</th>
<th>CVF</th>
</tr>
</thead>
<tbody>
<tr>
<td>DESCRIPTIONSTATUS</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DESCRIPTIONSTATUS</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DESCRIPTIONSTATUS</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CONCEPTSTATUS</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CONCEPTSTATUS</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Appendix - Metathesaurus relational files (RRF)