

One Medicine One Pathology: 2nd annual CASIMIR Symposium
on Human and Mouse Disease Informatics

UMLS and phenotype coding

Anita Burgun, Fleur Mouglin, Olivier Bodenreider

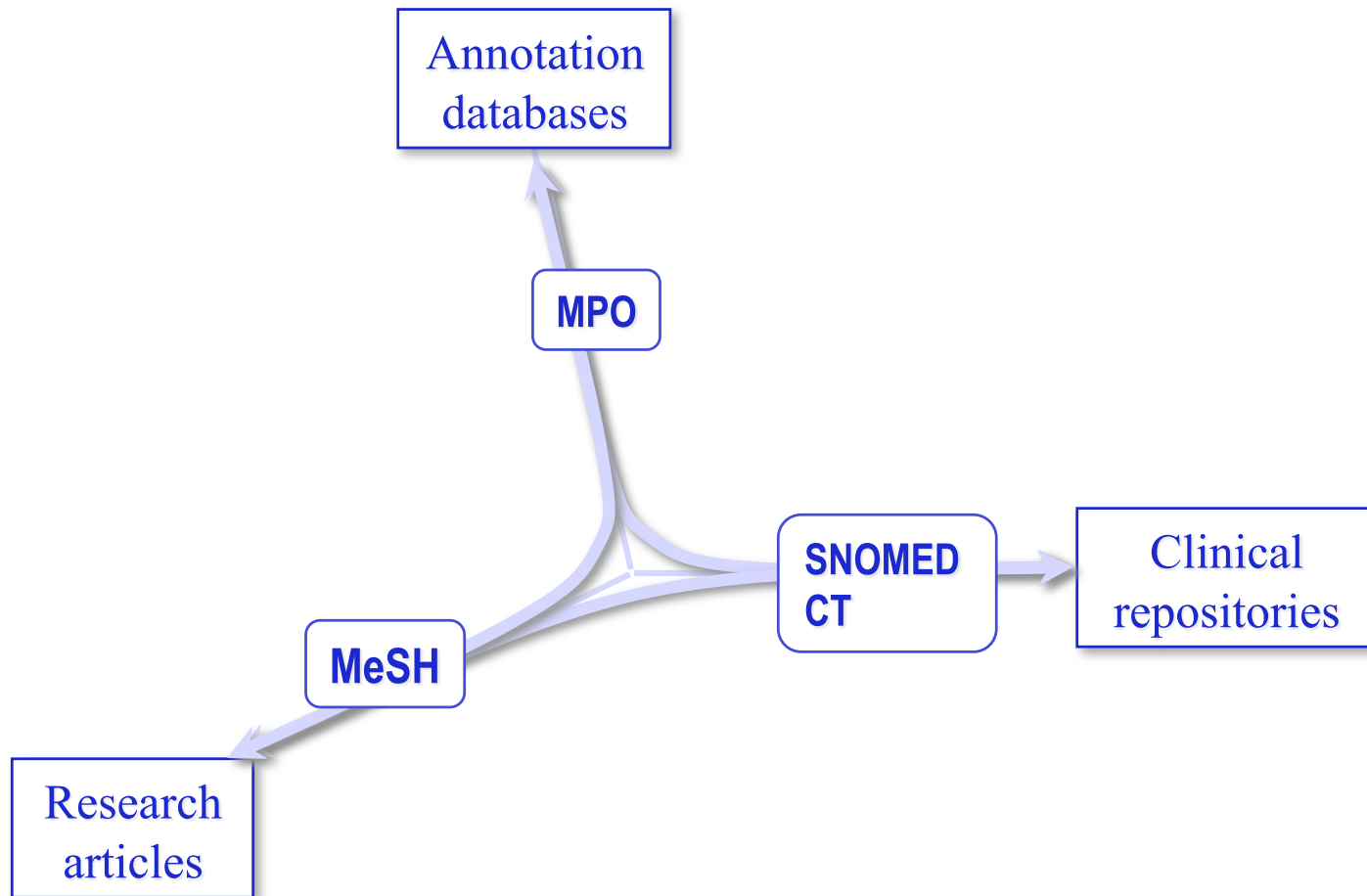
INSERM U936, EA 3888- Faculté de Médecine, Univ. Rennes1

ISPED, Univ. Bordeaux, France

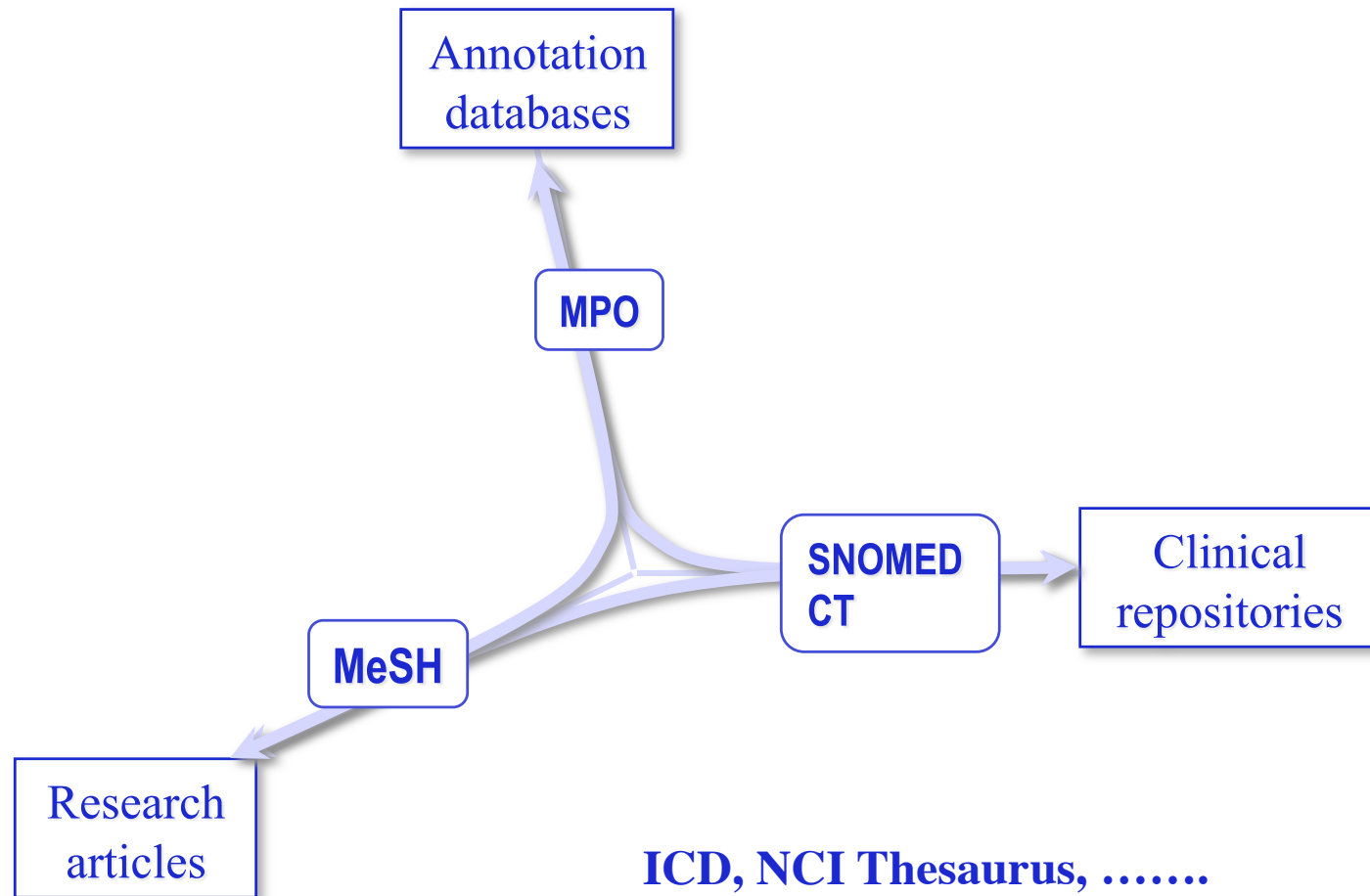
US National Library of Medicine, NIH, Bethesda, USA

December 2, 2008

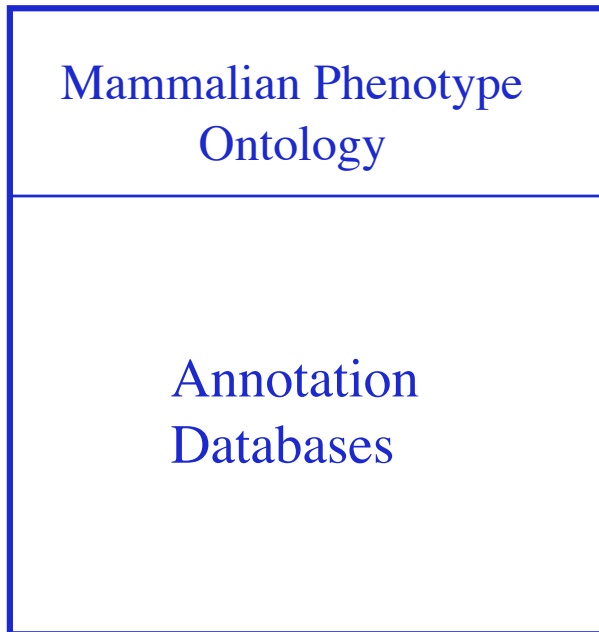
Phenotype coding



Phenotype coding



Phenotype coding: role of the UMLS

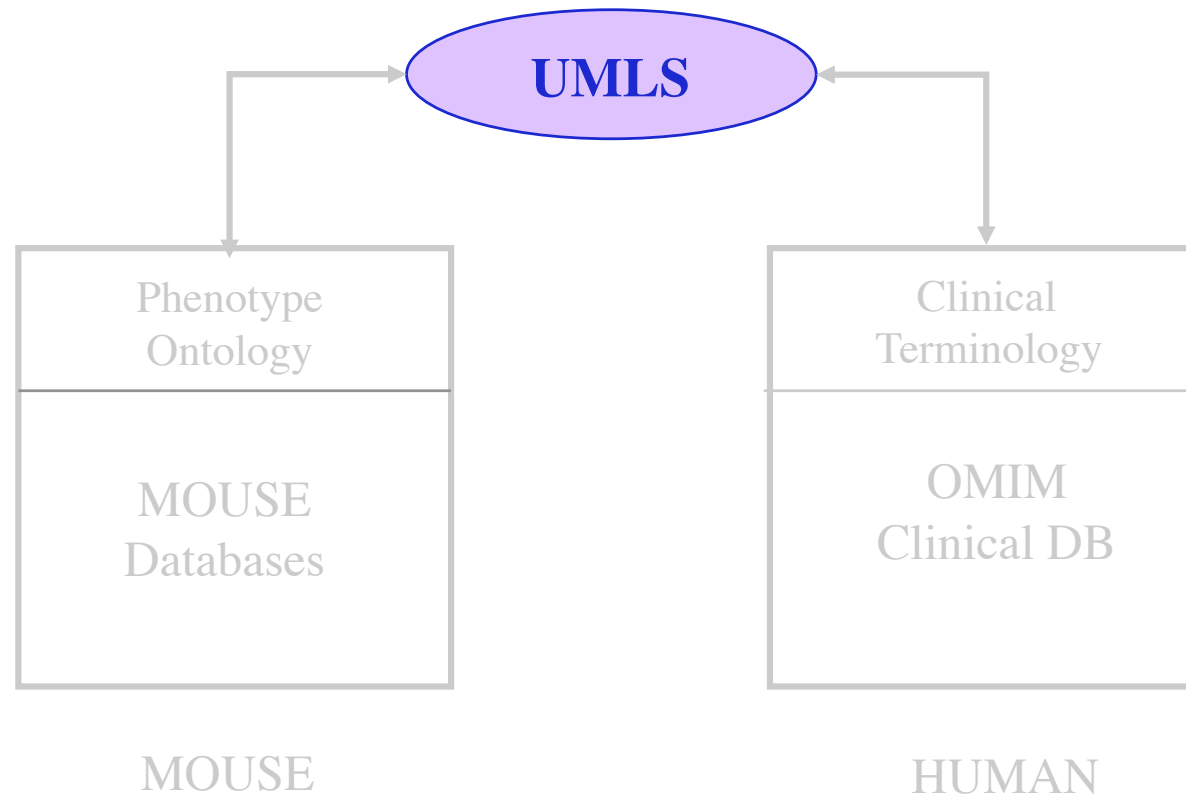


MOUSE



HUMAN

Phenotype coding: role of the UMLS



Unified Medical Language System

- Addresses heterogeneity issues
 - More than 100 source vocabularies
 - Unification
- Clusters terms into concepts
 - Metathesaurus: CUIs
- Organizes hierarchies
 - Metathesaurus: relations
- Categorizes concepts
 - Semantic Network: Semantic Types

Unified Medical Language System

- Addresses heterogeneity issues
 - More than 100 source vocabularies
 - Unification
- Clusters terms into concepts
 - Metathesaurus: CUIs
- Organizes hierarchies
 - Metathesaurus: relations
- Categorizes concepts
 - Semantic Network: Semantic Types

UMLS Metathesaurus

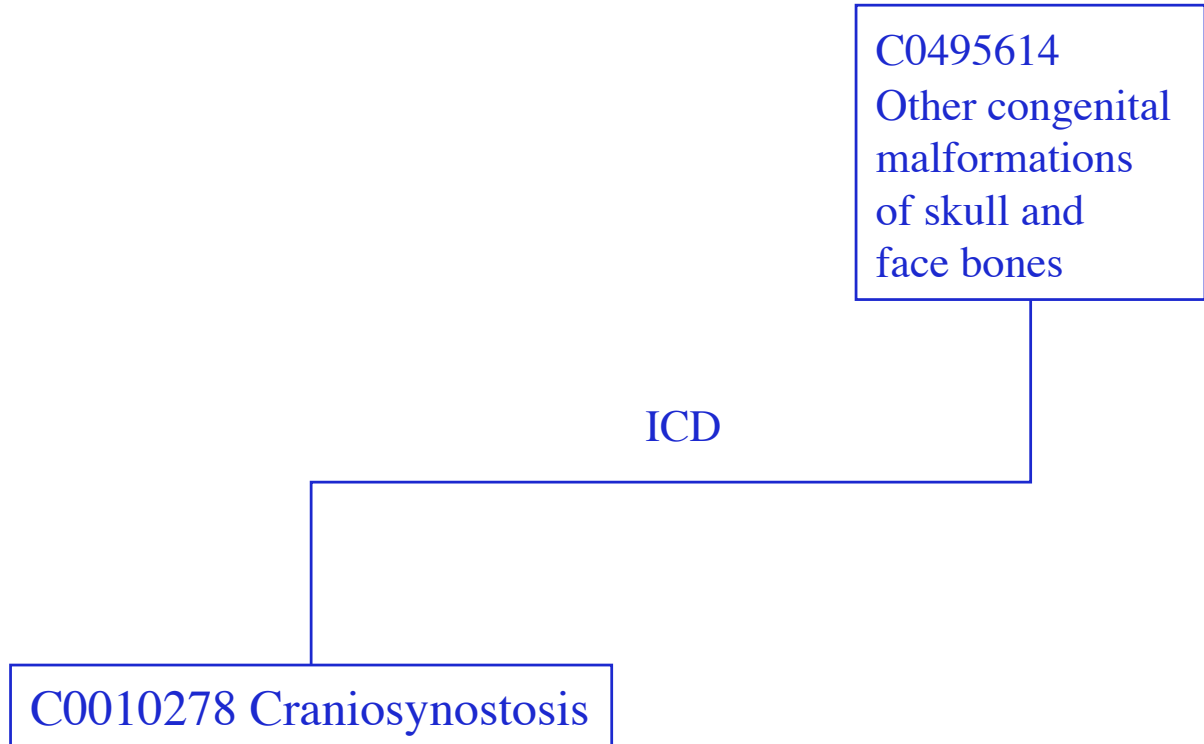
- Craniosynostosis in UMLS Release 2008AA
- Source vocabularies
 - ICD-10
 - ICPC
 - MedDRA
 - MeSH
 - OMIM
 - Read Codes
 - SNOMED CT
 -
- Definition (MeSH) : Premature closure of one or more sutures of the skull

- UMLS Metathesaurus
- CUI C0010278 Craniosynostosis (Preferred Term)
 - Craniostenosis (ICD, ICPC, OMIM, SNOMED CT)
 - Craniosynostosis syndrome (SNOMED CT)
 - Synostosis (cranial) (CRISP)
 - *Word phrases*
 - Premature closure of cranial sutures (MedDRA, SNCT)
 - Congenital ossification of cranial sutures
 - Congenital ossification of sutures
 - Congenital ossification of sutures of skull
 - Premature cranial suture closure (SNOMED)
 - *Abbreviations*
 - CRS, CSO, CRS1 (OMIM)
 - *More specific terms*
 - Craniosynostosis, type 1 (OMIM)
- Possible synonyms and related
 - Hurst syndrome (C0014077)
 - Christian syndrome 1 (C0795794)
 - SCARF (skeletal abnormalities, cutis laxa, craniostenosis, psychomotor retardation, facial abnormalities) syndrome (C0796146).....

Unified Medical Language System

- Addresses heterogeneity issues
 - More than 100 source vocabularies
 - Unification
- Clusters terms into concepts
 - Metathesaurus: CUIs
- **Organizes hierarchies**
 - **Metathesaurus: relations**
- Categorizes concepts
 - Semantic Network: Semantic Types

ICD-10



MedDRA

C0852332
Musculo
skeletal
and
connective
tissue
deformities
of skull,
face
and buccal
cavity

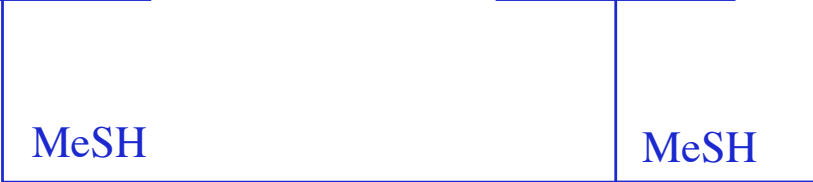
MedDRA

C0010278 Craniosynostosis

MeSH

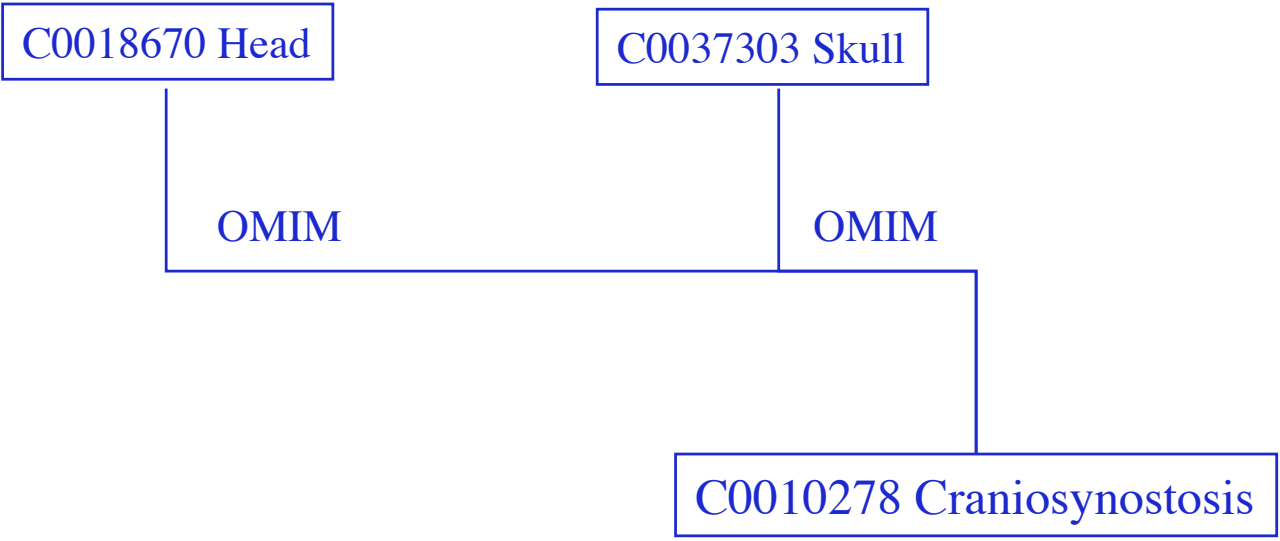
C0376634
Craniofacial
Abnormalities

C0039093
Congenital
abnormal
synostosis

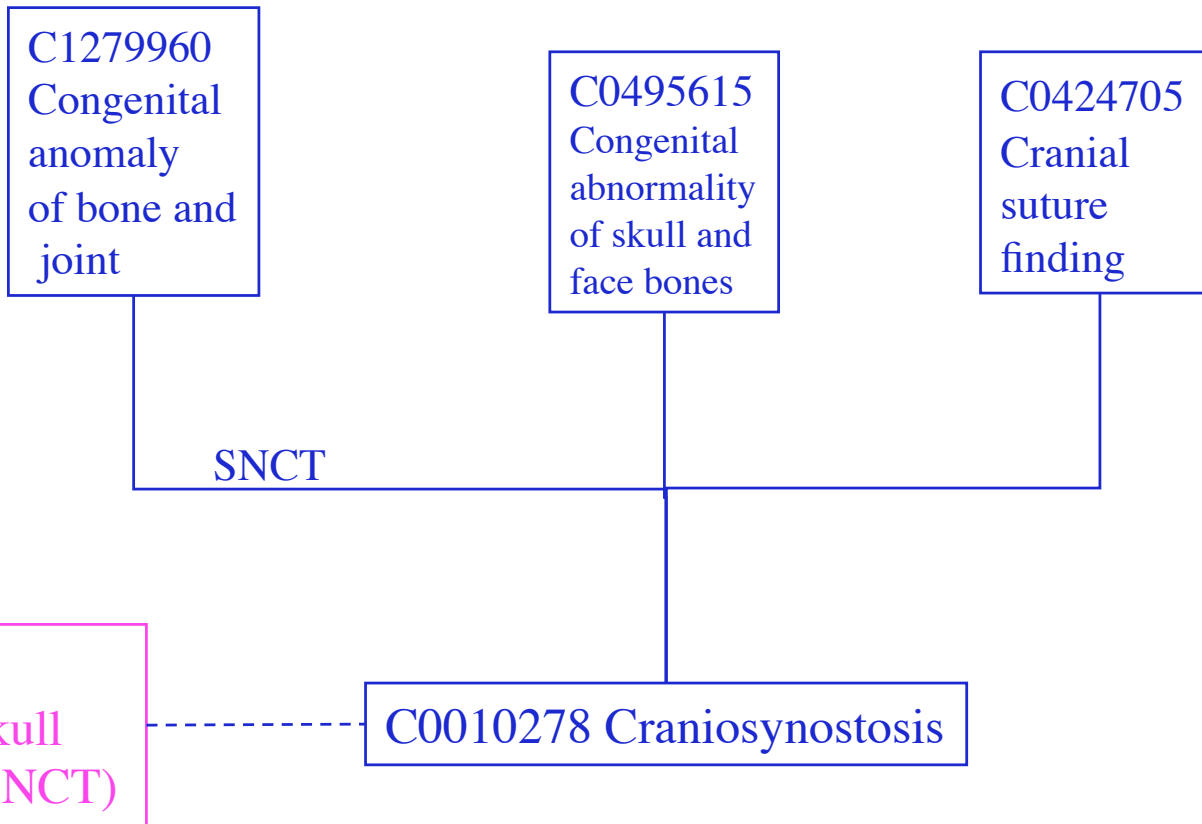


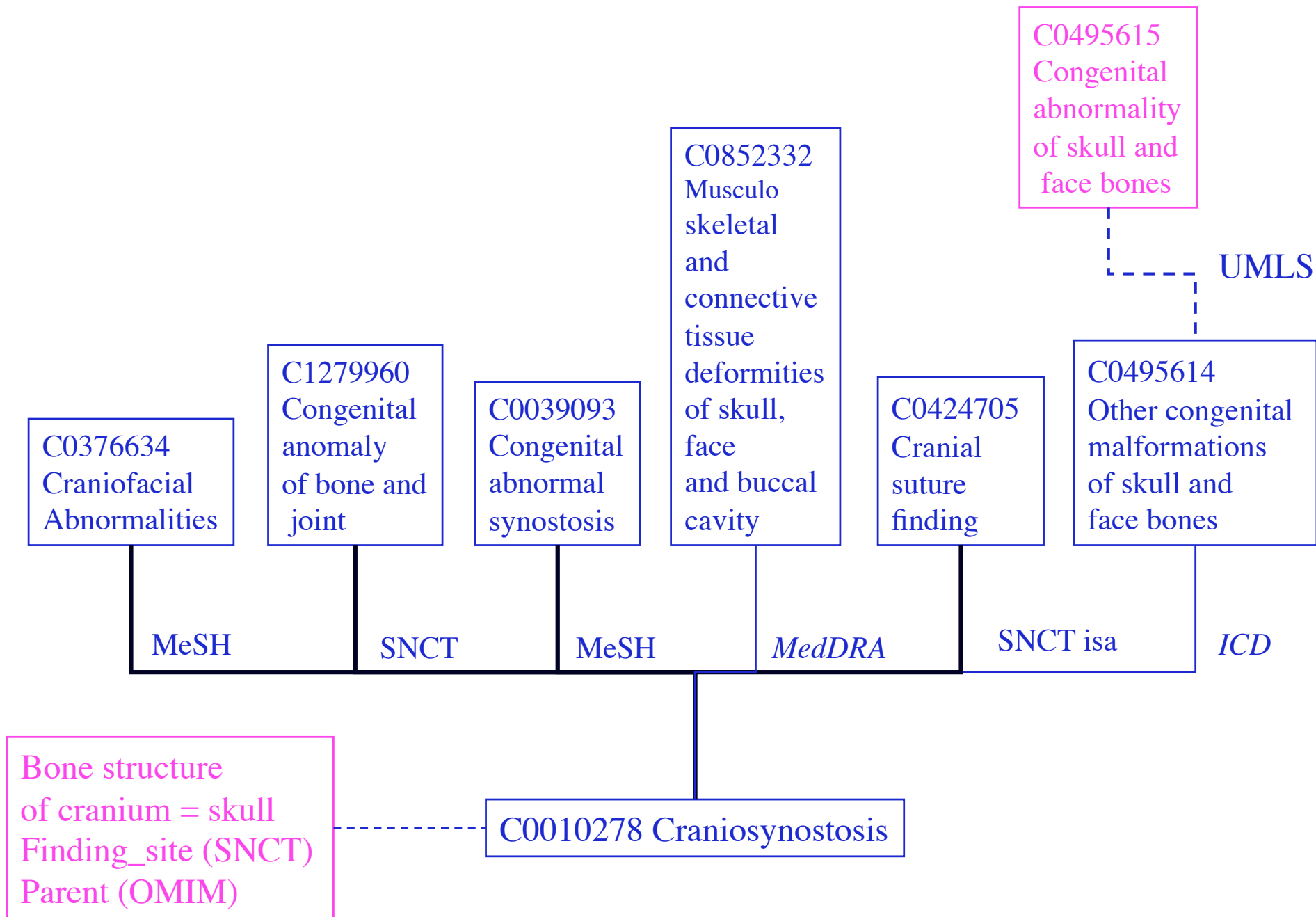
C0010278 Craniosynostosis

OMIM



SNOMED CT

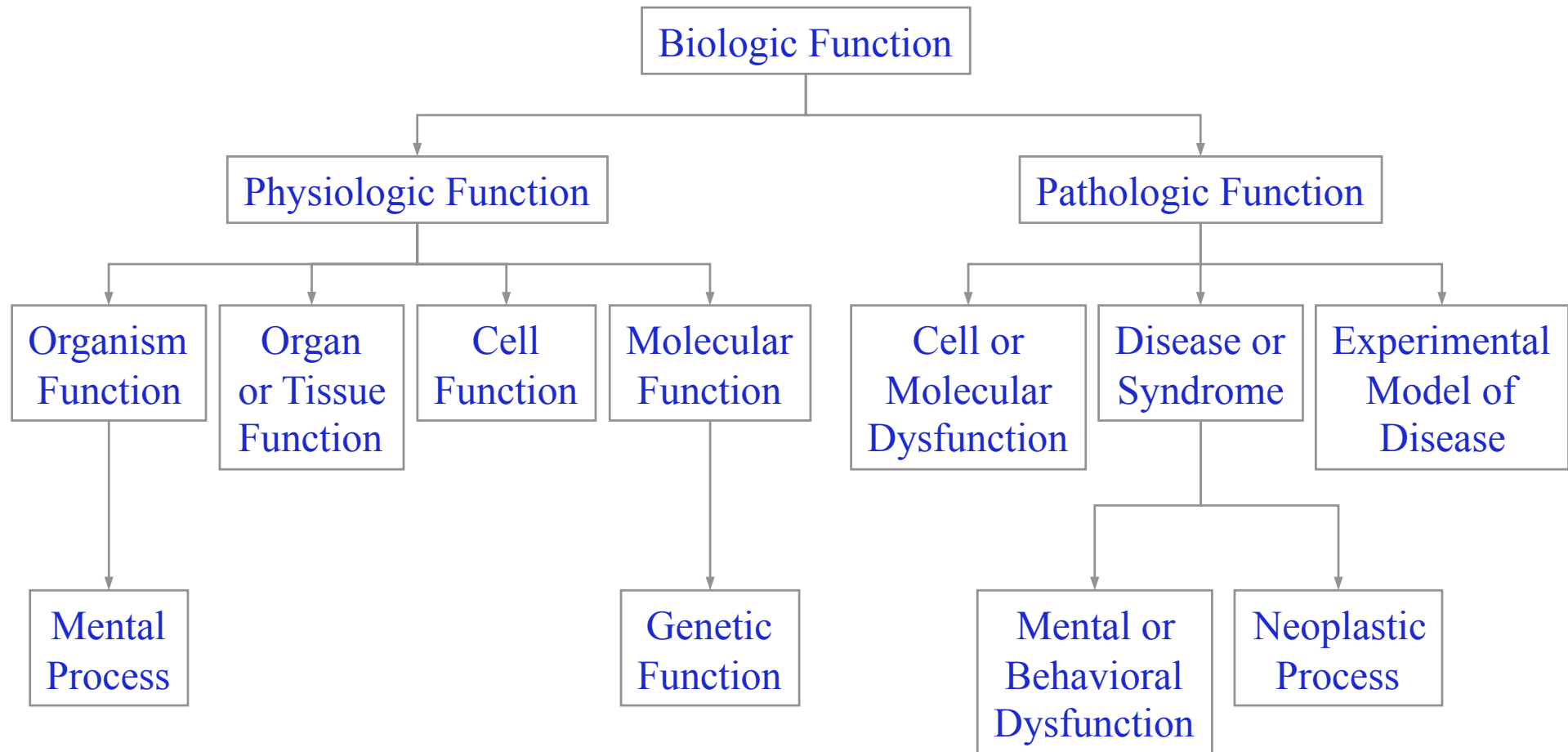




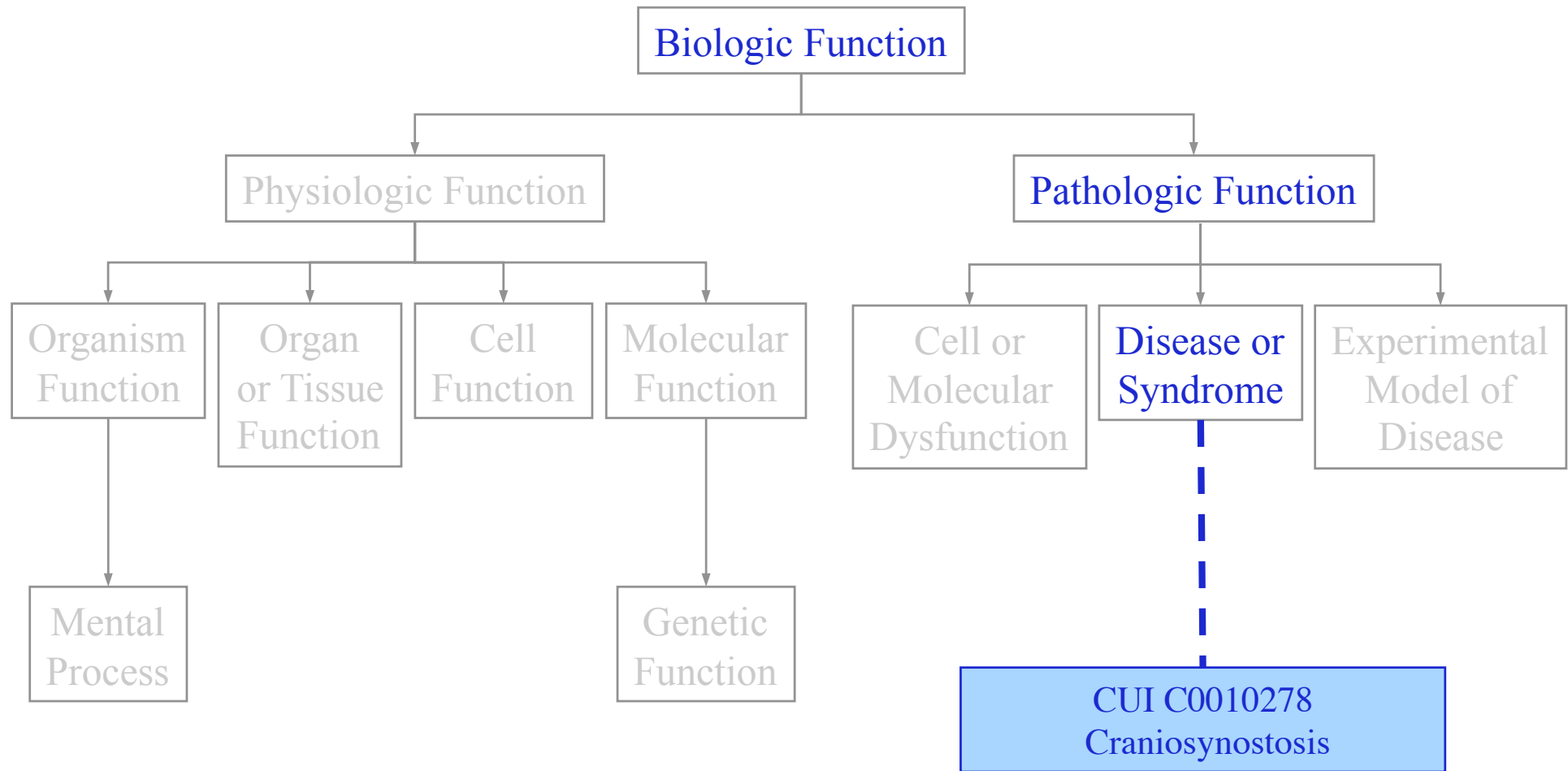
Unified Medical Language System

- Addresses heterogeneity issues
 - More than 100 source vocabularies
 - Unification
- Clusters terms into concepts
 - Metathesaurus: CUIs
- Organizes hierarchies
 - Metathesaurus: relations
- **Categorizes concepts**
 - **Semantic Network: Semantic Types**

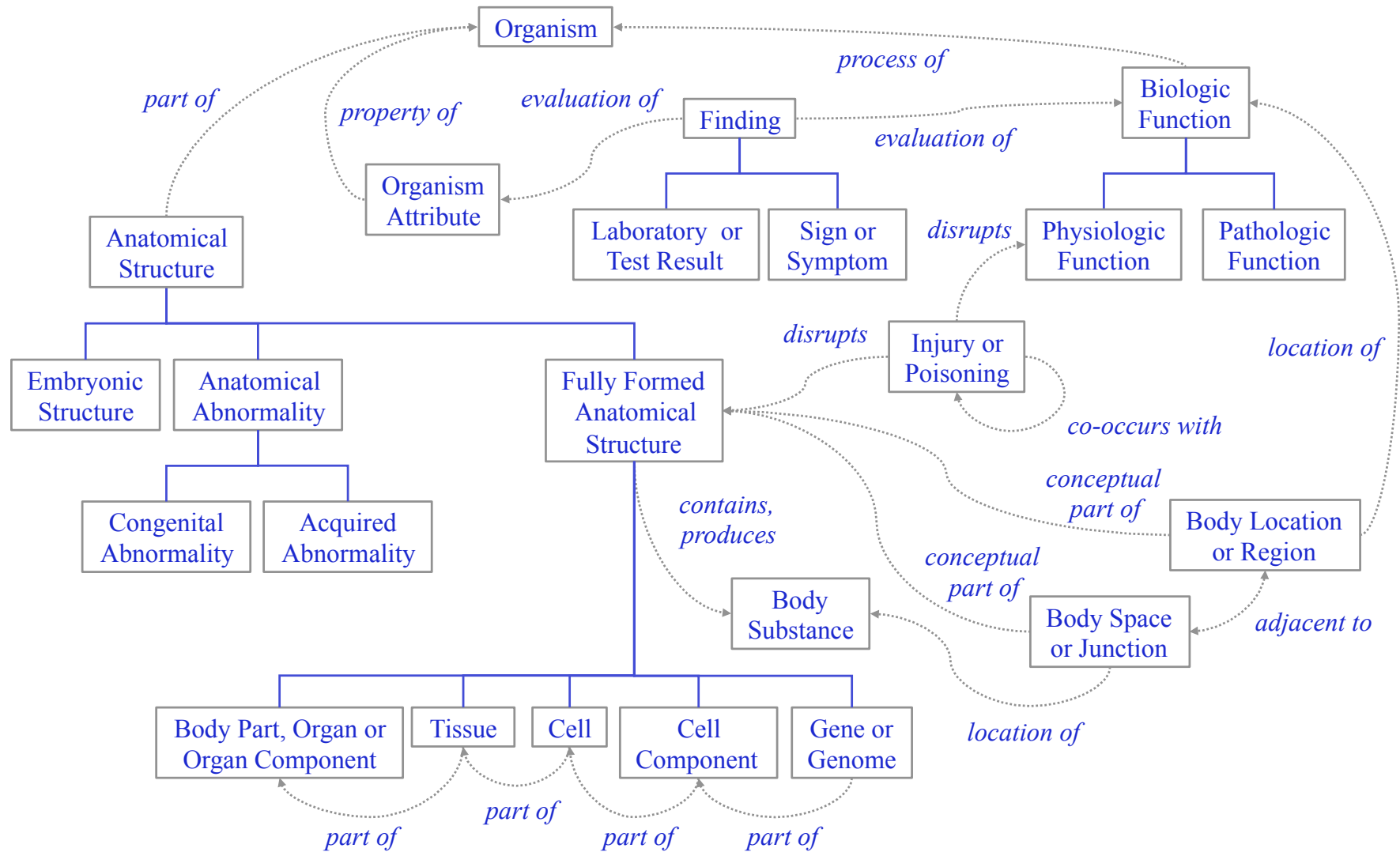
“Biologic Function” hierarchy (isa)



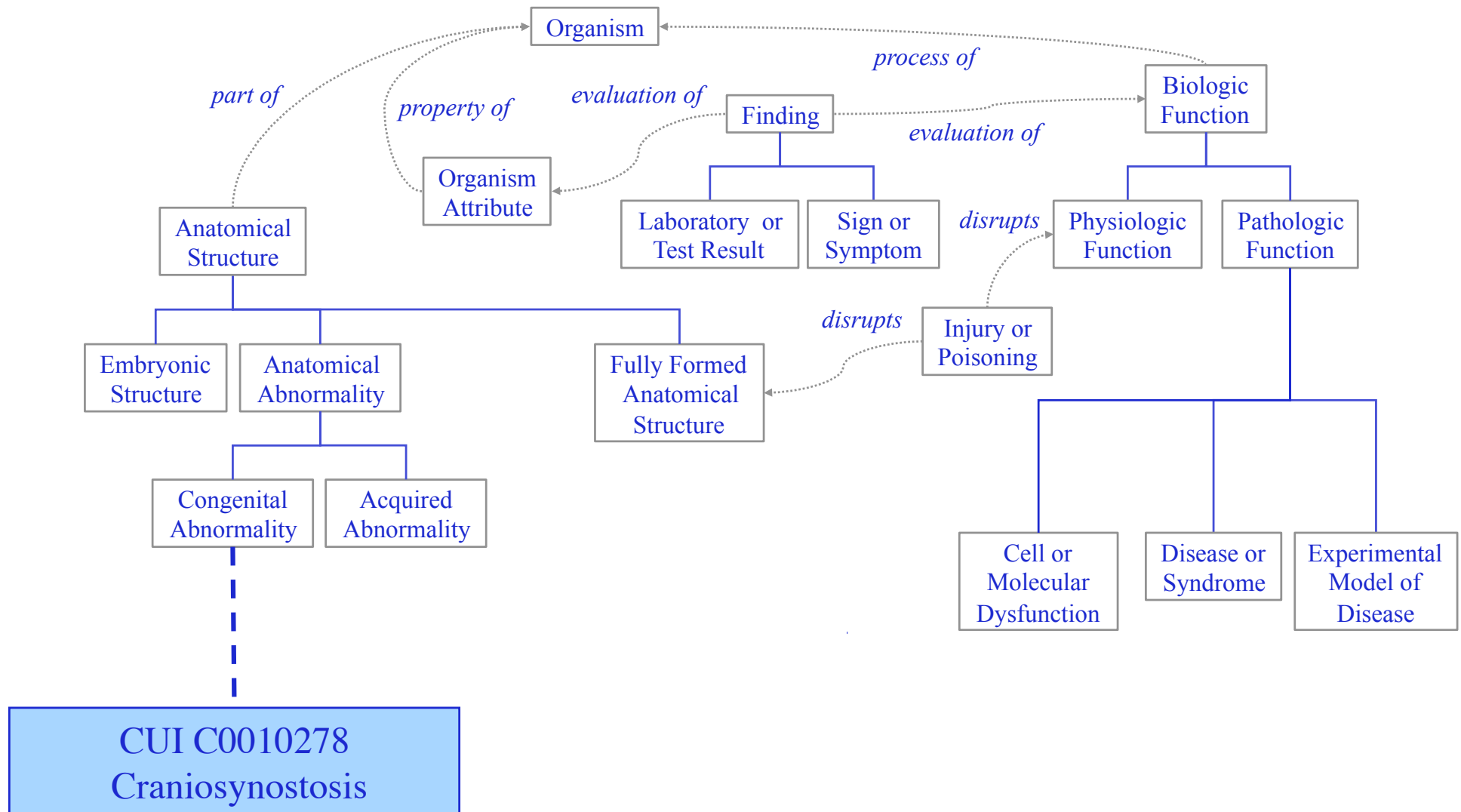
“Biologic Function” hierarchy (isa)



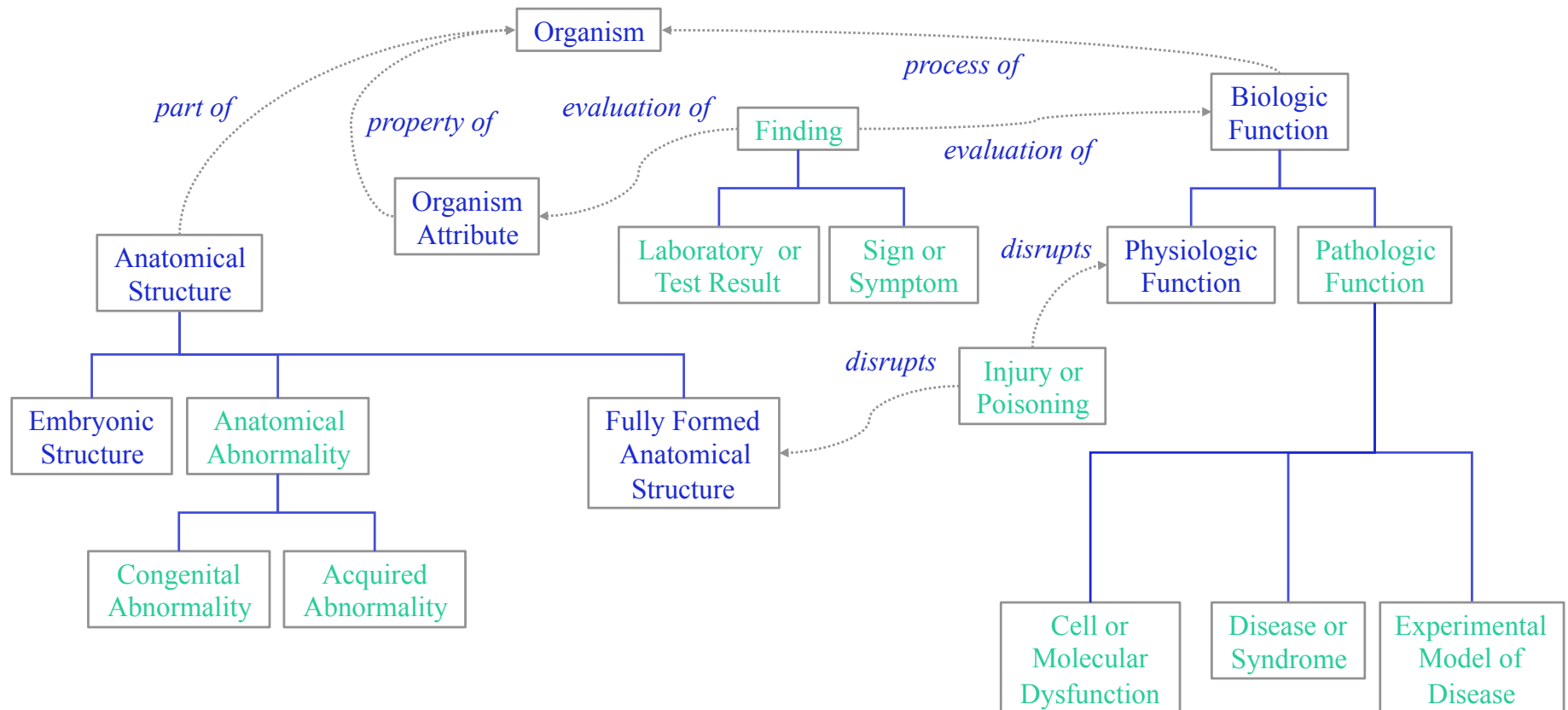
Overview of the Semantic Network



Overview of the Semantic Network

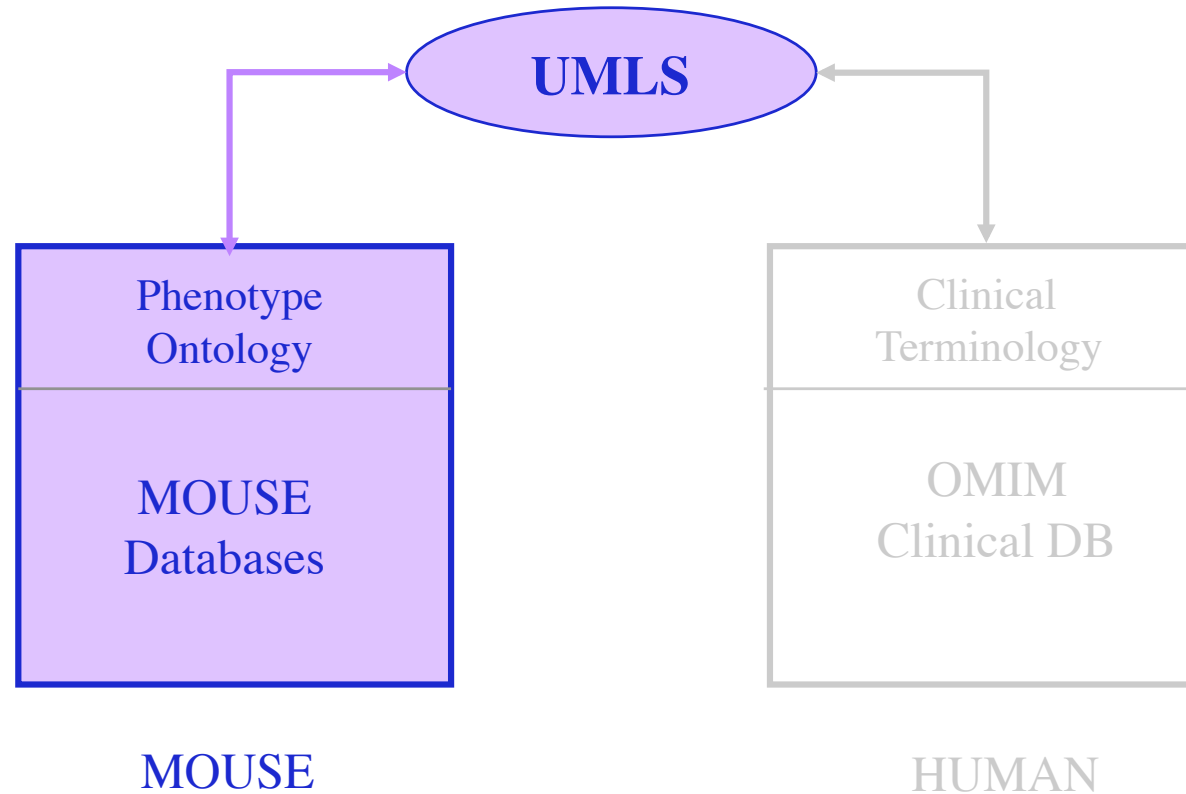


Semantic Groups



Disorders

Phenotype coding: role of the UMLS



Mammalian Phenotype Ontology (MPO)

- 14,662 terms
- 6,307 concepts
- MP:0003561 Rheumatoid arthritis
- MP:0000218 increased leukocyte cell number
 - increased leukocyte count
 - increased WBC count
 - increased WBC number
 - increased white blood cell number
 - leukocytosis
- MP:0000410 waved hair
 - curly hair
 - waved fur
 - wavy hair

Mapping to UMLS (1/3)

- Step 1 : Exact/normalized match
- Results
 - 2,065 MPO terms mapped successfully(14%)
 - 1,495 MPO concepts mapped successfully (24%)
 - Among them, 1,432 correspond to Disorders in UMLS (SG)
- Examples mapped successfully
 - MP:0000062 increased bone density-> C1141880 Bone density increased (NSI)
 - MP:0000081 craniostosis -> C0010278 Craniosynostosis (syn in MPO, EM)
 - MP:0000061 brittle bones -> C0029434 Osteogenesis Imperfecta (syn in UMLS, EM)
- Examples unmapped
 - MP:0000100 abnormal ethmoidal bone
 - MP:0000101 absent ethmoidal bone
 - MP:0000687 small lymphoid organs
 - MP:0000689 abnormal spleen structure

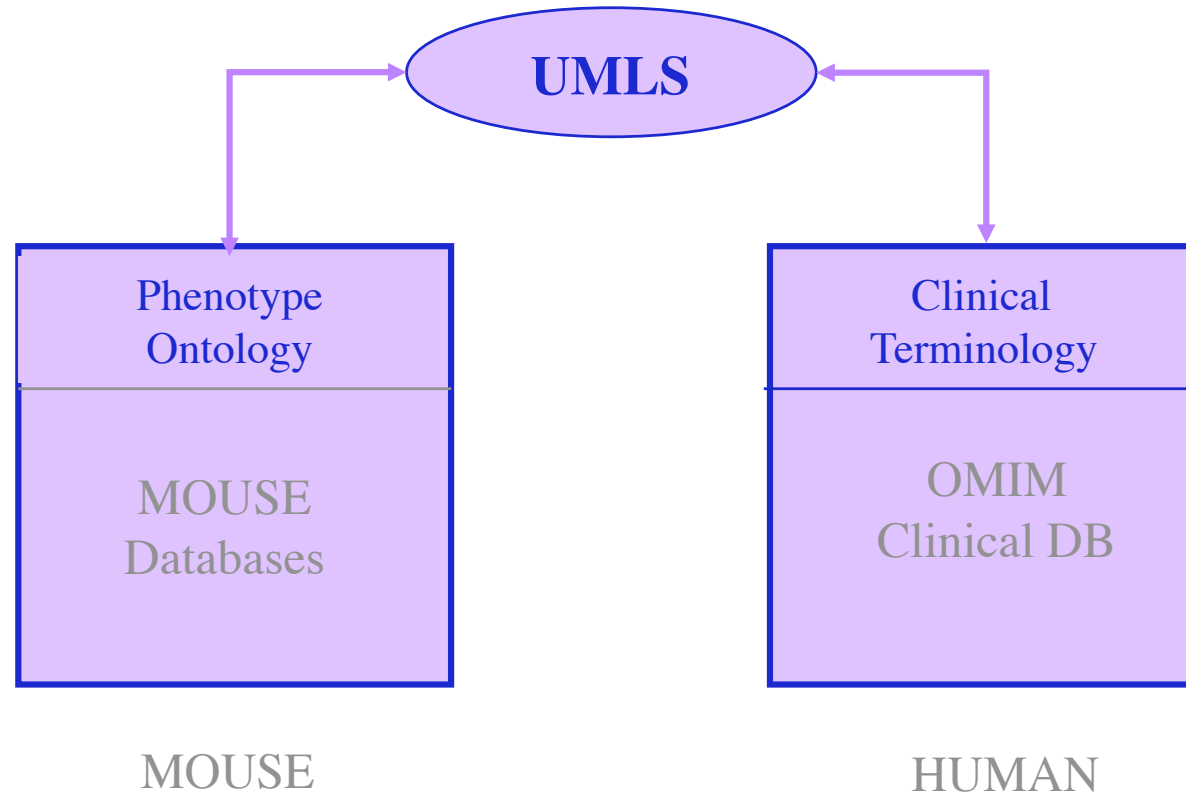
Mapping to UMLS (2/3)

- 11,466 unmapped terms (4,812 concepts)
 - MP:0000100 **abnormal** ethmoidal bone
 - MP:0000101 **absent** ethmoidal bone
 - MP:0000687 **small** lymphoid organs
 - MP:0000689 **abnormal** spleen structure
- Step 2 : demodification
 - 30 modifiers: abnormal, absent, small.....
 - Demodified terms
 - Mapping to UMLS
- Results : demodified terms
 - 9,845 <<modifier> xxx> terms in MPO
 - 7,925 unique terms (after demodification)

Mapping to UMLS (3/3)

- Results : mapping after demodification
 - 2,359 MPO terms mapped successfully after demodification (out of 11,466 , 20%)
 - Unique terms : 1,586 (20%)
 - 1,645 MPO concepts mapped successfully after demodification (out of 4,812, 34%)
- Demodified terms correspond mostly to:
 - Anatomical concepts
 - Semantic Group **Anatomy** in UMLS
 - 1410 terms, e.g., abnormal <anatomical structure>
 - MP:0000005 (**increased**) brown fat -> C0006298 Brown Fat (Tissue, **ANAT**)
 - Physiology
 - Semantic Group **Physiology** in UMLS
 - 516 terms, e.g., abnormal <physiological process>
 - MP:0000057 (**abnormal**) osteogenesis -> C0029433 Osteogenesis (Organ or Tissue Function, **PHYS**)

Phenotype coding: role of the UMLS



Discussion

- Compositionality in phenotype terms

< **abnormal** <anatomical structure > >

30 modifiers

Functional Concept

Qualitative Concept

abnormal
decreased
increased
reduced
small

Anatomy

Physiology

SGs

Class Name: abnormal (**PATO:0000460**)

Is A: deviation(from_normal)

Synonym: aberrant, atypical, defective

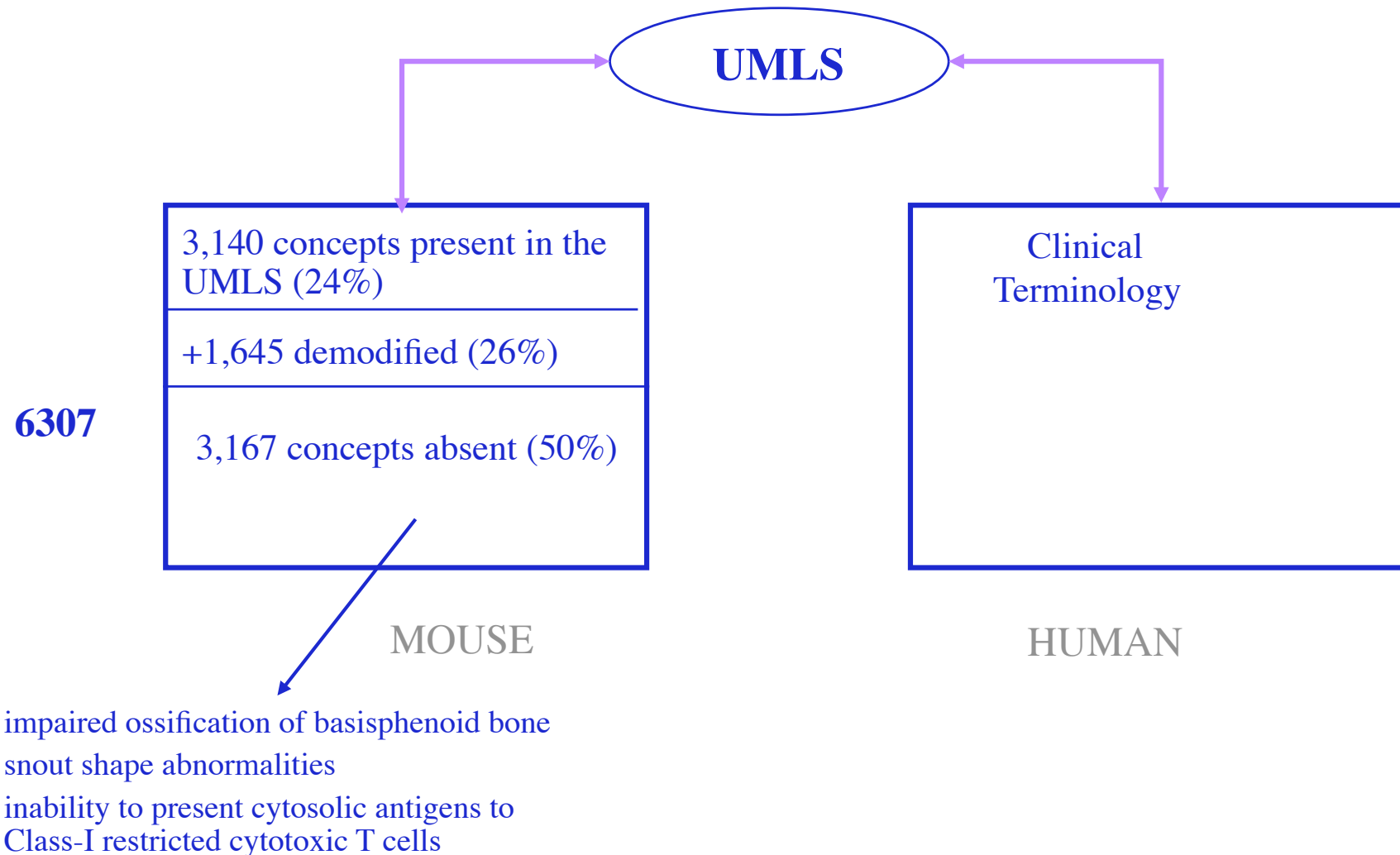
[R]Is A: pathological

entities from any ontology

Phenote : EQ model : combine entities from any ontology
with qualities / traits (such as those in PATO)

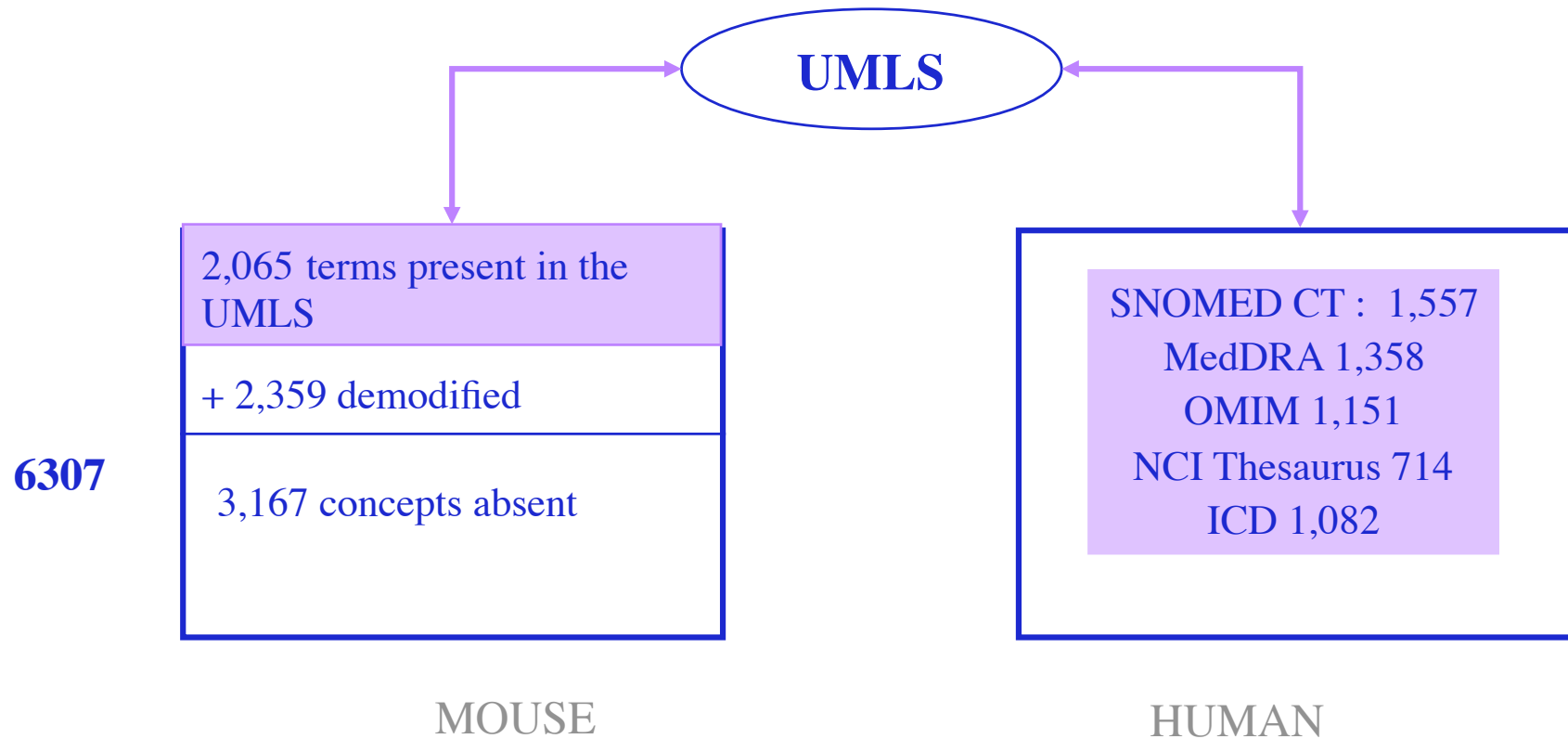
Discussion

- Role of the UMLS in integrating phenotype terminologies



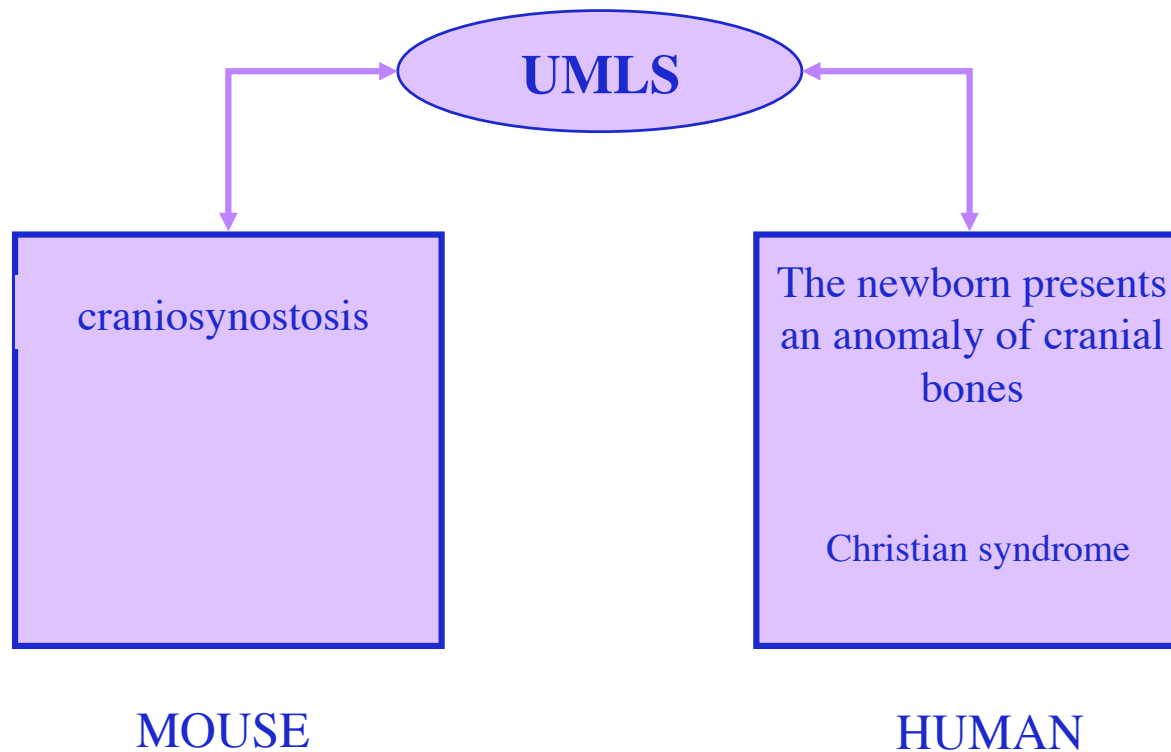
Discussion

- Role of the UMLS in integrating phenotype terminologies



Discussion

- Role of the UMLS in integrating phenotype data
 - OMIM, Clinical DBs, Mouse DBs



Acknowledgements

- Olivier Bodenreider, NLM
- Fleur Mouglin, ISPED

- Download/ customize/
browse the UMLS
- Knowledge Source Server

- umlsks.nlm.nih.gov/

