SNOMED CT Expo
Montevideo, Uruguay
October 29, 2015

# Extending the coverage of phenotypes in SNOMED CT through post-coordination

*Ferdinand Dhombres, Rainer Winnenburg, James T. Case and Olivier Bodenreider*

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

U.S. National Library of Medicine

# Disclaimer

The views and opinions expressed do not necessarily state or reflect those of the U.S. Government, and they may not be used for advertising or product endorsement purposes.

# SNOMED CT Expo 2014

SNOMED CT Implementation Showcase
Amsterdam, The Netherlands
October 30, 2014

ihtsdo Leading healthcare terminology, worldwide

## Coverage of human phenotypes in SNOMED CT

*Rainer Winnenburg and Olivier Bodenreider*

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

NIH U.S. National Library of Medicine

# Medinfo 2015

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

Stud Health Technol Inform. 2015;216:795-9.

**Extending the coverage of phenotypes in SNOMED CT through post-coordination.**

Dhombres F[1], Winnenburg R[2], Case JT[1], Bodenreider O[1].

**Author information**

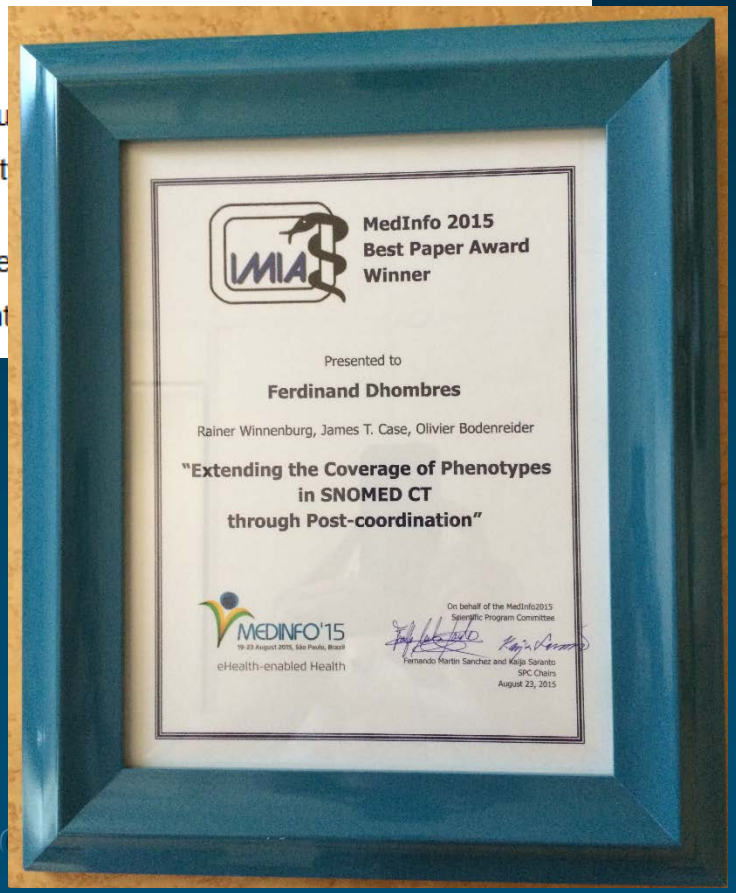
**Abstract**

**OBJECTIVES:** To extend the coverage of phenotypes in SNOMED CT throu

**METHODS:** We identify frequent modifiers in terms from the Human Phenot templates for post-coordinated expressions in SNOMED CT.

**RESULTS:** We identified 176 modifiers, created 12 templates, and generate

**CONCLUSIONS:** Through this novel approach, we can increase the current




NLM

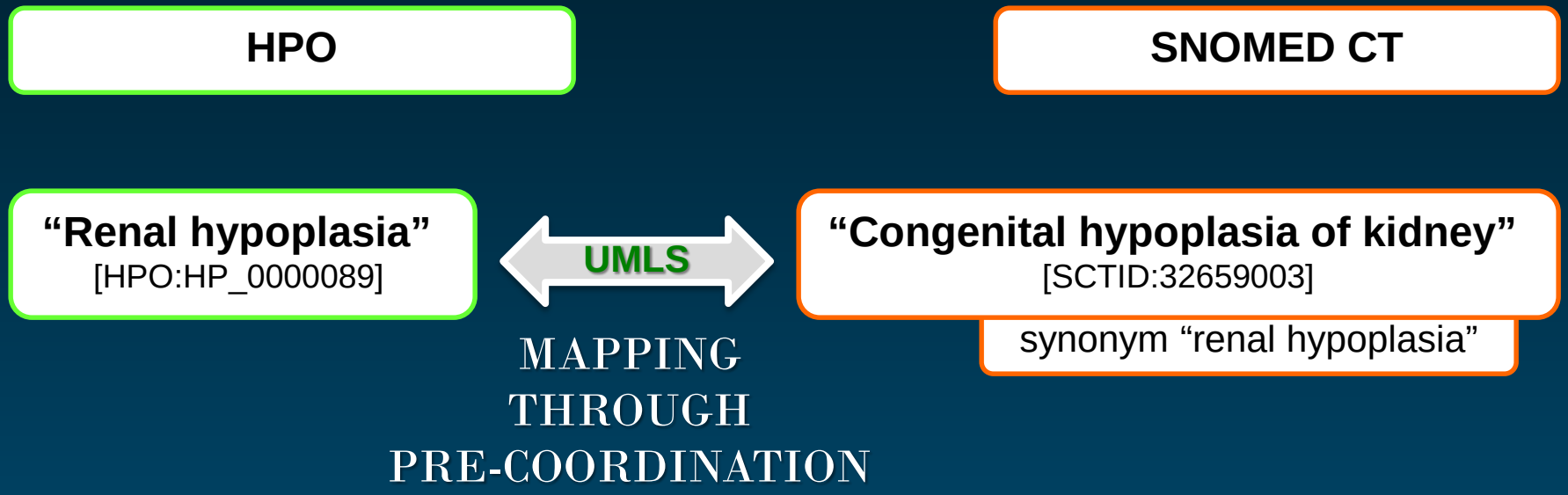
# Introduction

- Phenotype: observable characteristics of an organism (anatomy, physiology, behavior)
- Phenotyping is crucial to understanding how genetic variation relates to clinical manifestations
  - Precise phenotyping is required for the study of rare syndromes
  - Poor interoperability of phenotypic data
    - Across clinical data repositories
    - Between research and clinical data repositories
- Objectives: To extend the coverage of phenotypes in SNOMED CT through post-coordination

# Resources

| HPO | SNOMED CT |
| --- | --- |
| **Human Phenotype Ontology**<br>Developed collaboratively<br>(coordination: Peter Robinson) | Developed by the International Health Terminology Standard Development Organization |
| **Specialized terminology**<br>phenotypes for clinical genetics | **General terminology**<br>broad coverage of Clinical Medicine |
| 10,491 classes for phenotype<br>16,414 terms for phenotype<br>(one preferred term for each class,<br>5,923 exact synonyms) | ~300,000 concepts<br><br>clinical findings ~100,000 concepts<br>~169,000 names |
| Description logic formalism | Description logic formalism<br>supports post-coordination |
| Textual and logical definitions<br>for most concepts | Logical definitions provided<br>for most pre-coordinated concepts |

# Mapping through pre-coordination



# Mapping through pre-coordination

**HPO**

**SNOMED CT**

**“Renal hypoplasia”**
[HPO:HP_0000089]

UMLS

**“Congenital hypoplasia of kidney”**
[SCTID:32659003]

synonym “renal hypoplasia”

MAPPING THROUGH PRE-COORDINATION

**“Macular hypoplasia”**
[HPO:HP_00001104]

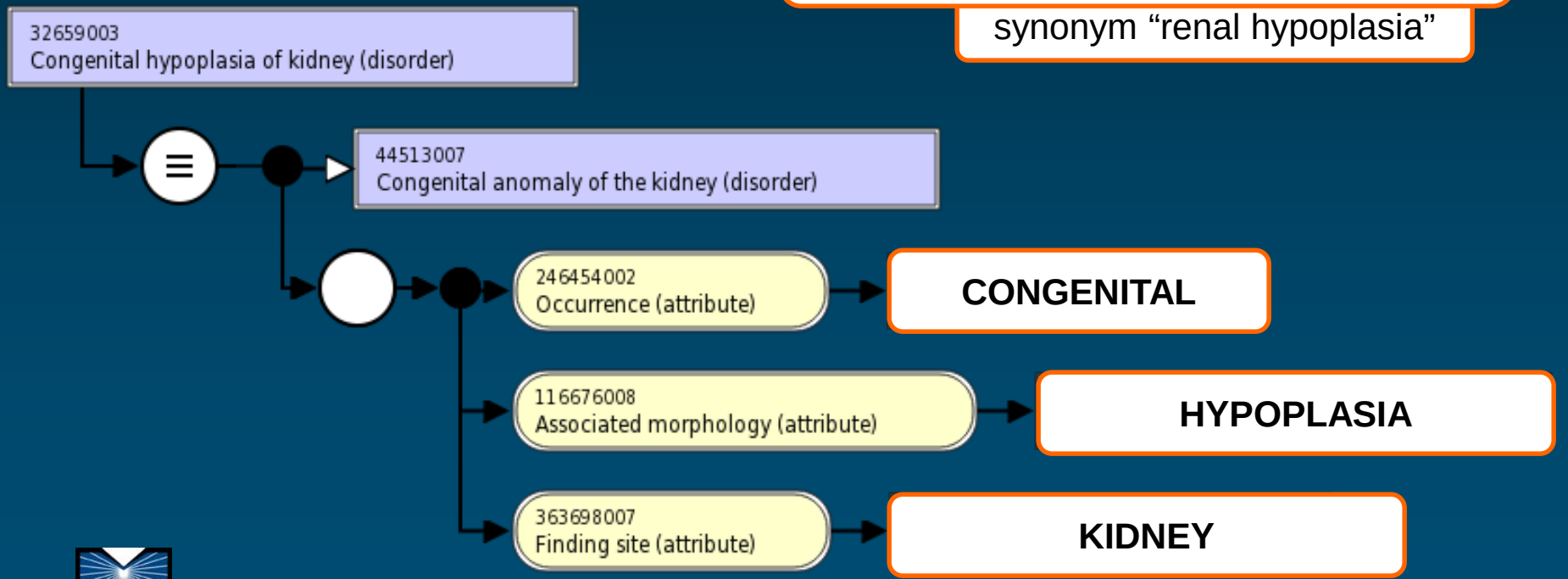
UMLS

# Logical definition

**SNOMED CT**

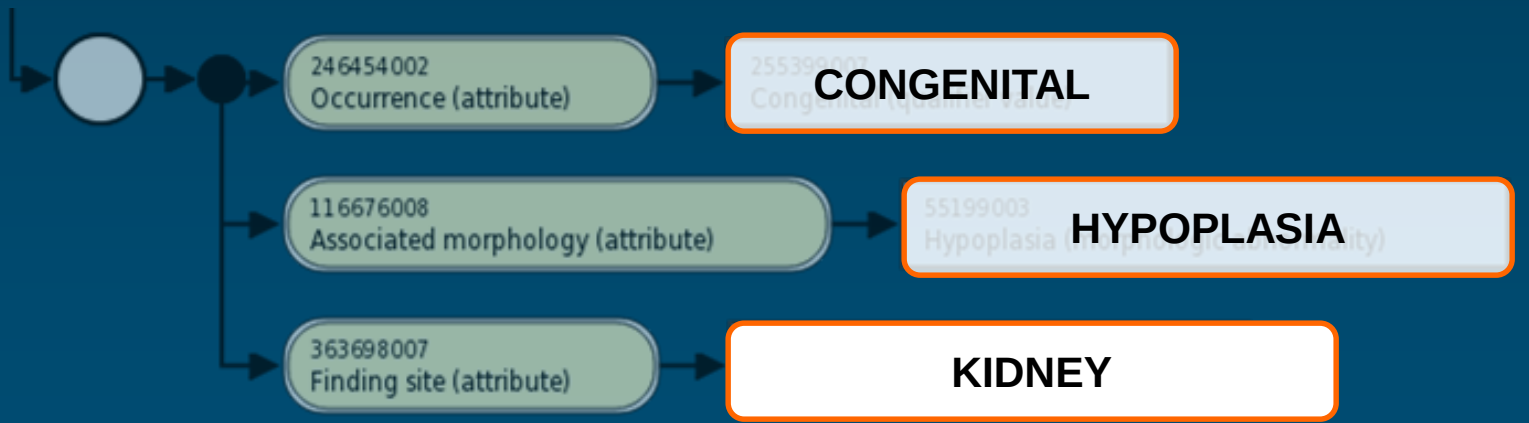
**“Congenital hypoplasia of kidney”**
[SCTID:32659003]

synonym “renal hypoplasia”

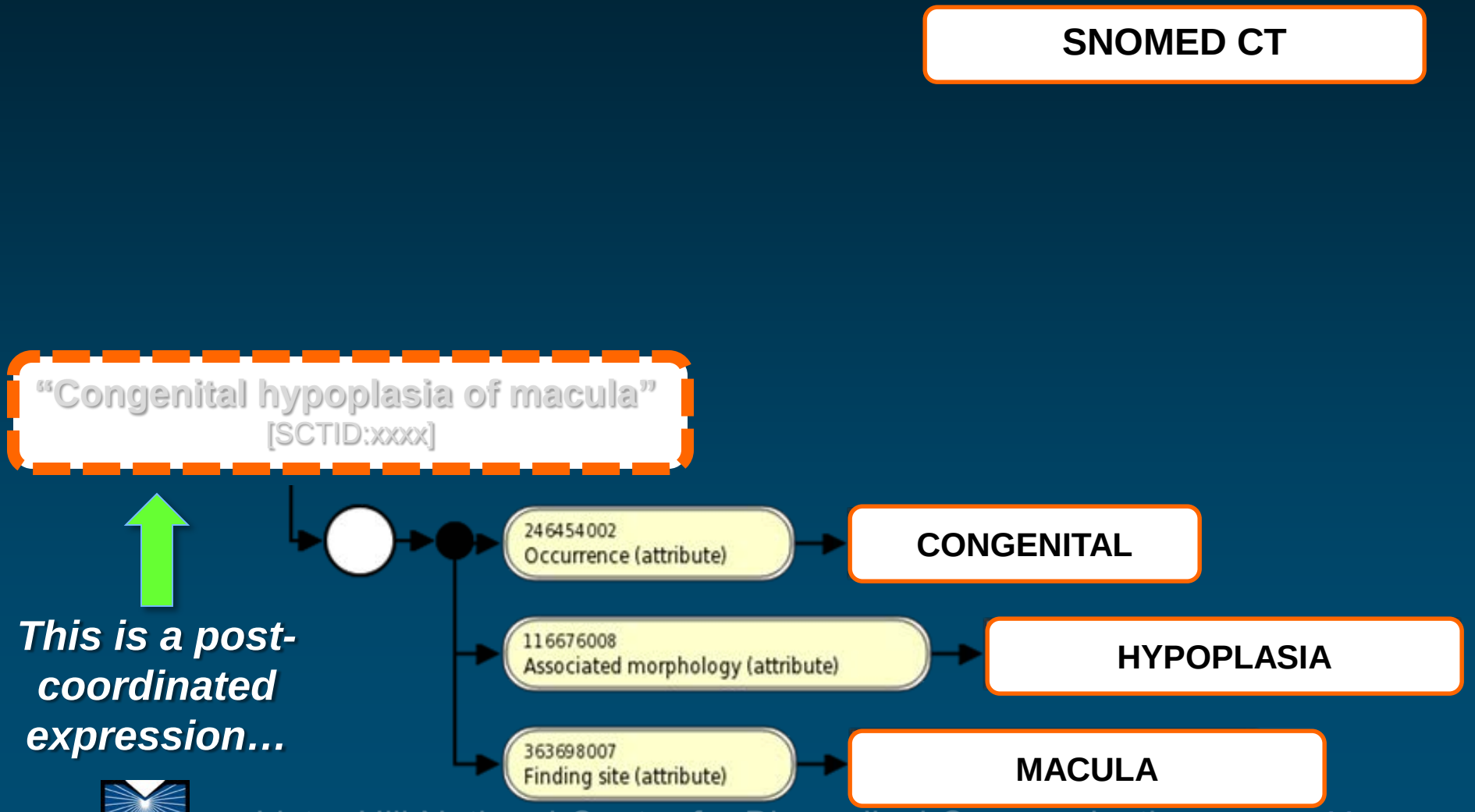


# Logical definition

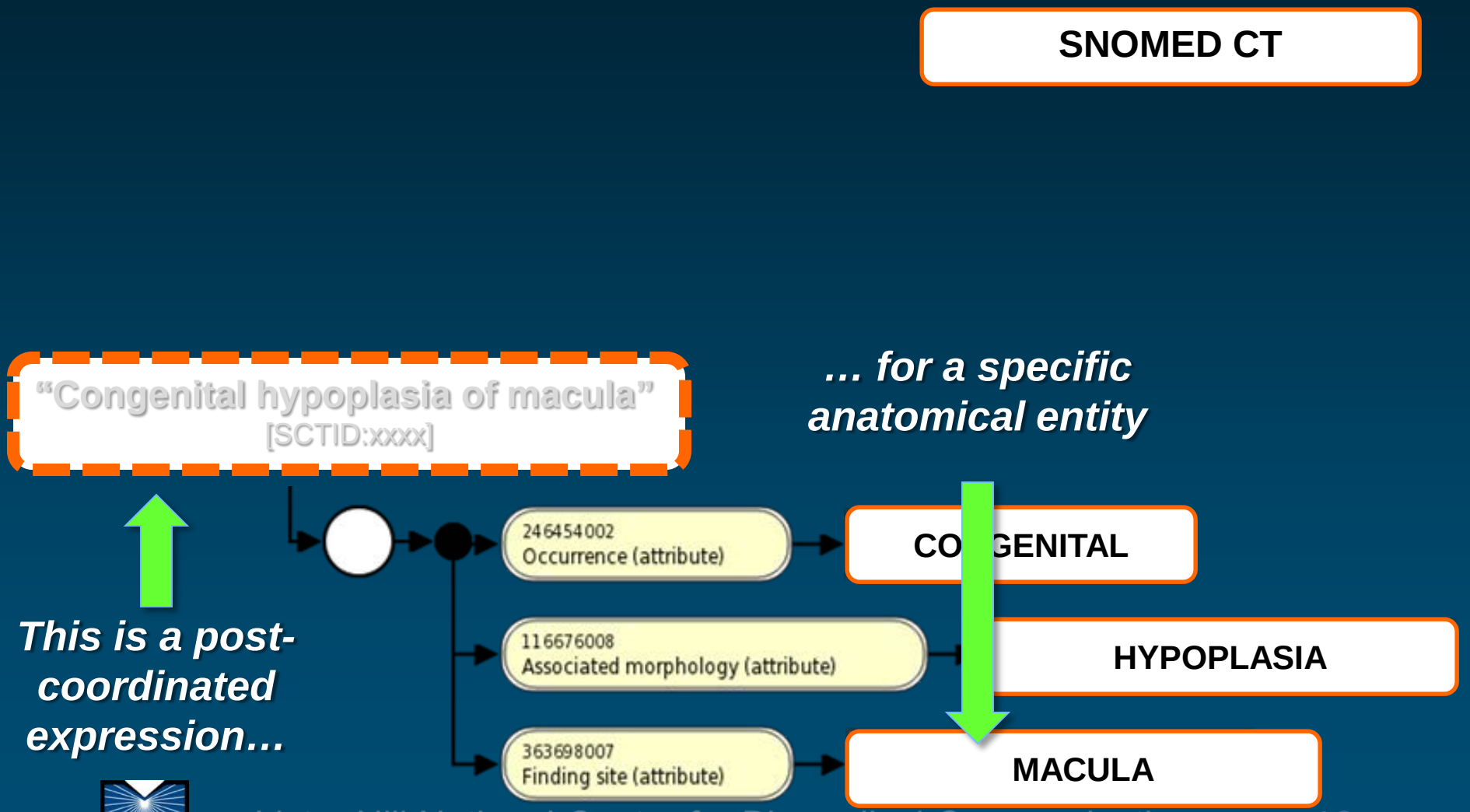
**SNOMED CT**



# Logical definition (modified)



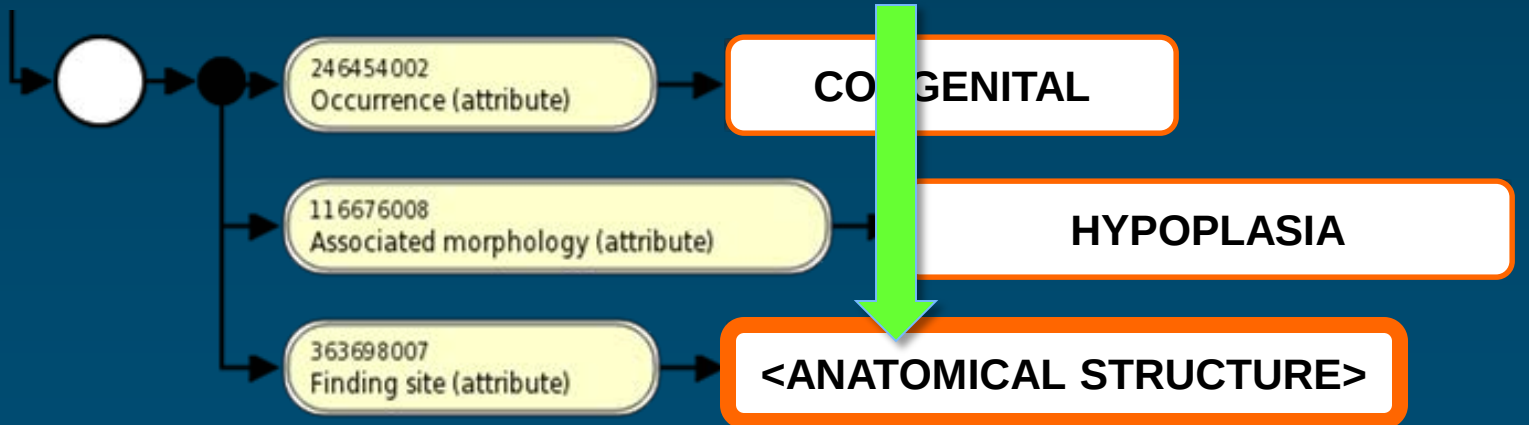
# Logical definition (modified)



# Logical definition (generalized)

**SNOMED CT**

***Generalization***



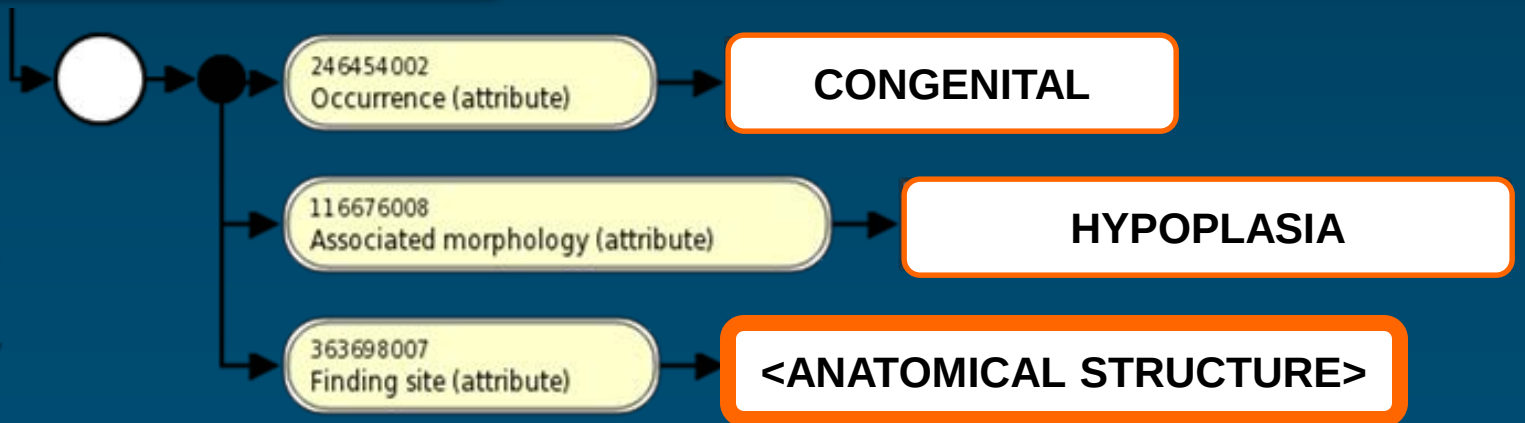
# Template

**SNOMED CT**

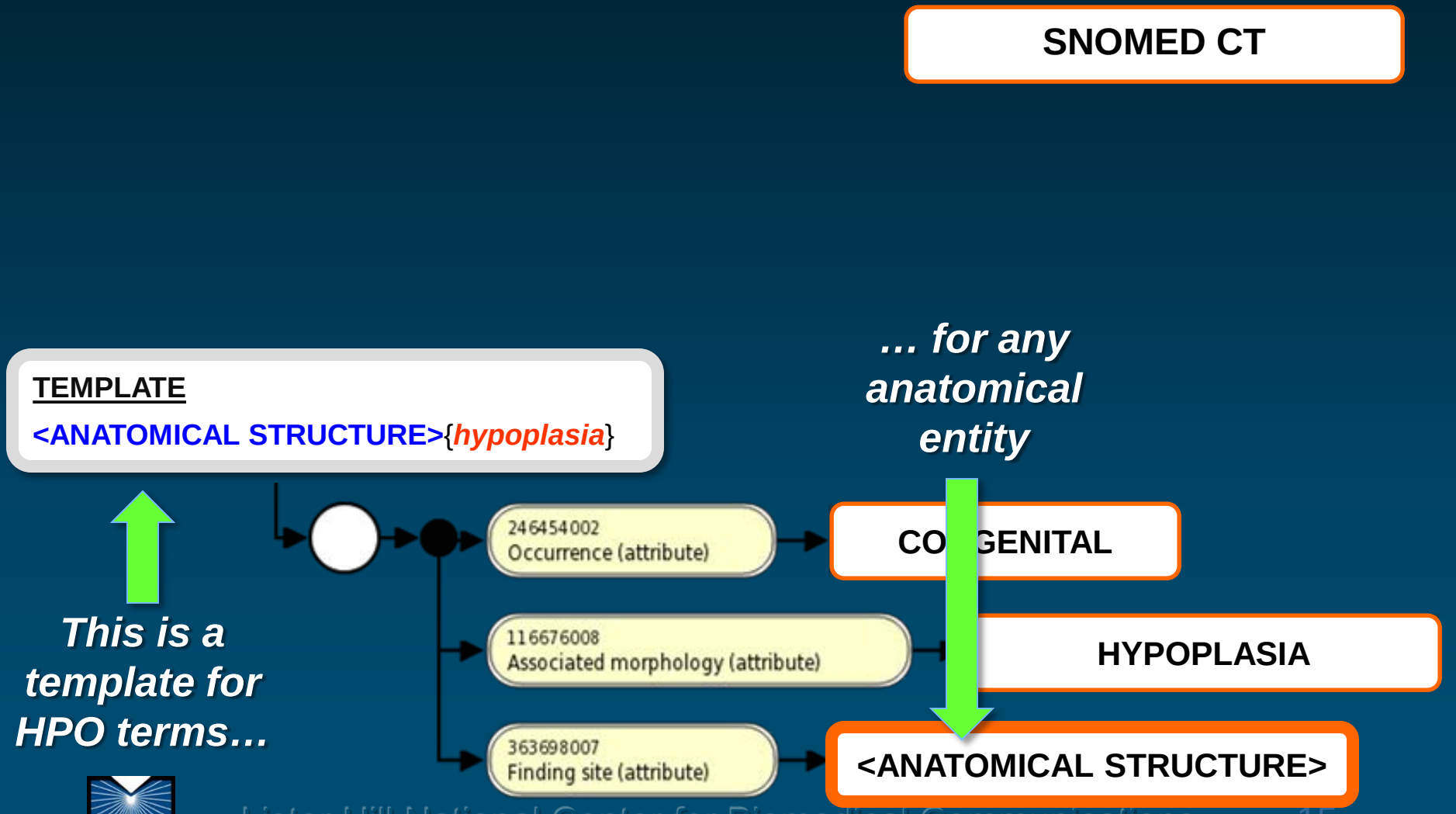
**TEMPLATE**

**<ANATOMICAL STRUCTURE>{*hypoplasia*}**

***This is a template for HPO terms…***



# Template


SNOMED CT
TEMPLATE
<ANATOMICAL STRUCTURE>{hypoplasia}
… for any anatomical entity
This is a template for HPO terms…
246454002
Occurrence (attribute)
CO GENITAL
116676008
Associated morphology (attribute)
HYPOPLASIA
363698007
Finding site (attribute)
<ANATOMICAL STRUCTURE>
NLM

# Mapping through post-coordination

**HPO** | **SNOMED CT**

**“Renal hypoplasia”**
[HPO:HP_0000089]


**“Congenital hypoplasia of kidney”**
[SCTID:32659003]
synonym “renal hypoplasia”

MAPPING THROUGH PRE-COORDINATION

**“Macular hypoplasia”**
[HPO:HP_00001104]

TEMPLATE
<ANATOMICAL STRUCTURE>{qualifier}


MAPPING THROUGH POST-COORDINATION

**“Congenital hypoplasia of macula”**
[SCTID:xxxx]

# Post-coordination in action

- With 12 post-coordination templates, we generated post-coordinated mappings to SNOMED CT for 1617 HPO concepts
- This is in complement to the 3081 HPO concepts for which there is a pre-coordinated mapping to SNOMED CT
- Overall mapping rate: 45%
  - Pre-coordination: 29.4%
  - Post-coordination: 15.4%
- Template-based mappings are usually of high quality

# 12 ontology templates

| Template | Logical definition | # Classes (terms) |
|---|---|---|
| {abnormal, abnormality of, abnormality of the, abnormality involving the }<ANATOMICAL STRUCTURE><br><ANATOMICAL STRUCTURE>{abnormal} | 'Disease (disorder)' and ('Role group (attribute)' some<br>(('Associated morphology (attribute)' some 'Developmental anomaly (morphologic abnormality)')<br>and ('Occurrence (attribute)' some 'Congenital (qualifier value)')<br>and ('Finding site (attribute)' some <ANATOMICAL STRUCTURE>))) | 618<br>(714) |
| {aplastic, aplasia of, aplasia of the, aplasia involving the, absence of}<ANATOMICAL STRUCTURE><br>ANATOMICAL STRUCTURE>{agenesis, aplasia, absent}<br>{congenital absence of, congenital aplasia of }<ANATOMICAL STRUCTURE> | 'Disease (disorder)' and ('Role group (attribute)' some<br>(('Associated morphology (attribute)' some 'Congenital absence (morphologic abnormality)')<br>and ('Occurrence (attribute)' some 'Congenital (qualifier value)')<br>and ('Finding site (attribute)' some <ANATOMICAL STRUCTURE>))) | 415<br>(977) |
| {hypoplastic, hypoplasia of, hypoplasia of the, hypoplasia involving, hypoplasia involving the, hypoplasia affecting the}<ANATOMICAL STRUCTURE><br><ANATOMICAL STRUCTURE>{hypoplasia}<br>{congenital hypoplasia of}<ANATOMICAL STRUCTURE> | 'Disease (disorder)' and ('Role group (attribute)' some<br>(('Associated morphology (attribute)' some 'Hypoplasia (morphologic abnormality)')<br>and ('Occurrence (attribute)' some 'Congenital (qualifier value)')<br>and ('Finding site (attribute)' some <ANATOMICAL STRUCTURE>))) | 409<br>(853) |
| {duplication of, duplication of the, duplication involving}<ANATOMICAL STRUCTURE><br><ANATOMICAL STRUCTURE>{duplication}<br>{complete duplication of, complete duplication of the}<ANATOMICAL STRUCTURE> | 'Disease (disorder)' and ('Role group (attribute)' some<br>(('Associated morphology (attribute)' some 'Double structure (morphologic abnormality)')<br>and ('Occurrence (attribute)' some 'Congenital (qualifier value)')<br>and ('Finding site (attribute)' some <ANATOMICAL STRUCTURE>))) | 140<br>(232) |
| {bilateral, D1} <ANATOMICAL STRUCTURE><br>examples :<br>{bilateral aplasia}<ANATOMICAL STRUCTURE><br>{bilateral}<ANATOMICAL STRUCTURE>{aplasia}<br>{bilateral absence of}<ANATOMICAL STRUCTURE> … | 'Disease (disorder)'<br>and ('Role group (attribute)' some … ***D1 logical definition here …***<br>and ('Finding site (attribute)' some 'left <ANATOMICAL STRUCTURE>'))<br>and ('Role group (attribute)' some … ***D1 logical definition here …***<br>and ('Finding site (attribute)' some 'right <ANATOMICAL STRUCTURE>')) | 35<br>(63) |
| | TOTAL = | 1,617<br>(2,839) |

# Examples I

- Congenital anomaly of testis [55631001] 618
  - AM=Developmental anomaly; OC=Congenital; FS=Testis st.
- Abnormality of the kidney [HP_0000077]
  - AM=Developmental anomaly; OC=Congenital; FS=Kidney st.

- Duplication of femur [253940004] 140
  - AM=Double structure; OC=Congenital; FS=Bone st. of femur
- Duplication involving bones of the feet [HP_0009136]
  - AM=Double structure; OC=Congenital; FS=Bone st. of foot

# Examples II

- Aplasia of spleen [702624008] 415
  - AM=Aplasia; OC=Congenital; FS=Splenic st.
- Hypoplastic/aplastic middle ear structures [HP_0008773]
  - AM=Aplasia; OC=Congenital; FS=Middle ear st.

- Congenital hypoplasia of kidney [32659003] 409
  - AM=Hypoplasia; OC=Congenital; FS=Kidney st.
- Aplasia/Hypoplasia of the macula [HP_0008059]
  - AM=Hypoplasia; OC=Congenital; FS=Macula lutea st.

# Challenges (lexical)

- **HPO terms**: Aplasia of the phalanges of the 4th toe
  - Rule: replace ordinal abbreviations
    - Aplasia of the phalanges of the **fourth** toe
  - Impact:1093 terms (1063 classes)

- **HPO terms**: Aplasia/hypoplasia of the thymus
  - Rule: split and create two terms for terms with a "/"
    - **Aplasia** of the thymus
    - **Hypoplasia** of the thymus
  - Impact: 543 terms (345 classes)

# Challenges (semantic)

- Implicit vs. explicit congenitality
  - Not an issue when a condition is always congenital
    - HPO: Aplasia/Hypoplasia of the macula
    - SNOMED CT: AM=Aplasia/Hypoplasia; OC=Congenital; FS=Macula lutea st.
  - Not (too much of) an issue when the context could be assumed from HPO, but problematic as HPO extends to phenotypes beyond genetic disorders
    - HPO: Abnormality of the elbow
    - SNOMED CT: AM=Developmental anomaly; OC=Congenital; FS=Elbow region st.

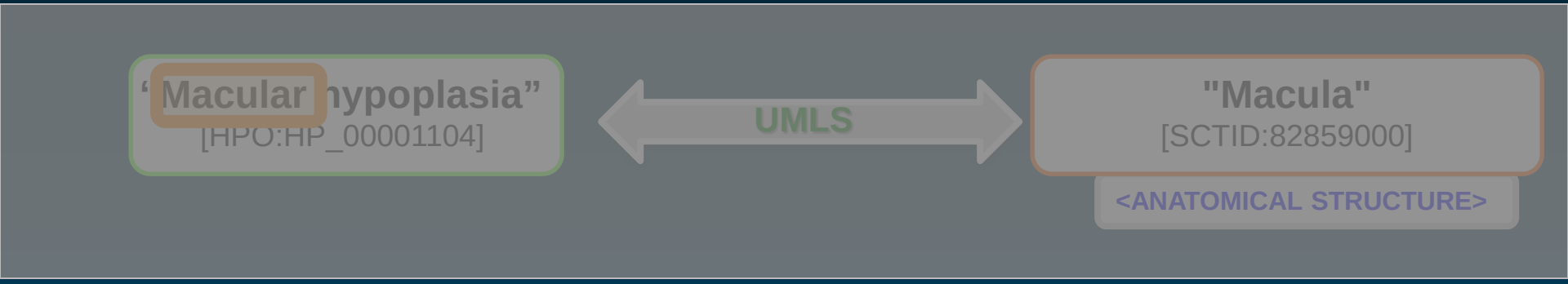
# Challenges (semantic)

- Issues with the mapping through UMLS synonymy
  - Mouth Abnormalities [C0026633]
    - *Abnormality of the mouth*
    - *Congenital anomaly of mouth*
  - Congenital Abnormality [C0000768]
    - *Anomaly*
    - *Congenital malformation*

# Challenges (pragmatic)

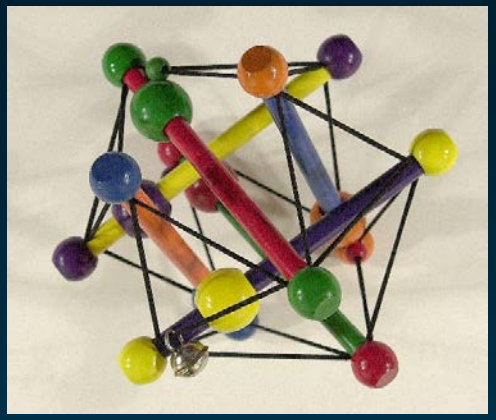
- With post-coordination
  - Not end user-friendly
  - Impractical in regular clinical data entry systems
  - “excessive pre-coordination” – perspective of terminologists vs. clinicians

# Future work


“Macular hypoplasia”
[HPO:HP_00001104]
UMLS
"Macula"
[SCTID:82859000]
<ANATOMICAL STRUCTURE>


“Abnormality of intracranial pressure”
[HPO:HP_00012640]
UMLS
"Intracranial pressure"
[SCTID:250844005]
<PHYSIOLOGICAL PROCESS>


“Abnormality of prothrombin”
[HPO:HP_00012200]
UMLS
"Prothrombin"
[SCTID:7348004]
<CHEMICAL SUBSTANCE>



# Medical Ontology Research


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

NATIONAL LIBRARY OF MEDICINE

*Olivier Bodenreider*

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