

Title: Investigating partial mappings between HPO and SNOMED CT

Presenter: Ferdinand Dhombres and Olivier Bodenreider, National Library of Medicine, USA

Audience

Developers and users of SNOMED CT interested in the representation of phenotypes.

Objectives

1. To learn about the coverage of phenotypes in HPO and SNOMED CT;
2. To understand how partial mappings to SNOMED CT can be created for HPO phenotypes;
3. To describe the contribution of lexical and logical approaches to creating partial mappings.

Abstract

Background. Identifying partial mappings between two terminologies is of special importance when one terminology is finer-grained than the other, as is the case for the Human Phenotype Ontology (HPO), mainly used for research purposes, and SNOMED CT, mainly used in healthcare.

Objectives. To investigate and contrast lexical and logical approaches to deriving partial mappings between HPO and SNOMED CT.

Methods. 1) Lexical approach: We identify modifiers in HPO terms and attempt to map demodified terms to SNOMED CT through UMLS; 2) Logical approach: We leverage subsumption relations in HPO to infer partial mappings to SNOMED CT; 3) Comparison: We analyze the specific contribution of each approach and evaluate the quality of the partial mappings through manual review.

Results. There are 7358 HPO concepts with no complete mapping to SNOMED CT. We identified partial mappings lexically for 33% of them and logically for 82%. We identified partial mappings both lexically and logically for 27%. The clinical relevance of the partial mappings (for a cohort selection use case) is 49% for lexical mappings and 67% for logical mappings.

Conclusions. Through complete and partial mappings, 92% of the 10,454 HPO concepts can be mapped to SNOMED CT (30% complete and 62% partial). Equivalence mappings between HPO and SNOMED CT allow for interoperability between data described using these two systems. However, due to differences in focus and granularity, equivalence is only possible for 30% of HPO classes. In the remaining cases, partial mappings provide a next-best approach for traversing between the two systems. Both lexical and logical mapping techniques produce mappings that cannot be generated by the other technique, suggesting that the two techniques are complementary to each other. Finally, this work demonstrates interesting properties (both lexical and logical) of HPO and SNOMED CT and illustrates some limitations of mapping through UMLS.

References

1. Dhombres F, Bodenreider O. Interoperability between phenotypes in research and healthcare terminologies—Investigating partial mappings between HPO and SNOMED CT. *J Biomed Semantics*. 2016 Feb 9;7:3. (PMID: 26865946)
2. Human Phenotype Ontology (HPO) <http://human-phenotype-ontology.github.io/>