Issues in Representing Biological and Clinical Phenotypes Using Formal Models

Ying Tao[1,*], M.D., Chintan Patel[1,*], M.S, Carol Friedman[1,†], PhD, Yves A. Lussier[1,2,†],M.D.

1- Dept. of Biomedical Informatics, Columbia University, New York, NY,USA
2- Center for Biomedical Informatics and Dept. of Medicine, The University of Chicago, Chicago, IL, USA

Representing phenotypes in a structured and standardized manner across different biological species poses significant challenges. We performed a modeling experiment to compare a model called the Canon model, and the PATO for representing a range of biological and clinical phenotypes. The formal nature of Canon model allows for complex representations, but lacks the simplicity offered by PATO. A phenotype model allowing flexible representation with unique semantic interpretation is desired.

BACKGROUND

The Phenotype Attribute and Value Ontology (1) (PATO) is an emerging standard to annotate assayed phenotypes in a structured and coherent manner across different biological species. Canon group (2) developed a model for the formal (canonical) representation of clinical information for data exchange and medical applications.

METHODS

We selected a diverse set of phenotypes from Wormbase, OMIM and chest radiology report (radiographic findings/phenotypes). We then evaluated the PATO and Canon models by encoding the phenotypes into each model.

RESULTS AND DISCUSSION

Examples of the phenotype modeling experiment are described in table 1. The flexibility of choosing the entity from an external ontology in PATO can lead to multiple representations, for example, *vulval_differentiation* (mammalian phenotype ontology) or *vulva* (anatomy ontology); it is not clear how semantic equivalences can be inferred from such representations. Developing a symbolic model that can represent and reason with complex concepts such as '*penetrance*' is challenging. Furthermore, concepts having deep nested structures need a more formal representation framework to capture the knowledge at finer granularity (e.g. *slight interval decrease*). The Canon model with its logic based representation allows for formal and complex representations but the familiarity and acceptance of such a model among end-users remains an open issue. We conclude that using PATO with a formal description logic language, as the one provided in Canon, would provide a more expressive and less ambiguous framework for representing clinical and biological phenotypes, however additional studies are required to evaluate the usability aspects of the combined model.

References

1. www.bioontology.org/wiki/index.php/PATO:Main_Page
2. Friedman C, Huff SM, Hersh WR, Pattison-Gordon E, Cimino JJ: The Canon Group's effort: working toward a merged model. J Am Med Inform Assoc 2:4-18 (1995)..

Table 1. Modeling biological and clinical phenotypes using the PATO and the CANON model

Phenotype	**PATO** (observable entity | attribute | value) **Note: the latest version of PATO does not have notion of 'attributes' (1)**	**CANON Model** (conceptual graph)
negatively regulates vulval differentiation ***(WormBase)***	Vulva Differentiation | regulation | negative	[phenotype: #ark1Fun] - (has-observation) → [differentiation] (has-location) → [vulval] (has-process*) → [negatively regulated]
Cystic Fibrosis with pancreatic insufficiency in 80% ***(OMIM)***	Cystic Fibrosis Pancreas | enzyme_function | Insufficient*	[phenotype: MIM:219700]- (has-observation) → [enzyme function] (has-location) → [pancreas] (has-degree) → [insufficient] (has-penetrance) → [80%]
Slight interval decrease in left pleural effusion ***(Radiology Report)***	Pleural effusion | local_qualifier | left Pleural effusion | temporal | decrease Left Pleural Cavity | pathological change | pleural effusion Left Pleural Cavity | temporal | decrease	[phenotype: # BWH22.09] - (has_observation) → [pleural_effusion] (has_location_qualifier) → [left] (has_temporal) → [decrease_in] - (has_degree) → [slight] (has_temporal) → [interval]

* represents concepts not present in the model

* These authors contributed equally to the work † Corresponding authors who have contributed equally