Mondo Disease Ontology: harmonizing disease concepts around the world

Nicole A. Vasilevsky[1], Shahim Essaid[1], Nico Matentzoglu[2], Nomi L. Harris[3], Melissa A. Haendel[1,4], Peter N. Robinson[5], Christopher J. Mungall[6]

[1]Oregon Health & Science University, [2]European Bioinformatics Institute, [3]Lawrence Berkeley National Laboratory, [4]Oregon State University, [5]Jackson Laboratory

What is Mondo?

Mondo is a unified disease ontology, encompassing many disease terminologies, which aims to harmonize disease definitions across the world.
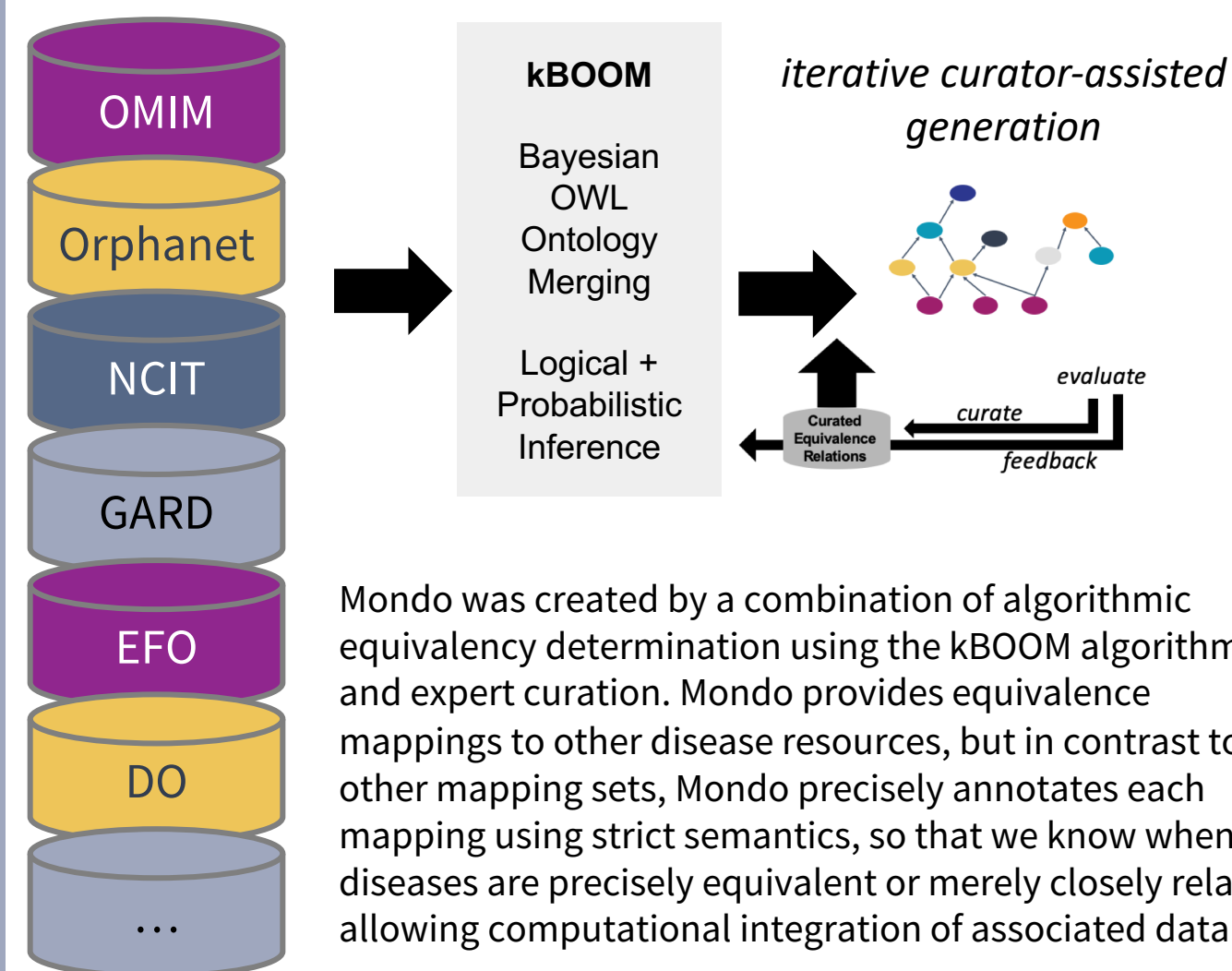
Why did we create Mondo?

- Lack of definitive standard for encoding diseases for information exchange
- Proliferation of mappings between disease terminologies

| NCIT:C3269 schwannoma | equivalent? | ORPHA:252164 Benign schwannoma |

In order to computationally utilize our collective knowledge sources for diagnostics and to reveal underlying mechanisms of diseases, we need to understand which terms are truly equivalent across different resources

How was Mondo created?

OMIM, Orphanet, NCIT, GARD, EFO, DO, …

kBOOM
Bayesian OWL Ontology Merging
Logical + Probabilistic Inference

iterative curator-assisted generation

Curated Equivalence Relations — evaluate — curate — feedback

Mondo was created by a combination of algorithmic equivalency determination using the kBOOM algorithm, and expert curation. Mondo provides equivalence mappings to other disease resources, but in contrast to other mapping sets, Mondo precisely annotates each mapping using strict semantics, so that we know when two diseases are precisely equivalent or merely closely related - allowing computational integration of associated data.

The Mondo Disease Ontology provide a logic-based structure for unifying multiple disease resources, for use in structuring disease descriptions in childhood cancers and beyond.

Hierarchical and Graphical Disease Classification

Terms in Mondo are classified in a hierarchy with multiple parentage, meaning terms can be classified in multiple ways. The figure shows a partial representation of the classification of **MONDO:0021054 'bone sarcoma'** (yellow). This type of classification allows for query expansion – for example, if you query on bone sarcoma, you will retrieve data annotated to subclasses, such as MONDO_0002625 'Ewing sarcoma of bone' (dark grey). In addition, users can search on synonyms of bone sarcoma, such as 'skeletal sarcoma' and return the same results.
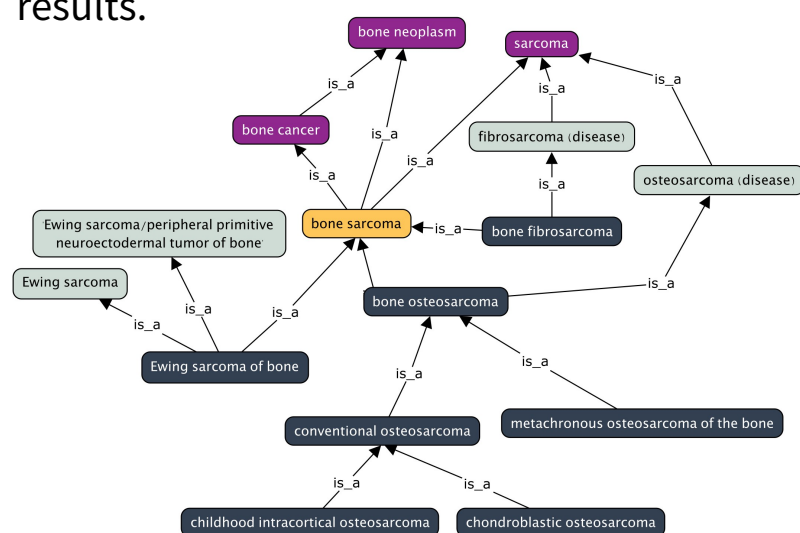
Users

Mondo, along with NCIt and the Human Phenotype Ontology are used to structure disease and phenotype descriptions for a number of resources, including the Kids First Data Resources Center, which is a curated database of clinical and genetic sequence data from pediatric patients with structural abnormalities or childhood cancers.

Get Involved

Mondo Website:
https://monarch-initiative.github.io/mondo/

GitHub Repository:
https://github.com/monarch-initiative/mondo

Join our Mailing List:
https://groups.google.com/d/forum/mondo-users