Innovation Across the Translational Divide Webinar

NOTES

JOIN INFO:	WEBINAR DATE:	Agenda:
WebEx https://nih.webex.com/nih/e.php?MTID=m688450e2b8d0019ab462dff37f86fd40	April 09, 2020	https://monarch-initiative.github.io/phenomics/pages/clin-phen-webinar.html Slides: bit.ly/kfworkshop

Notes:

Melissa H: set the stage for a fruitful discussion.
Why interoperability and utility and challenges. Speakers will address particulars of phenotype data and how we can make those data more useful for searching and cohort selecting.
Learning health systems context.
Terminologies versus data models.
Knowledge goes into clinical guidelines and experts for how to manage a clinical case. Depends on interop of how we model data and the
HPO. Developed by Monarch. Standardized terminology by different in that ..more than one parent to any given term. Logical definition based on terms coming from other ontologies…
Smell function is annotated to 22 species etc…
Lots of data that can be logically related to a phenotype observed in a clinical setting.
Different types of phenotypic information. Semantic, but can also have quantitative. Can use FHIR.. has an observation resource to record phen features.

- PSA value, serum value, each are the outputs of tests that indicate a value that has a range for abnormal high or low for subject - refernce ranges migh tvary based on age, sex etc.
- Quant. Phenotypes can be evaluated for many different tests. Blood nitrite.

[missed a lot]

Adam Resnick

- As the DRC, we have not yet solved some of the challenges. Will show you where we are and where we are heading and set the stage to brainstorm how we can be better
- Kids First is unique in bringing together 2 segregated sciences but also community who think differently about phenotypes. Challenges and opportunities
- Phil lUpo paper substantiated this link between cancer and birth defects. We are trying to empower in the context of real time research through platforms and interfaces

- Talk about how data come into the platform.
    - Thining as a platform user.
    - My interpretation.
    - When we bring these landscapes together, tension. Cancer focus is on diagnosis and clinical outcomes in ways that most of the non-genomic data is on the MONDO format side of structuring data. Congenital birth defects researchers have spent an inordinate amount of time in terms of genomic… syndromic context. Changes that have many contexts for human representation
        - Tension comes when trying to empower this connection in part because of the source of data and how that data comes into platforms.
        - Cancer and development.
        - COSMIC
            - Somatic mutations in cancer. Curated lists of bonified cancer genes
        - and DECIPHER
            - Dedicated effort towards phen landscape. Can browse and download data associated with particular genes and phenotypes
        - Take two lists together, a large portion of curated cancer genes overlap with development context
        - Can do a small experiment. Take overlapping context. Do something similar with HPO. extra all HPO terms from all KF datasets and look to see which are overlapping with curated
        - Take HPO terms… genes associated with development and compare with cancer and ask within ped cancer cohorts, what do we see.
        - Challenging for me to describe and even more challenges for a user to hop around data
        - Challenges to undertake
- KF Phenotypes
    - Besides community … where data comes from and comes in.
    - HPO IDs by KF studies. Wide range
    - KF Diagnoses (MONDO) - distribution changes and tilt toward cancer landscape. Cancer diagnosis also end up having many others.
    - In the Portal to highlight…
        - Explore data
        - Neuroblastoma data from KF and targe. Almost 200 subjects.
        - Scoliosis. As observed phenotype.
            - 2 cancer cohorts. Clinical data pull from electronic health record. ICD-9 or ICD-10 then curated to be HPO.
            - EMR and data from research process
        - Screen shot examples, trying to implement user challenges
        - Instead of looking for direct match, want to empower (similarities)
            - Number of patients that fit into each

        - As a researcher. Want to start a higher level to see what data you want to bring into view.
    - Can start getting really rich cohorts to prioritize gene lists.
    - Vincent at St Justine Much more to do, additional visualization.
    - MONDO terms… in blue cancer studies.
        - Deep phenotypes.

PCGC

- Not speaking on behalf of PCGC but as someone who (think about) structure of these data.
- ...ToF might miss detailed anomalies. Very well understood and described.
    - The way we approached this for 15-20 years before PCGC. We knew CHD, forced into a primary diagnosis to bin and then easier later on to assign them to a category. Pick one only. Systematic structure each of the following…
        - Need a hierarchy of priorities.
        - PCGC does provide where data was obtained from
        - Hybrid approach to say for each cardiac segmant known/unknown. Fyler codes - good for being comprehensive, but doesn't require you to mark upon each segment.
        - Example of what you end up pulling out. Laundry list of diagnoses, not hierarchy, and maybe not consistent (not sure if reported or not).
            - Personal questions about fyler codes - not sure how consistent it was between groups.
        - Another coding system developed by international nomenclature society. Other clinical registries. With WHO for 11th ICD. Tree structure all the way down. Considers all components of heart. Improtatn for coding and outcome analyses.
            - May not be as important for granularity for genetics
            - Middle segment: this is how you find ToF. What could be helpful is detailed definitions of features of heart disease.

Joaquin DS-CHD

- Pan-omics
- Online participatn surveys. REDCap. Follow up on the phone. Deep chart reviews.
    - Clincial coordinators look at and curating medical record data. Many contradictions, different doctors. Very different nomenclatures.
    - www.trisome.org/explorer
    - Phenotyping for heart disease.
    - Categorize at level of class of condition.
    - CHD. Can get very granular and then for researcher

- Very important; these conditions rarely happen on their own in DS. can select the conditions i want to look at.
- Realtime get a picture of what a cohort is looking like.
- How this impact analysis of omics datasets. Proteome dataset. Assay for somatic markers in the blood.

….
Nara Sobreira

- https://phenodb.org/

- Took the information the DRC was requesting and formatted the database accordingly so she can export with no changes.

Mary Marazita

- No ontology for these phenotypes in HPO. FaceBase created an ontology to cover human, mouse, zebrafish.
- 2000 proband trios
- 3D facial imaging. Rating system, variables added to the dataset.
- Sub clinical phenotypes, physical features within range of normal (muscle discontinuity). Anyone can have. Hypothesis increased prevelance in unaffected reletaives compared to controls. Not a new concept. Left side, … hope can assist in heterogeneity.
- Harmonizing phenotypes will be important
- https://www.facebase.org/ocdm/

Carol Wise

-