Clinical and Phenotypic Data Capture

Co-chairs: David Hansen & Melissa Haendel
Clinical and Phenotypic Data Capture

• Key problems:
• Crossing the Clinical/Research divide
  • Clinicians are looking for answers about their patients
  • Researchers/informaticians want to give the answers – but need really good phenotype data
• Two aspects of this data conundrum:
  • Clinical and non-genomic data capture – need to capture rich, high quality data
  • Clinical data exchange – in a way that is computable and interoperable
Clinical and Phenotypic Data Capture

Previous Task Teams – doing great stuff!

• Meta-data
• Phenotype Ontologies
• e-Health
Clinical and Phenotypic Data Capture

Meta-Data Task Team

• ArrayMap
  • cancer genome array data

• Beacon+
  • built on ArrayMap, incorporating structural genomic variants

• Biosamples
  • 5 million samples, linking to EMBL-EBI data
Clinical and Phenotypic Data Capture

Phenotype Ontologies Task Team

• **HPO**
  • Already in use in MME, Beacon, ...

• **NCI-t**
  • Cancer ontology interoperable with HPO

• **Monarch initiative**
  • Exomiser
  • Patient Archive
  • Phenotypes patient app (emergent)
How do all these ontologies fit into our notion of disease?

Phe

• HPO
• Already in use in MME, Beacon, ...

• Monarch

Large scale data integration

Disease 1

Disease 2

Genes

Environment

Phenotypes
Clinical and Phenotypic Data Capture

e-Health Task Team

• FHIR for genomics / Sync for Science
• FHIR Family health history pilot
  • Best practice for capturing Family Health history
• ..
Clinical and Phenotypic Data Capture

What the Driver Projects told us

• Interaction with Electronic Health Records
  • Capturing phenotypic information for sharing
  • Reports and decision support
  • Getting standardized data out of the EHR

• Need for
  • Standard ontologies for data capture
  • Standards for meta-data to describe data
  • minimal data set to be collected for specific diseases
Key issues we see as a work stream

• We want to be use case driven
  • identify use cases which solve a problem for multiple Driver Projects

• We’re keen to avoid duplication of effort
  • Especially across the work streams

• Avoid building from scratch
  • The previous task teams have given us plenty to work with

• Need to broaden our engagement – g2mc etc
Clinical and Phenotypic Data Capture

- Workstream co-leads
  - David Hansen, Australian Genomics Health Alliance / CSIRO
  - Melissa Haendel, Monarch Initiative / OHSU
- Steering committee – capture the enthusiasm from the previous task teams
  - Grant Wood
  - John Mattison
  - Melanie Courtot
  - ...
Clinical and Phenotypic Data Capture

Sub-groups and the previous task teams

• Clinical (and non-genomic) data representation
  • HPO, NCI-T (from Phenotype); SNOMED-CT, ICD etc (from e-Health)
  • Family health history (from e-Health); Bio-sample meta data (meta data)
  • Machine learning and NLP – getting to clinical data from notes (Phenotype, e-Health, meta-data)

• Clinical data exchange
  • Phenopackets (from Phenotype)
  • FHIR for Genomics (from e-Health)

• Education / outreach
  • ??
Activities and milestones / Engagement with Driver projects

- **Driver Project (+) survey**
  - Build on the survey from the Global Leaders in Genomics Medicine conference in 2013
  - Interviews with DP leads
  - Followed by an electronic survey
  - Aim is to identify common use cases across the Driver Projects
Clinical and Phenotypic Data Capture

Activities and milestones / Engagement with Driver projects

• Current initiatives which will continue
  • Cancer ontology integration with HPO
  • FHIR Genomics and FHIR Family Health History Pilot are underway
  • Minimum data sets are starting to be shared between the national initiatives – and could be a activity and milestone within this workstream
  • Phenopackets exchange and FHIR interoperability
Q&A