Using human genetic evidence to identify and develop medicines that make a difference in people’s lives

March 22, 2016
Meg Ehm on behalf of Genomic Resources for Drug Discovery Consortium
Background

- Using human genetic evidence to identify and develop medicines.
- Work is built on technology advances, understanding of biological pathways, genomics/genetics, progress in gene editing, informatics and emergence of large scale human health databases.

Previous efforts assess the impact of genetic variation on clinical phenotypes have used population studies & genetic analysis consortiums

- Phenotypes are reduced to a “common denominator” and rarely have longitudinal information.
- Very few collections have drug usage and/or drug response information.
- Research is slow
- Many conditions haven’t been studied
Example – Predicting Clinical Response

- Identified low-frequency missense variant (Ala316Thr) in *GLP1R* gene, the target of *GLP1R* agonists:

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>N</th>
<th>Beta (95% CI)</th>
<th>Pvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fasting glucose</td>
<td>39469</td>
<td>-0.15 (-0.20, -0.11)</td>
<td>2.2x10^{-10}</td>
</tr>
<tr>
<td>2h glucose</td>
<td>39600</td>
<td>0.04 (-0.02, 0.10)</td>
<td>0.15</td>
</tr>
<tr>
<td>Fasting insulin</td>
<td>55854</td>
<td>0.02 (-0.01, 0.05)</td>
<td>0.24</td>
</tr>
</tbody>
</table>

Beta - SDs per minor-allele
Association with clinical endpoints

We investigated the association of the variant with other disease outcomes:

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Odds ratio (95% CI)</th>
<th>N cases</th>
<th>N controls</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 2 diabetes</td>
<td>0.83 (0.76, 0.91)</td>
<td>25868</td>
<td>122393</td>
<td>9.4x10^{-5}</td>
</tr>
<tr>
<td>Coronary heart disease</td>
<td>0.93 (0.87, 0.98)</td>
<td>61846</td>
<td>163728</td>
<td>9.2x10^{-3}</td>
</tr>
<tr>
<td>Pancreatic cancer</td>
<td>1.15 (0.82, 1.61)</td>
<td>4987</td>
<td>8627</td>
<td>0.43</td>
</tr>
</tbody>
</table>

March 4, 2016 press release

Victoza® significantly reduces the risk of major adverse cardiovascular events in the LEADER trial
Wish to study genetic influence of potential drug target genes on medically relevant traits.
Potential of EMR-linked Biobanks to Influence Drug Discovery & Development

- EMR-linked biobanks enable study of clinical characteristics and healthcare utilization:
  - conditions not studied – RSV hospitalization
  - disease trajectories – heart failure -> pulmonary edema -> renal conditions

- Pharma companies envision using EMR-linked biobanks to:
  - Recruit patients for further study by phenotype or genotype.
  - Study of the effects of gene variants on clinical phenotypes which can anticipate effects of drug treatment.
  - Study the effect of gene variants on the trajectory of disease and use of the health care system supporting indication optimization.
We see value in integrated medical and genomic resources to identify/prioritize targets and understand drug response:

- Comprehensive and diverse health data
- Comprehensive genetic data
- Ability to recontact patients
- Biological sample availability
- Informatics
Genomic Resources for Drug Discovery Consortium

- View developing a research enabled environment to facilitate work is a pre-competitive activity and will require more resources than any one company can provide.

- Agree it is critical to engage early in the design and development of genomic/EMR resources to leverage unique insight into use of resources for drug discovery and development questions.

- Agree that harmonization of multiple resources (ethnicities, health settings, …) will be needed to fully realize value – no one resource will be enough.
EMR-Linked Biobanks for Drug Discovery

Basic Research and Pre-Clinical Studies

Evidence that target is associated with outcomes

Evidence that target contributes to disease progression.

Evidence that modulating target impacts disease & progression.

Simulate/predict outcomes/progression when modulating drug target.

Use of systems biology/systems pharmacology with genetic, frequently sampled clinical, biomarker data.

Differentiating

Minimal
EMR-Linked Biobanks for Drug Discovery & Development

Human Clinical Trials – Pragmatic Trials

Recruit participants using phenotype, genotype, diagnosis for studies of all types with future unspecified analyses.

Recruit participants with different rates of progression, biomarker/genomic profiles & exposures.

Use trial ready population to efficiently recruit & complete clinical trial.

Minimal

Differentiating

Game-changing
Emerging Requirements for Genomic Resources

Minimal

- GWAS array
- Prescribing info
- Demographics
- Clinical Genetics
- Basic genotype QC & association testing
- Disease state, ICD-9/10 codes
- Standard EHR format
- 100,000 participants

Differentiating

- Comprehensive phenotyping
  - RNAseq, WES data
  - Lab, imaging data
  - EHR formal shared with other institutions
- 500,000 proactive participants

Game changing

- Tissue RNAseq data
- WGS data
- Mobile Health data
- Exposure information
- Plug data directly into predictive modeling platforms & analytics
- Cognitive and ML tools for phenomics
- Capture of real time updates and participatory information
- Systems connected to sysbio/sylypharmacology data resources
- Integration of other HCO data types
- Automated phenotype derivation
- Temporally based analytics
- 1M proactive participants with providers open to trials

HCO Data

- NLP capabilities
- Extensive ontologies
- Cloud computing environment
- Search by genotype/phenotype
- DNA, Plasma/serum, Paxgene (RNA)
- Minimally invasive specimens collected during normal clinical care & on demand
- Longitudinal biospecimens

Informatics capabilities

Bio-specimens & patient recall

Size & scope

• Recall by genotype or phenotype
• More invasive sample types collected during the course of care
• Recall for clinical evaluation
• Wide range of sample types available on demand

• NLP capabilities
• Extensive ontologies
• Cloud computing environment
• Search by genotype/phenotype
• DNA, Plasma/serum, Paxgene (RNA)
• Minimally invasive specimens collected during normal clinical care & on demand
• Longitudinal biospecimens

• Tissue RNAseq data
• WGS data
• Mobile Health data
• Exposure information
• Plug data directly into predictive modeling platforms & analytics
• Cognitive and ML tools for phenomics
• Capture of real time updates and participatory information
• Systems connected to sysbio/sylypharmacology data resources
• Integration of other HCO data types
• Automated phenotype derivation
• Temporally based analytics
• 1M proactive participants with providers open to trials

• NLP capabilities
• Extensive ontologies
• Cloud computing environment
• Search by genotype/phenotype
• DNA, Plasma/serum, Paxgene (RNA)
• Minimally invasive specimens collected during normal clinical care & on demand
• Longitudinal biospecimens

• Tissue RNAseq data
• WGS data
• Mobile Health data
• Exposure information
• Plug data directly into predictive modeling platforms & analytics
• Cognitive and ML tools for phenomics
• Capture of real time updates and participatory information
• Systems connected to sysbio/sylypharmacology data resources
• Integration of other HCO data types
• Automated phenotype derivation
• Temporally based analytics
• 1M proactive participants with providers open to trials
Consortium Deliverables: Research-Enabled Environment

- Capability to assess if target is associated with clinical conditions & drug response enabling prediction of possible outcomes.

- Consent and infrastructure to identify and recontact patients for clinical trials of all types using genetic and clinical information

- Access model that enables member companies to easily perform scoping and analysis

- Informatics capabilities that enable harmonization of phenotypes and meta-analysis of results across studies.

- Option to leverage existing or generate additional data for all or sub-sets of the resource (e.g. biomarkers, metabolomics)
Next Steps

- Assemble parties that are interested in participating in the consortium organization/planning phase.

- Work together to answer the question:
  - What are the models that would bring together EMR-linked biobank resources to enable us to achieve our goals of a research enabled environment?

- We intend to develop a business plan and funding model which will realize our goals by the end of 2016.
Acknowledgments

Aaron Day-Williams
- John Carulli
- Sally John
- Michelle Penny
- Hank Wu

Eisai
- Janna Hutz
- Nadeem Sarwar

Lilly
- Laura Nisenbaum
- Sreekumar Pillai

gsk
- Lon Cardon
- Matt Nelson

Takeda
- Eric Lai
- Daniel Kemp
- Jatinder Kaur

Biogen

MERCK
- Rebecca Blanchard
- Caroline Fox
- Robert Plenge
- Heiko Runz

Pfizer
- Nan Bing
- Craig Hyde
- Katrina Loomis
- Cliona Molony
- Serena Scollen
- Jemma Wilk

Arthur Holden – Co-Chairman of Genomic Resources for Drug Discovery Consortium Organizing Committee

GLP1R Example
- Daniel Frietag
- Robert Scott
- Nick Wareham
Genomics Resources

**Minimal**

- 100,000 participants
- GWAS array
- Prescribing info
- Demographics
- Basic genotype QC & association testing
- Disease state, ICD-9/10 codes
- Standard EHR format

**Differentiating**

- 500,000 proactive participants
- Comprehensive phenotyping
  - RNAseq, WES data
  - Lab, imaging data
  - EHR formal shared with other institutions

**Game changing**

- 1M proactive participants with providers open to trials
- Tissue RNAseq data
- WGS data
- Mobile Health data
- Exposure information
- Plug data directly into predictive modeling platforms & analytics
- Cognitive and ML tools for
  - Phenomics
  - Capture of real-time updates
  - NLP capabilities
  - Extensive ontologies
  - Cloud computing environment
  - Search by genotype/phenotype

**HCO Data**

- DNA, Plasma/serum, Paxgene (RNA)
- Minimally invasive specimens collected during normal clinical care & on demand
- Longitudinal biospecimens

**Bio-specimens & patient recall**

- 500,000 proactive participants
- Recall by genotype or phenotype
- More invasive sample types collected during the course of care
- Recall for clinical evaluation
- Wide range of sample types available on demand
- Systems connected to sysbio/sypharmacology data resources
- Integration of other HCO data types
- Automated phenotype derivation
- Temporally based analytics
- NLP capabilities
- Extensive ontologies
- Cloud computing environment
- Search by genotype/phenotype

**Size & scope**

100,000 participants

Informatics capabilities

Bio-specimens & patient recall
Genomics Resources

Minimal

- 100,000 participants connected to HCO database
- GWAS array
- Prescribing info
- Demographics/family history
- Clinical Genetics
- Basic genotype QC & association testing
- Disease state, ICD-9/10 codes
- Standard EHR format in HCO

Differentiating

- 500,000 proactive participants connected to HCO database
- Comprehensive phenotyping
  - RNAseq data
  - WES data
  - LOINC data
  - Imaging data
  - EHR formal & implementation shared with other institutions

Game changing

- 1M proactive participants
- Facilitated participation of providers & patients in clinical trials
- Tissue RNAseq data
- WGS data
- Mobile Health data
- Complete suite of "wellness"
- Exposure information
- Subset of population deeply phenotyped
- Plug data directly into predictive modeling platforms & analytics
- Cognitive and ML tools for phenomics
- Capture of real time updates and participatory information
- Inter-operability with other bioresources
- iPSC capabilities
- Recall for comprehensive clinical evaluation
- Wide range of sample types available on demand
- Recall by genotype or phenotype
- More invasive sample types collected during the course of care
- Microbiome capabilities

HCO Data

- NLP capabilities
  - Extensive ontologies
  - Cloud computing environment
  - Search by genotype/phenotype
  - EHR search within HCO
  - Standardized variables, disease data

Informatics capabilities

- DNA, Plasma/serum, Paxgene (RNA)
- Minimally invasive specimens collected during normal clinical care (e.g. urine) & on demand
- Longitudinal biospecimens
- Tumor/normal paired samples

Bio-specimens & patient recall

Size & scope
## Questions RWD and Genetics can address

<table>
<thead>
<tr>
<th>Discovery – Basic Research / Pre-Clinical Studies</th>
<th>Development – Human Clinical Trials / Pragmatic Trials</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Game-changing</strong></td>
<td></td>
</tr>
<tr>
<td>– Can we simulate/predict possible outcomes (efficacy and safety) when we modulate a drug target?</td>
<td>– Use trial ready population to efficiently recruit and complete clinical trial</td>
</tr>
<tr>
<td></td>
<td>– Detailed natural history data</td>
</tr>
<tr>
<td></td>
<td>– Imaging data</td>
</tr>
<tr>
<td></td>
<td>– Cognitive tests on regular basis</td>
</tr>
<tr>
<td>– Use of systems network biology /systems pharmacology using genetic data, frequently sampled biomarkers and simulated medicine use.</td>
<td></td>
</tr>
<tr>
<td><strong>Differentiates</strong></td>
<td></td>
</tr>
<tr>
<td>– What targets or pathways contribute to disease pathogenesis, progression, complications and recurrence?</td>
<td>– Use resource to identify participants with different rates of progression, biomarker profiles, clinical genetic diagnoses, genetic/genomic (germline &amp; somatic) profiles and exposures (family history, taking specific medications, ...) which contribute to an efficient clinical trial.</td>
</tr>
<tr>
<td></td>
<td>– What are the biomarkers that can track early signs of efficacy in the clinic?</td>
</tr>
<tr>
<td></td>
<td>– Which biomarkers, clinical phenotypes will select patients most likely to develop severe conditions?</td>
</tr>
<tr>
<td></td>
<td>– Which biomarkers are prognostic?</td>
</tr>
<tr>
<td>– What is the evidence that modulating a specific drug target will prevent disease, slow progression or prevent complications?</td>
<td></td>
</tr>
<tr>
<td>– What is the evidence that diseases (i.e. Cancers) could benefit from alternative or combination treatments?</td>
<td></td>
</tr>
<tr>
<td><strong>Minimal</strong></td>
<td></td>
</tr>
<tr>
<td>What is the evidence that a given target/pathway/molecule is associated with clinical conditions using cross sectional data.</td>
<td>– Need to be able to recruit subjects with specific characteristics including phenotype, genotype (germline, somatic), diagnosis for trials/specialized studies with future unspecified analyses.</td>
</tr>
<tr>
<td>Size and Scope</td>
<td>HCO data</td>
</tr>
<tr>
<td>----------------</td>
<td>----------</td>
</tr>
<tr>
<td><strong>Game changing</strong></td>
<td>• 1M participants • Facilitated participation of health care providers and patients in research and clinical trials</td>
</tr>
<tr>
<td><strong>Differentiates</strong></td>
<td>• 500,000 participants • Demographic similarity to target population • Pragmatic participants • Ability to contact and recruit subjects into clinical trials</td>
</tr>
<tr>
<td><strong>Minimal</strong></td>
<td>• 100,000 genotyped participants connected to larger HCO database • Consent for specific genotyping &amp; future unspecified research</td>
</tr>
<tr>
<td>Name</td>
<td>Objective</td>
</tr>
<tr>
<td>------</td>
<td>-----------</td>
</tr>
<tr>
<td>COPD target validation</td>
<td>Evaluate if variation in respiratory infection/COPD drug targets is associated with susceptibility to bacterial and viral infections.</td>
</tr>
<tr>
<td>GWAS of dermatology conditions</td>
<td>Identify genetic risk factors associated with disease risk or progression for understudied dermatology conditions.</td>
</tr>
<tr>
<td>Progression of renal diseases</td>
<td>Identify and study rapid and slow progressors for renal diseases of interest.</td>
</tr>
<tr>
<td>Study of factors influencing the preservation of muscle mass</td>
<td>Identify factors influencing muscle mass for patients with COPD</td>
</tr>
<tr>
<td>Feasibility of PGx study of recently marketed GSK compound</td>
<td>Evaluate the #s of patients taking recently marketed GSK compound that could be enrolled into pharmacogenetic studies.</td>
</tr>
<tr>
<td>Heart failure (HF) sub-types and disease</td>
<td>Identify and study typical HF patients with reduced ejection fraction (EF)</td>
</tr>
</tbody>
</table>