Studying Rare Phenotypes

Opportunities Presented by the Cross Cohort Collaboration
Multi-cohort collaborations for uncommon conditions

Assumptions:

• Low prevalence/incidence conditions are difficult to study, primarily because of identification and recruitment challenges.

• Identification of affected participants from community representative cohorts avoids some selection biases that may accompany targeted recruitment.

• Hypothesis-driven multi-cohort consortia may allow for the power necessary to draw conclusions about uncommon conditions.
Advantages of Multi-cohort collaborations

Assumptions:

• Established cohorts offer distinct advantages for both the investigator and funder
  • Participants are
    • already enrolled
    • phenotyped (at least to some extent).
    • may have relevant biospecimens
    • may have longitudinal data/specimens
    • the above are particular advantages over cohort assembly from EMRs.
  • Coordinating center functions are established
  • Track record of productivity and prior ancillary funding
Rare Diseases vs. Rare Phenotypes

• Very hard to study diseases of low prevalence/incidence in traditional cohorts
  • Other methods more suited for this
    • Rare disease networks (i.e. CF)
    • Linked, EMR-based methods

• However, excellent opportunity to study uncommon phenotypes placing individual at risk for - or protection from - a common, well-described, easier-to-adjudicate disease (i.e., CVD)
Importance of Disease Hereterogeneity

One of the central findings of MESA
Distribution of CAC by RF Burden

% of individuals in each CAC group

<table>
<thead>
<tr>
<th>RF</th>
<th>CAC &gt; 100</th>
<th>CAC 1-100</th>
<th>CAC 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0RF</td>
<td>12%</td>
<td>20%</td>
<td>68%</td>
</tr>
<tr>
<td>1RF</td>
<td>20%</td>
<td>25%</td>
<td>55%</td>
</tr>
<tr>
<td>2RF</td>
<td>27%</td>
<td>29%</td>
<td>45%</td>
</tr>
<tr>
<td>≥3RF</td>
<td>34%</td>
<td>31%</td>
<td>35%</td>
</tr>
</tbody>
</table>

P < 0.001

Hard CHD Event Rates (per 1,000 person-years) by CAC score according to Risk Factor Burden

March 7, 2015
The extremes design permits a focus on patients who are of interest with respect to identifying new disease pathways and therapeutic targets. To obtain similar statistical power, the number needed to phenotype is 4 times greater when using a population-based approach compared with an approach based on targeted phenotyping of individuals at the extremes.

Zhang et al  Genetic Implication of a Novel Thiamine Transporter in Human Hypertension, Journal of the American College of Cardiology 2014
Lanktree et al  Extremes of Unexplained Variation as a Phenotype: An Efficient Approach for Genome-Wide Association Studies of Cardiovascular Disease, Circulation 2010
Rana et al  Population-Based Sample Reveals Gene–Gender Interactions in Blood Pressure in White Americans. Hypertension 2006
Lipids and Atherosclerotic CVD

Importance of Disease Heterogeneity
LDL Cholesterol and CAC

Rate of CVD per 1,000 person-years

LDL-C (mmol/L)

<1.80 1.80 to 2.57 2.58 to 3.34 ≥3.35

0 5 10 15 20 25 30

≥100 1 to 99 CAC

4.7 26.1 26.1 28.9

14.0 10.5 10.2

3.5 2.6 3.0

6.7 2.6 3.0

10.5 2.6 3.0

2.6 3.0 3.0

Martin, Blaha, et al.
Circulation. 2014.
LDL Cholesterol Rare Phenotypes

• Lifetime exposure to low LDL and high risk
  • What is optimum LDL?
  • <70 mg/dL or <50 mg/dL (N=345 and N=67 in MESA)

• Exposure to very high LDL and very low risk
  • What is familial hypercholesterolemia (FH)?

• Opportunity for deep phenotyping
  • Genetic factors
  • Inflammatory factors
  • Lipoprotein size and function
Age and Atherosclerotic CVD

Importance of Disease Heterogeneity
The prevalence of coronary artery calcium in asymptomatic patients across age groups
All-cause Mortality in Different Age Groups Stratified by Increasing CAC

Mortality/1000 person years

Age

- <45
- 45-54
- 55-64
- 65-74
- ≥75

CAC=0
CAC 1-100
CAC 101-400
CAC>400

Rare Phenotypes in Aging

- Advanced disease in young patients
  - Apparent “malignant” form of the disease
- “Healthy Agers” free of subclinical CVD >75 years old
  - Insights toward aging, interaction of age and risk factors

- Opportunity for deep phenotyping
  - Genetic factors
  - Other aging markers, like telomere length
  - Inflammatory factors
  - Other imaging markers of vascular compliance and function
Brief, Incomplete List of Other Rare Phenotypes

- Questionable importance of isolated low HDL
  - N=158 in MESA (N=781 with “optimal” lipid profiles)
- Long-term subclinical disease non-progressors
- CHD events in the absence of overt or subclinical disease
  - Increasing interest in MINOCA (“MI with no [obstructive] coronary artery disease”)
- Unexpected very low/high bone density
- Accelerated/protection from sarcopenia
Practical Considerations

• Volunteer recruitments – skew prevalence or identity of rare phenotypes?
• Rare phenotypes would have to be exquisitely defined
• Requisite data may not be available in all cohorts
• Data analysis would likely need to be done at a single, central location
• Funding would be necessary to gain mechanistic insights from “deep phenotyping”
  • Preliminary data accumulation prior to a funding request may be attractive and feasible with existing resources
• An effective prime mover/PI for each project would be essential
Thank you!

Michael J. Blaha (Johns Hopkins)
Eric Orwoll (OHSU)