Health and Retirement Study: Biological Markers
PAA Biomarker Meeting
March 30, 2016

David Weir, PI
Jessica Faul, Co-Investigator
Institute for Social Research
HRS

- Began collecting dried blood spots and salivary DNA in 2006
- Half-sample rotation: half assigned to biomarkers starting in 2006, half in 2008
- New cohort added in 2010 was also split across the half-sample
- Goal was to build a longitudinal biomarker resource
- And a genetic repository
## HRS DBS Collection by wave

<table>
<thead>
<tr>
<th>Sample</th>
<th>First eligible</th>
<th>2006</th>
<th>2008</th>
<th>2010</th>
<th>2012</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel</td>
<td>2006</td>
<td>6735</td>
<td>5709</td>
<td>5133</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>6329</td>
<td></td>
<td></td>
<td>4958</td>
<td></td>
</tr>
<tr>
<td>New cohort</td>
<td>2010</td>
<td></td>
<td>2073</td>
<td></td>
<td></td>
<td>2373</td>
</tr>
<tr>
<td></td>
<td>2012</td>
<td></td>
<td></td>
<td></td>
<td>2285</td>
<td></td>
</tr>
</tbody>
</table>
### HRS DBS Collection: Longitudinal patterns

<table>
<thead>
<tr>
<th>Sample</th>
<th>First eligible</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel</td>
<td>2006</td>
<td>2,196</td>
<td>2,295</td>
<td>3,597*</td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>2,735</td>
<td>4,276</td>
<td>[2016]</td>
</tr>
<tr>
<td>New cohort</td>
<td>2010</td>
<td>1,204</td>
<td>1,621</td>
<td>[2016]</td>
</tr>
<tr>
<td></td>
<td>2012</td>
<td>2,285</td>
<td></td>
<td>[2016]</td>
</tr>
</tbody>
</table>

* Data not yet released
By the time this grant cycle ends in 2017

• Expect ~20,000 people to have given at least one DBS sample
• Expect ~7,000 old panel cases with three consecutive waves of DBS
• Expect ~3,000 new cohort cases with two consecutive waves of DBS
Assays

• 2006-14
  – HbA1c, CRP, cystatin-C, total and HDL cholesterol

• 2016
  – Same plus IL-6

• We await this network’s validation of the UW IL-6 assay, but we will also be able to validate in our own sample
PLAN FOR HRS 2016
VENOUS BLOOD COLLECTION

Key partners:
Hooper Holmes (phlebotomy)
University of Minnesota (lab)
Advantages of whole blood over DBS

- Assay reliability
- Range of analytes
- Storage
Motivation in HRS

1. Improve biological assessment of health/disease status
2. Continue search for the elusive biological pathways connecting social experience to health
3. Explore experimental markers of aging
1. Clinical markers

• Much broader set of standard markers that can be reported to respondents to encourage participation
• Markers that harmonize to ELSA and other studies
• Major organ system functioning
• (Alzheimer’s)
2. Pathways

- Gene expression
- Modifications at cellular level, and specific to cell type
- Focus on the immune system, which is significantly managed in blood
- Markers of inflammation and immune function
- Cryopreserve cells for future analysis
- RNA (paxgene)
3. Experimental markers

- DNA methylation
- Telomere length
- mtDNA copy number
Overview of plan for blood collection and processing

- Consent must be obtained at time of HRS 2016 interview
- Contact info for consented respondents will be sent to Hooper Holmes, a contract phlebotomy service selected by competitive bid
- Attempt blood draw within 4 weeks of core interview
- Fasting will be recommended and preferred but not required
  - Most of our immediate assays do not require fasting but it will enhance potential of stored samples
- We project 9850 collections (proxy and NH excluded)
- Prepare samples in field and ship overnight to lab
  - Centrifuge
  - Cold / room temp shipping
Overview of plan for processing at the lab

- Minnesota to receive samples within 24-48 hours
- Perform assays that must be done immediately
- Freeze serum and plasma
- Cryopreserve white blood cells
- Over course of 2016/17 perform our selected assays on frozen samples
- Retain remaining material (about half) in repository for future use
HRS Venous Blood Collection

(50.5 mL Venous Blood in six tubes)
## Revised Sample/Assay Plan

<table>
<thead>
<tr>
<th>Panel Sample</th>
<th>Innovative Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=7850</td>
<td>N=2000</td>
</tr>
<tr>
<td>Metabolic Panel</td>
<td></td>
</tr>
<tr>
<td>Lipid Panel</td>
<td></td>
</tr>
<tr>
<td>CBC</td>
<td></td>
</tr>
<tr>
<td>High sensitivity CRP (hsCRP)</td>
<td></td>
</tr>
<tr>
<td>Ferritin (FRTN)</td>
<td></td>
</tr>
<tr>
<td>IGF-1,</td>
<td></td>
</tr>
<tr>
<td>DHEA-S</td>
<td></td>
</tr>
<tr>
<td>Cystatin C</td>
<td></td>
</tr>
<tr>
<td>Vitamin D (25 Hydroxy)</td>
<td></td>
</tr>
<tr>
<td>Cytokine panel</td>
<td></td>
</tr>
<tr>
<td>Flow cytometry (cryopreserved cells)</td>
<td></td>
</tr>
<tr>
<td>CMV seroprevalence</td>
<td></td>
</tr>
<tr>
<td>B-type natriuretic peptid (NT-proBNP)</td>
<td></td>
</tr>
<tr>
<td>mtDNA copy number</td>
<td></td>
</tr>
<tr>
<td>DNA Methylation</td>
<td></td>
</tr>
<tr>
<td>Homocysteine</td>
<td></td>
</tr>
<tr>
<td>Telomere length</td>
<td></td>
</tr>
<tr>
<td>P16</td>
<td></td>
</tr>
</tbody>
</table>

1. Traditional biochemical /harmonized marker
2. Immune system and inflammation marker
3. Innovative aging marker
Enabling Genetic-based Social Science Research Using HRS

- Over 130 approved research projects at dbGaP
- Mostly bio-medical, narrowly focused
- Full genetic database: large, unwieldy, restricted data
- We are working on data products that will be easier to use in social science
- And not a risk for respondent identification
Beyond Candidate Genes

- Complex health outcomes or behaviors of interest to the research community are often highly polygenic, or reflect the aggregate effect of many different genes (Visscher, Hill, & Wray 2008)

  - Individuals fall somewhere on a continuum of genetic risk that reflects small contributions from many genetic loci (Gibson 2012)

  - Genetic loci influencing the etiology of complex phenotypes have low penetrance—i.e. one single gene does not produce a symptom or trait at a detectable level (ibid)

- Use of single genetic variants or candidate genes may not capture the dynamic nature of more complex phenotypes

- Providing even a few direct genotypes is a confidentiality risk
Polygenic Risk Score (PRS)

• Using the published effect sizes from a GWAS, researchers can construct a polygenic risk score (PRS)

• A PRS aggregates millions of individual loci across the human genome and weights them by the strength of their association to produce a single quantitative measure of genetic risk

  • PRS: Weighted average across the number of SNPs ($n$) of the number of reference alleles $x$ (0,1 or 2) at that SNP multiplied by the score for that SNP ($\beta$):

$$PRS_i = \sum_{j=1}^{n} (\beta_j x_{ij})$$

Source: Schmitz & Conley 2015
Attractive features of a PRS

• Hypothesis-free measures
  • Ex ante knowledge about the biological processes involved is not needed to estimate a score for a particular phenotype
  • Allows researchers to explore how genes operate within environments where the biological mechanisms are not fully understood (Belsky & Israel 2014)

• Maximizes statistical power when modeling gene-environment (G x E) interactions
  • PRSs use the raw statistical power from huge consortia to generate one measure of genetic risk
  • The statistical power needed to model a candidate G x E study for biologically distal, social phenotypes is not possible in social surveys that contain the level of detailed information about respondents that motivates GxE inquiry (Belsky, Moffitt, & Caspi 2013)
Issues to consider when constructing PRSs

• Constructing a PRS based off of a GWAS your study is in
  • Mathematically remove your study (see attached PowerPoint)
    • Problems: dealing with adjustment for genomic inflation, incorporation of stratified meta-analyses (e.g. sex-specific scores)
  • Approach consortia to rerun meta-analysis without your study

• If there is no seminal GWAS on a particular phenotype

• If your study does not have the related phenotype

• Size of the study, discovery, replication, meta-analysis or joint analysis for selection of appropriate beta weights or effect sizes

• HRS is currently in the process of designing the methods and pipeline to create publicly available PRSs
HRS - PRS in Development

• Will be a public data release (summer 2016)

• 1st priority (Data publicly available)
  Psychiatric Genomics Consortium (PGC)
  BMI - Locke (2015)
  Alzheimer’s Disease - IGAP
  Height - HapMap GIANT - (2014)
  Waist - GIANT
  Waist to hip - GIANT
  Educational attainment

• 2nd priority (Need additional data)
  Blood Pressure (ICBP)
  Smoking
  Longevity
  General cognition (Davies)
  Subjective Wellbeing
  Kidney Function (Cystatin C)
THANK YOU!

http://hrsonline.isr.umich.edu/