The M.U.R.D.O.C.K. Study
Measurement to Understand Reclassification of Disease Of Cabarrus/Kannapolis

Jessica Tenenbaum, PhD
and many, many others
MURDOCK Study—Measurement to Understand the Reclassification of Disease Of Cabarrus / Kannapolis

• A longitudinal health study identifying genomic linkages across major chronic diseases

• Gift from Mr. David H. Murdock, Chairman of Dole Foods and visionary behind the North Carolina Research Campus

• Organized as Horizons
  - Initiated in Duke facilities
  - Centered in Kannapolis/Cabarrus County, NC

• PI: Robert M. Califf, MD
  - Vice Chancellor of Clinical Research, Duke Medicine, and Director, Duke Translational Medicine Institute (DTMI)
The MURDOCK Study: A Framingham for the omics era

- Stratify patient populations
  - Guide therapy, preventive care
- Identify biomarkers
  - Develop diagnostic, prognostic tests
- Insights into underlying mechanisms
  - Identification of new treatment targets or strategies
- $35 million gift to Duke

PI: Rob Califf, MD

Executive Manager: Victoria Christian
Design: MURDOCK Study Horizons

**Horizon 1**
- Generated molecular biosignatures using legacy samples
- Mined associated clinical data
- Generated hypotheses to reclassify diseases:
  - Liver disease
  - Cardiovascular disease
  - Osteoarthritis
  - Obesity

**Horizon 2**
- Prospective cohort studies
  - Multiple sclerosis
  - Alzheimer’s disease
  - Physical functioning
  - Severe acne
  - Centenarians
- Drawing from subjects accessible via Horizon 1.5 registry

**Horizon 1.5**
- Building a Community Registry and Biorepository
  (Informatics and Sample Management)

**Horizon 3+**
- New waves of molecular data from diverse patient populations
- New cohort studies
- Data meta-analyses
- Multi-institutional and multi-national collaborations
- Measurement of public health impact
Horizon One

- Cardiovascular Disease
- Hepatitis C
- Obesity
- Osteoarthritis
Quick biology refresher

Genomics

Transcriptomics

Proteomics

Metabolomics

Image adapted from wikipedia

DNA

RNA

Protein

Active Protein

Metabolites
## Horizon 1 data types

<table>
<thead>
<tr>
<th></th>
<th>Obesity</th>
<th>Cardiovascular Disease</th>
<th>Hepatitis C</th>
<th>Osteoarthritis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proteomics- targeted</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proteomics- discovery</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transcriptomics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metabolomics- targeted</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Imaging</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Horizon 1.5: A Population Registry & Biorepository

- >10,300 participants and follow-up
- ~380,000 samples banked at LabCorp Kannapolis Biorepository

- Median follow-up: 2.6 years
- EHR coming soon
Horizon 1.5: Community registry

- 50,000 subjects
  - 35,000 volunteers
  - 15,000 randomized
- Collect blood, urine
- Clinical data, family history
- Annual follow up
- Re-contact ≤ 4 times/year
- Electronic health record, when available, through partnerships with local health providers
### Snapshot of MURDOCK Study Participants

<table>
<thead>
<tr>
<th></th>
<th>% MURDOCK</th>
<th>% Cabarrus Pop¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>66</td>
<td>51</td>
</tr>
<tr>
<td>African American</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>Hispanic</td>
<td>11.5</td>
<td>9</td>
</tr>
<tr>
<td>Age [median]</td>
<td>56 years</td>
<td>37 years</td>
</tr>
</tbody>
</table>

Self-reported data collected at baseline enrollment, N=10,315. ¹Estimates from U.S. Census Bureau Cabarrus County 2010 reports.
## Snapshot 3: Self-Reported Cancers & Other Diseases

<table>
<thead>
<tr>
<th>Disease</th>
<th>Number</th>
<th>% of Total Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oncology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast</td>
<td>242</td>
<td>2.5%</td>
</tr>
<tr>
<td>Colon</td>
<td>74</td>
<td>0.8%</td>
</tr>
<tr>
<td>Lung</td>
<td>33</td>
<td>0.3%</td>
</tr>
<tr>
<td>Prostate</td>
<td>157</td>
<td>1.6%</td>
</tr>
<tr>
<td>Cervical</td>
<td>111</td>
<td>1.1%</td>
</tr>
<tr>
<td>Melanoma</td>
<td>249</td>
<td>2.6%</td>
</tr>
<tr>
<td>Diabetes</td>
<td>1491</td>
<td>15.3%</td>
</tr>
<tr>
<td>High Cholesterol</td>
<td>3794</td>
<td>38.9%</td>
</tr>
<tr>
<td>Osteoarthritis (OA)</td>
<td>1688</td>
<td>17.5%</td>
</tr>
<tr>
<td>Depression</td>
<td>2310</td>
<td>23.7%</td>
</tr>
<tr>
<td>Other Mental Illness</td>
<td>340</td>
<td>3.5%</td>
</tr>
</tbody>
</table>

N=9746 MURDOCK Study Participants
### Lifestyle Habits: Diet, Past Year

#### Servings, fruit & veggies (#)

<table>
<thead>
<tr>
<th>Servings</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1%</td>
</tr>
<tr>
<td>2</td>
<td>21%</td>
</tr>
<tr>
<td>3</td>
<td>27%</td>
</tr>
<tr>
<td>4</td>
<td>17%</td>
</tr>
<tr>
<td>5</td>
<td>10%</td>
</tr>
<tr>
<td>6</td>
<td>5%</td>
</tr>
<tr>
<td>6-10</td>
<td>3%</td>
</tr>
<tr>
<td>11-20</td>
<td>0%</td>
</tr>
<tr>
<td>21-40</td>
<td>0%</td>
</tr>
<tr>
<td>Don't know</td>
<td>14%</td>
</tr>
</tbody>
</table>

#### Sugar sweetened beverages (#)

<table>
<thead>
<tr>
<th>Servings</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>46%</td>
</tr>
<tr>
<td>2</td>
<td>23%</td>
</tr>
<tr>
<td>3</td>
<td>11%</td>
</tr>
<tr>
<td>4</td>
<td>5%</td>
</tr>
<tr>
<td>5</td>
<td>2%</td>
</tr>
<tr>
<td>6</td>
<td>1%</td>
</tr>
<tr>
<td>6-10</td>
<td>1%</td>
</tr>
<tr>
<td>11-20</td>
<td>0%</td>
</tr>
<tr>
<td>21-40</td>
<td>0%</td>
</tr>
<tr>
<td>Don't know</td>
<td>9%</td>
</tr>
</tbody>
</table>

#### Read nutrition labels

- Always: 16%
- Very often: 30%
- Sometimes: 29%
- Rarely: 13%
- Never: 10%

N=9476 MURDOCK Study Participants
Geospatial data

• Expand upon work done in Durham, NC by Dr. Marie Lynn Miranda (now at U Michigan)
• Know address- pinpoint where subjects live
• Include environment in outcomes analysis
Electronic health records

• 85% of local population covered by 3 major regional health care providers
  – MURDOCK has collaborations with these providers
• Different EHR systems
• Identify set of variables to request
• Pilot project (n ≈ 100) to compare EHR values to patient-provided values
• Status: in progress
Horizon Two

Exemplars of Collaboration
Horizon 2: Prospective Cohort Studies

- Objectives:
  - Support prospective cohort studies, registries, interventional studies
  - Offer community participation in clinical trials
  - Discover, test, and validate biomarkers for various diseases

- Current cohort studies:
  - Multiple sclerosis
  - Centenarians project
  - Memory and cognitive impairment
  - Severe Acne
  - Healthy Aging
  - Gout and Hyperuricemia
  - Attitudes toward genetic testing
Pulling it all together: The MIDR

MURDOCK Integrated Data Repository
## Study search results

<table>
<thead>
<tr>
<th>Code</th>
<th>Study Name</th>
<th>Dept</th>
<th>PI</th>
<th>Start Year</th>
<th>Disease</th>
<th>Type</th>
<th>N</th>
<th>Data types</th>
<th>Sample Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV4</td>
<td>CVD outcomes</td>
<td>Cardiology</td>
<td>Kris Thompson</td>
<td>1989</td>
<td>CVD</td>
<td>Prospective Cohort, Case only</td>
<td>6700</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBV</td>
<td>RESPOND</td>
<td>Hepatology</td>
<td>Jane McDoe</td>
<td>2001</td>
<td>Hep B</td>
<td>Interventional cohort</td>
<td>1000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FIZ</td>
<td>Reflux</td>
<td>Gastrology</td>
<td>Georgia Kraut</td>
<td>2005</td>
<td>Sore tummy</td>
<td>Prospective Case-Control</td>
<td>1888</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OBE</td>
<td>Weight Watching</td>
<td>Medicine</td>
<td>Mary Olson</td>
<td>2003</td>
<td>Obesity</td>
<td>Interventional cohort</td>
<td>1000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
MURDOCK Integrated Data Repository (MIDR) Overview

- CRF/Clinical Data
- Electronic Health Records
- Consent Data
- Biospecimen Data
- Study Metadata
- Omics, Imaging Metadata

Discovery

- Data mining
- Disease stratification
- Predictive algorithms
- Pathway modeling
- Systems biology

Information retrieval

- Metadata queries
- Cohort selection
- Administrative reports
Example use case queries

• Current focus:
  – How many samples do we have from women over 50 who were later diagnosed with hypertension and who have a family history of cardiovascular disease?
  – How many male subjects do we have under the age of 50 who have been diagnosed with prostate cancer, for whom we have GWAS data?

• Future:
  – What pathways appear to be transcriptionally upregulated in subjects diagnosed with COPD versus healthy controls?
  – What is the representative genotype at SNP rs62653620 for males under 50 with prostate cancer versus health controls?
Governance

• MURDOCK participation is conditional upon agreeing to share resulting data
• Governance document defines policies for data access, publication, acknowledgement, etc.
• MURDOCK Leadership team reviews proposals for data use
Future work

- Add data sources, types to MIDR
  - omics, EHR, biospecimen, consent, etc.
- Incorporate structured omics data
- Upgrade query interface
- Use MIDR as proof of concept and lessons learned for enterprise-level resource at Duke
- Establish additional collaborations with computational biologists and biostatisticians to enable additional analysis and discovery
Priorities for 2014

• **Transfer of electronic health record** data from major healthcare systems
• Random sampling of households, initially a pilot to evaluate a **full representative sample**
• **National exposure** of the MURDOCK Study at highly visible conferences/meetings and new federally funded collaborations
• **Industry-sponsored funding** – big pharma, biotechnology and data companies
• Improved **strategies for follow-up/retention**
• Additional **disease cohorts** of study
• **Publications** in highly regarded scientific journals
• **New website** and expansion of social media sites and other health innovations
In Summary

- MURDOCK study aims to stratify patient populations using molecular biomarkers, redefining disease and human health
- Volunteer registry has enrolled 10k+ subjects, on the way to 50,000
- Novel opportunity to analyze longitudinal data across modalities and cohorts
- Translational team science in action
Thanks!

jessie.tenenbaum@duke.edu