Design and Analytical Issues in family-based longitudinal cohort studies: The FHS

Vasan S. Ramachandran, MD
BU and NHLBI’s Framingham Heart Study
Boston University School of Medicine

No Conflicts to Disclose
### Framingham Heart Study

**Longitudinal Community-Based Family Study**

<table>
<thead>
<tr>
<th>Year/Exam</th>
<th>Original cohort</th>
<th>Offspring cohort</th>
<th>Omni study 1</th>
<th>Third Gen cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>1948</td>
<td>N = 5209 men and women (ages 28-62)</td>
<td>5124</td>
<td>506</td>
<td>N~4000</td>
</tr>
<tr>
<td>1972</td>
<td>1644 spouse pairs, 596 extended families</td>
<td>2016</td>
<td>1995</td>
<td>New Offspring Spouses, N=100</td>
</tr>
</tbody>
</table>

**New Offspring Spouses, N=100**

1576 spouse pairs, 3514 biological offspring
Framingham Heart Study

• Longitudinal
  - Lifetime measures & Lifestyle measures

• Deeply Phenotyped
  - FHS is ~ The Human Phenome Project

• Extensive Genetic/Genomic Resources
  - Unique tissue resources

• Family-based study
FHS: Dense Phenotypic Characterization

- Dementia
- Stroke
- Cardiac
- Pulmonary
- Vascular
- Osteoporosis
- Osteoarthritis
- Alcohol

- Cancer
- Depression
- Eye
- Renal
- Endocrine
- Aging
- Diabetes

- Proteome
- Methylome
- CMS
- Metabolome

The Human Phenome Project

Funded grants
# Framingham Heart Study Cohort Retention

<table>
<thead>
<tr>
<th>COHORT</th>
<th>ENROLLED</th>
<th>% Exam or update or dead</th>
<th>Lost to date</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORIGINAL</td>
<td>5209</td>
<td>99.6</td>
<td>19</td>
</tr>
<tr>
<td>OFFSPRING</td>
<td>5124</td>
<td>99.8</td>
<td>15</td>
</tr>
<tr>
<td>Gen 3</td>
<td>4095</td>
<td>99</td>
<td>0</td>
</tr>
<tr>
<td>NOS</td>
<td>103</td>
<td>99.5</td>
<td>0</td>
</tr>
<tr>
<td>OMNI 1</td>
<td>507</td>
<td>99.2</td>
<td>30</td>
</tr>
<tr>
<td>OMNI 2</td>
<td>410</td>
<td>98.9</td>
<td>17</td>
</tr>
</tbody>
</table>
FHS is a Family-based Study

<table>
<thead>
<tr>
<th>Number of Pairs</th>
<th>Relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>6477</td>
<td>Parent-offspring</td>
</tr>
<tr>
<td>5530</td>
<td>Siblings</td>
</tr>
<tr>
<td>1267</td>
<td>Grandparent-grandchild</td>
</tr>
<tr>
<td>6849</td>
<td>Avuncular</td>
</tr>
<tr>
<td>306</td>
<td>Half siblings</td>
</tr>
<tr>
<td>8882</td>
<td>3rd degree</td>
</tr>
<tr>
<td>4503</td>
<td>4th degree</td>
</tr>
<tr>
<td>3876</td>
<td>5th degree</td>
</tr>
</tbody>
</table>
FHS design and statistical issues

- Correlated observations
  - Multilevel modeling
- Related individuals
  - Family-based analysis
- Longitudinal observations
  - Lifecourse approach, trajectories
  - Lifetime risk
  - Antecedent RF profile
  - Pooled repeated observations
- Risk functions
FHS design and statistical issues

- Temporal Trends
- Follow-up to extinction
  - Competing causes
- Genetic Studies
  - Linkage and Association analyses, including FBAT
Genomic Resources

- Marshfield genome scan (~400 markers/every 10cM)
- GWAS: 100K, 550K imputed to 1000K, 5M Omni
- Exome Chip Illumina v1.0
- Whole Exome Sequences
- Whole Genome Sequences
- MediSeq
- Candidate genes/SNPs

Access locally if working at BU
Or through dbGaP
SABRe CVD Initiative: Resources

- High-throughput technology to measure
  - Project 1: Discovery proteomics, metabolomics & lipomics
  - Project 2: Targeted immunoassays
  - Project 3: Gene expression profiling
  - Project 4: microRNA profiling

<table>
<thead>
<tr>
<th>Gene Expression</th>
<th>Methylome</th>
<th>Metabolome</th>
<th>miRNA</th>
<th>Protein Biomarkers*</th>
</tr>
</thead>
<tbody>
<tr>
<td>5,622</td>
<td>2,726</td>
<td>2,650</td>
<td>5,718</td>
<td>7,315</td>
</tr>
</tbody>
</table>

*170 immunoassay proteins.
FHS design and statistical issues

- Correlated observations
  - Multilevel modeling
- Related individuals
  - Family-based analysis
- Longitudinal observations
  - Lifecourse approach, trajectories
  - Lifetime risk
  - Antecedent RF profile
  - Pooled repeated observations
- Risk functions
• Level 1 = individual observations

• Level 2 = clustering units (participant)
  - 2 types of variability (between participant and within participant)

• Level 3 = Type of cohort
  (Original, Offspring, Third Generation)
FHS: Multilevel modeling of Echo traits

Longitudinal Tracking of Left Ventricular Mass Over the Adult Life Course
Clinical Correlates of Short- and Long-Term Change in the Framingham Offspring Study

Wolfgang Lieb, MD; Vanessa Xanthakis, MS; Lisa M. Sullivan, PhD; Jayashri Aragam, MD;
Michael J. Pencina, PhD; Martin G. Larson, ScD;
Emelia J. Benjamin, MD, ScM; Ramachandran S. Vasan, MD
*Circulation*. 2009;119:3085-3092

Longitudinal Tracking of Left Atrial Diameter Over the Adult Life Course: Clinical Correlates in the Community

David D. McMarus, MD; Vanessa Xanthakis, MS; Lisa M. Sullivan, PhD;
Justin Zacharich, MD; Jayashri Aragam, MD; Martin G. Larson, ScD;
Emelia J. Benjamin, MD, ScM; Ramachandran S. Vasan, MD
*Circulation*. 2010;121:667-674

Correlates of Echocardiographic Indices of Cardiac Remodeling Over the Adult Life Course
Longitudinal Observations From the Framingham Heart Study

Susan Cheng, MD; Vanessa Xanthakis, MS; Lisa M. Sullivan, PhD;
Wolfgang Lieb, MD; Joseph Massaro, PhD; Jayashri Aragam, MD;
Emelia J. Benjamin, MD, ScM; Ramachandran S. Vasan, MD
*Circulation*. 2010;122:570-578

Aortic Root Remodeling Over the Adult Life Course
Longitudinal Data From the Framingham Heart Study

Carolyn S.P. Lam, MBBS, MRCP; Vanessa Xanthakis, MS; Lisa M. Sullivan, PhD;
Wolfgang Lieb, MD; Jayashri Aragam, MD; Margaret M. Redfield, MD; Gary F. Mitchell, MD;
Emelia J. Benjamin, MD, ScM; Ramachandran S. Vasan, MD
*Circulation*. 2010;122:884-890
Longitudinal tracking of LV mass

Cumulative risk factor burden and longitudinal LV mass tracking

High CVD risk factor burden –
- Current smoker
- Hypertensive systolic BP of 140 mm Hg; not on antihypertensive treatment
- With diabetes
- BMI of 30.0 kg/m² (women) and 27.5 (men)

Low CVD risk factor burden –
- Non-smoker
- Non-hypertensive systolic BP of 130 mm Hg; not on antihypertensive treatment
- Free of diabetes
- BMI of 25 kg/m²

Lieb et al. Circulation, 2009; 119:3085
FHS design and statistical issues

• Correlated observations
  - Multilevel modeling

• Related individuals
  - Family-based analysis

• Longitudinal observations
  - Lifecourse approach, trajectories
  - Lifetime risk
  - Antecedent RF profile
  - Pooled repeated observations

• Risk functions
Residual Lifetime Risk for Developing Hypertension in Middle-aged Women and Men: The Framingham Heart Study

Ramachandran S. Vasan, MD; Alexa Beiser, PhD; Sudha Seshadri, MD; Martin G. Larson, ScD; William B. Kannel, MD; Ralph B. D'Agostino, PhD; Daniel Levy, MD

Estimating Lifetime Risk of Developing High Serum Total Cholesterol: Adjustment for Baseline Prevalence and Single-Occasion Measurements

Michael J. Pencina¹, Ralph B. D’Agostino¹, Alexa S. Beiser², Mark R.Cobain³, and Ramachandran S. Vasan⁴
Lifetime Risk of HTN

Statistical Methods

• Modified Kaplan-Meier method
  - cumulative incidence (CI)
  
  » age-specific incidence
  
  • \( f_A = h_A S_{A-1} \)
  
  \( F = \sum_{j=A \text{ min}} f_j \)
  
  - CI adjusted for competing risk
Unadjusted Cumulative Incidence of HTN defined as \{ \text{Stage I or RX} \}

Men and Women Free of HTN to age 65

Cumulative Incidence (%) vs. Years

- Exam 3
- Exam 14
### Lifetime Risk of HTN

**Data for Men & Women, Baseline Age 65 Yrs**

<table>
<thead>
<tr>
<th>F/U</th>
<th>Cumulative Incidence HTN (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 years</td>
<td>36</td>
</tr>
<tr>
<td>10 years</td>
<td>59</td>
</tr>
<tr>
<td>15 years</td>
<td>73</td>
</tr>
<tr>
<td>20 years</td>
<td>81</td>
</tr>
</tbody>
</table>
Lifetime Risk of CHF

Framingham Heart Study

Men

Women

Cumulative Risk

0% 10% 20% 20.6% 20.2%

Attained Age

0% 10% 20%

Attained Age

Lloyd-Jones Circulation 2002;106:3068
FHS design and statistical issues

- Correlated observations
  - Multilevel modeling
- Related individuals
  - Family-based analysis
- Longitudinal observations
  - Lifecourse approach, trajectories
  - Lifetime risk
  - Antecedent RF profile
  - Pooled repeated observations
- Risk functions
Antecedent Blood Pressure, Body Mass Index, and the Risk of Incident Heart Failure in Later Life


Study Sample
Framingham Heart Study
Original Cohort; Age >50 yrs; no CHF

Follow-up
Baseline Examination
1970

1950-60 remote
1960-70 recent
1948
1970
### Effect of Baseline or Antecedent BP on HF Events, Adjusted for Baseline Risk Factors except BP

<table>
<thead>
<tr>
<th></th>
<th>Hazards Ratio per SD Increment (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Current BP</strong></td>
<td></td>
</tr>
<tr>
<td>SBP</td>
<td>1.26 (1.13-1.40)*</td>
</tr>
<tr>
<td>DBP</td>
<td>0.93 (0.84-1.04)</td>
</tr>
<tr>
<td>PP</td>
<td>1.21 (1.11-1.31)*</td>
</tr>
<tr>
<td><strong>Previous BP</strong></td>
<td></td>
</tr>
<tr>
<td>SBP</td>
<td>1.36 (1.22-1.53)*</td>
</tr>
<tr>
<td>DBP</td>
<td>0.97 (0.87-1.09)</td>
</tr>
<tr>
<td>PP</td>
<td>1.31 (1.19-1.43)*</td>
</tr>
<tr>
<td><strong>Remote BP</strong></td>
<td></td>
</tr>
<tr>
<td>SBP</td>
<td>1.24 (1.11-1.39)*</td>
</tr>
<tr>
<td>DBP</td>
<td>1.02 (0.91-1.13)</td>
</tr>
<tr>
<td>PP</td>
<td>1.24 (1.13-1.36)*</td>
</tr>
</tbody>
</table>

Adjusted for standard RF including intercurrent MI, Age stratified

‡ P<0.05, † P<0.01, * P<0.001
Effect of Antecedent BP on HF Events

Adjusted for Baseline Risk Factors including BP

<table>
<thead>
<tr>
<th></th>
<th>Hazards ratio per SD (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Previous BP</strong></td>
<td></td>
</tr>
<tr>
<td>SBP</td>
<td>1.31 (1.11-1.55)†</td>
</tr>
<tr>
<td>DBP</td>
<td>1.02 (0.88-1.19)</td>
</tr>
<tr>
<td>PP</td>
<td>1.33 (1.14-1.54)*</td>
</tr>
<tr>
<td><strong>Remote BP</strong></td>
<td></td>
</tr>
<tr>
<td>SBP</td>
<td>1.17 (1.04-1.31)‡</td>
</tr>
<tr>
<td>DBP</td>
<td>1.05 (0.93-1.18)</td>
</tr>
<tr>
<td>PP</td>
<td>1.17 (1.06-1.31)†</td>
</tr>
</tbody>
</table>

Adjusted for standard RF including intercurrent MI, Age stratified

‡ P<0.05, † P<0.01, * P<0.001
Effect of Baseline or Antecedent BMI on HF Without & with adjustment for Baseline BMI

<table>
<thead>
<tr>
<th></th>
<th>(95% CI)</th>
<th>* P&lt;0.001</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No Adjustment for Current BMI</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current BMI</td>
<td>1.05 (1.03-1.07)*</td>
<td></td>
</tr>
<tr>
<td>Previous BMI</td>
<td>1.06 (1.04-1.08)*</td>
<td></td>
</tr>
<tr>
<td>Remote BMI</td>
<td>1.07 (1.05-1.10)*</td>
<td></td>
</tr>
</tbody>
</table>

| **Adjusted for Current BMI**   |                   |           |
| Previous BMI                  | 1.15 (1.08-1.23)* |           |
| Remote BMI                    | 1.09 (1.05-1.14)* |           |

Adjusted for standard RF including intercurrent MI, Age stratified

* P<0.001
Effect of Baseline or Antecedent BP & BMI on HF

Comparison of single random measure
FHS design and statistical issues

- Correlated observations
  - Multilevel modeling
- Related individuals
  - Family-based analysis
- Longitudinal observations
  - Lifecourse approach, trajectories
  - Lifetime risk
  - Antecedent RF profile
  - Pooled repeated observations
- Risk functions
General Cardiovascular Risk Profile for Use in Primary Care
The Framingham Heart Study

Ralph B. D’Agostino, Sr, PhD; Ramachandran S. Vasan, MD; Michael J. Pencina, PhD; Philip A. Wolf, MD; Mark Cobain, PhD; Joseph M. Massaro, PhD; William B. Kannel, MD

A Risk Score for Predicting Near-Term Incidence of Hypertension: The Framingham Heart Study

Nilsha I. Parikh, MD, MPH; Michael J. Pencina, PhD; Thomas J. Wang, MD; Emelia J. Benjamin, MD, ScM; Katherine J. Lanier, BS; Daniel Levy, MD; Ralph B. D’Agostino Sr., PhD; William B. Kannel, MD; and Ramachandran S. Vasan, MD
FHS design and statistical issues

• Temporal Trends

• Follow-up to extinction
  - Competing causes

• Genetic Studies
  - Linkage and Association analyses, including FBAT
LONG-TERM TRENDS IN THE INCIDENCE OF AND SURVIVAL WITH HEART FAILURE


**Table 1. Temporal Trends in the Age-Adjusted Incidence of Heart Failure.***

<table>
<thead>
<tr>
<th>PERIOD</th>
<th>MEN INCIDENCE OF HEART FAILURE</th>
<th>MEN RATE RATIO</th>
<th>WOMEN INCIDENCE OF HEART FAILURE</th>
<th>WOMEN RATE RATIO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rate/100,000 person-yr</td>
<td></td>
<td>rate/100,000 person-yr</td>
<td></td>
</tr>
<tr>
<td>1950–1969†</td>
<td>627 (475–779)</td>
<td>1.00</td>
<td>420 (336–504)</td>
<td>1.00</td>
</tr>
<tr>
<td>1970–1979</td>
<td>563 (437–689)</td>
<td>0.87 (0.67–1.14)</td>
<td>311 (249–373)</td>
<td>0.63 (0.47–0.84)</td>
</tr>
<tr>
<td>1980–1989</td>
<td>536 (448–623)</td>
<td>0.87 (0.67–1.13)</td>
<td>298 (247–350)</td>
<td>0.60 (0.45–0.79)</td>
</tr>
<tr>
<td>1990–1999</td>
<td>564 (463–665)</td>
<td>0.93 (0.71–1.23)</td>
<td>327 (266–388)</td>
<td>0.69 (0.51–0.93)</td>
</tr>
</tbody>
</table>
FHS design and statistical issues

- Temporal Trends

- Follow-up to extinction
  - Competing causes

- Genetic Studies
  - Linkage and Association analyses, including FBAT
Parental Heart Failure & Risk of CHF/LV Remodeling in Offspring

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Association of Parental Heart Failure with Risk of Heart Failure in Offspring


Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche

This manuscript highlights use of FHS menarche data in a consortium that identified genes implicated in BMI & various diseases, including rare disorders of puberty. Signals in imprinted regions identified for first time.
Using Family-Based Imputation in Genome-Wide Association Studies with Large Complex Pedigrees: The Framingham Heart Study

Ming-Huei Chen¹,²,³, Jie Huang³,⁴,⁸, Wei-Min Chen⁵, Martin G. Larson²,³,⁶, Caroline S. Fox³,⁷, Ramachandran S. Vasan³,⁸, Sudha Seshadri¹,³, Christopher J. O’Donnell³,⁹, Qiong Yang²,³,*

Using 8998 Framingham Heart Study (FHS) participants genotyped with Affymetrix 550K SNPs, we imputed genotypes of the same set of SNPs for an additional 3121 participants.

We developed a novel algorithm for splitting large pedigrees for imputation and found a plausible imputation quality filtering threshold based on FHS.

We could employ WGS in a selected 60-75% of FHS participants, using imputation, to have a whole genome in >>10K participants.
Thank you