Overview of Add Health for New Data Users

Kathleen Mullan Harris

Add Health
The National Longitudinal Study of Adolescent to Adult Health
National Longitudinal Study of Adolescent to Adult Health

- On-going program project that began in 1994.
- Developed in response to a congressional mandate to fund a study of adolescent health.
- Funded by the National Institute of Child Health and Human Development (NICHD) with co-funding from 23 other federal agencies and foundations.
- Fifth follow-up wave 2016-2018.
Key Features of Add Health

- Nationally representative study that explores the causes of health and health-related behaviors of adolescents and their outcomes in young adulthood.
- Multi-survey, multi-wave inter-disciplinary design.
- Direct measurement of the social contexts of adolescent life and their effects on health and health behavior.
- Unprecedented racial and ethnic diversity and genetically informed sibling samples.
Initial Goal:
Putting the Individual Into Context
Sampling Structure

School Sampling Frame = QED

- HS
  - Feeder

Sampling Frame of Adolescents and Parents N = 100,000+ (100 to 4,000 per pair of schools)

- Genetic Samples
  - Saturation Samples from 16 Schools
  - Disabled Sample
  - Main Sample 200/Community
    - Identical Twins
    - Fraternal Twins
    - Full Sibs
    - Half Sibs
  - Unrelated Pairs in Same HH

- Ethnic Samples
  - High Educ Black
  - Puerto Rican
  - Chinese
  - Cuban
Add Health Contextual model
The Social Structure of “Countryside” School District

Points Colored by Race

Unique Features of Study Design

• Ethnic oversamples produce
  – unprecedented diversity in race and ethnicity in a representative population of adolescents
  – large numbers of youth in immigrant families
• Embedded genetic sample enables researchers to sort out genetic from environmental effects and explore gene-environment interactions.
Race and Ethnic Diversity in Add Health

<table>
<thead>
<tr>
<th>Race/Ethnicity</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mexico</td>
<td>1,767</td>
<td>8.5</td>
</tr>
<tr>
<td>Cuba</td>
<td>508</td>
<td>2.5</td>
</tr>
<tr>
<td>Central-South America</td>
<td>647</td>
<td>3.1</td>
</tr>
<tr>
<td>Puerto Rico</td>
<td>570</td>
<td>2.8</td>
</tr>
<tr>
<td>China</td>
<td>341</td>
<td>1.7</td>
</tr>
<tr>
<td>Philippines</td>
<td>643</td>
<td>3.1</td>
</tr>
<tr>
<td>Other Asia</td>
<td>601</td>
<td>2.9</td>
</tr>
<tr>
<td>Black (Africa/Afro-Caribbean)</td>
<td>4,601</td>
<td>22.2</td>
</tr>
<tr>
<td>Non-Hispanic White (Eur/Canada)</td>
<td>10,760</td>
<td>52.0</td>
</tr>
<tr>
<td>Native American (non-Hispanic)</td>
<td>248</td>
<td>1.2</td>
</tr>
<tr>
<td>Total N</td>
<td>20,686</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Missing on race/ethnicity=59
Diversity in Add Health: Immigrant Status

<table>
<thead>
<tr>
<th>Immigrant Status</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st generation</td>
<td>1,707</td>
<td>8.34</td>
</tr>
<tr>
<td>2\textsuperscript{nd} generation</td>
<td>2,987</td>
<td>14.59</td>
</tr>
<tr>
<td>3\textsuperscript{rd} generation +</td>
<td>15,774</td>
<td>77.07</td>
</tr>
<tr>
<td>Total N</td>
<td>20,468</td>
<td>100.0</td>
</tr>
</tbody>
</table>
### Diversity of Family Forms at Wave I

<table>
<thead>
<tr>
<th>Family Structure</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 biological parents</td>
<td>10,339</td>
<td>53.3</td>
</tr>
<tr>
<td>2 adoptive parents</td>
<td>403</td>
<td>0.7</td>
</tr>
<tr>
<td>Bio Mom/Step Dad</td>
<td>2,756</td>
<td>13.6</td>
</tr>
<tr>
<td>Bio Dad/Step Mom</td>
<td>591</td>
<td>2.6</td>
</tr>
<tr>
<td>Single Mom</td>
<td>4,520</td>
<td>20.4</td>
</tr>
<tr>
<td>Single Dad</td>
<td>637</td>
<td>3.1</td>
</tr>
<tr>
<td>Surrogate parent(s)</td>
<td>1,499</td>
<td>6.3</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>20,745</td>
<td>100.0</td>
</tr>
</tbody>
</table>
Wave I
1994-1995 (79%)
Students 90,118
School Admin 144
Adolescents in grades 7-12 20,745
Parent 17,670

Wave II
1996 (88.6%)
School Admin 128
Adolescents in grades 8-12 14,738

Wave III
2001-2002 (77.4%)
Partners 1,507
Young Adults Aged 18-26 15,197

Wave IV
2008 (80.3%)
IIV Study ~100
Adults Aged 24-32 15,701

Wave V
2016-18
IIV Study ~100
Adults Aged 32-42 Target: 12,000
Parent 3,000
# Questionnaire Content Across Waves

<table>
<thead>
<tr>
<th>Waves I, II</th>
<th>Wave III</th>
<th>Wave IV</th>
<th>Wave V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic</td>
<td>Demographic</td>
<td>Demographic</td>
<td>Demographic</td>
</tr>
<tr>
<td>Family, siblings, friends</td>
<td>Family, siblings, friends</td>
<td>Family, siblings, friends</td>
<td>Family, siblings, friends</td>
</tr>
<tr>
<td>Education, work</td>
<td>Education, work, <strong>military</strong></td>
<td>Educ, work, <strong>military (records)</strong></td>
<td>Educ, work, military</td>
</tr>
<tr>
<td>Physical and mental health</td>
<td>Physical and mental health</td>
<td>Physical and mental health</td>
<td>Physical and mental health</td>
</tr>
<tr>
<td>Daily activities and sleep</td>
<td>Daily activities and sleep</td>
<td>Daily activities and sleep</td>
<td>Daily activities and sleep</td>
</tr>
<tr>
<td>Relationships</td>
<td>Relationships</td>
<td>Relationships</td>
<td>Relationships</td>
</tr>
<tr>
<td>Sexual, &amp; fertility histories</td>
<td>Sexual, &amp; fertility histories</td>
<td>Sexual, &amp; fertility histories</td>
<td>Sexual, &amp; fertility histories</td>
</tr>
<tr>
<td>Substance use</td>
<td>Substance use</td>
<td>Substance use and abuse</td>
<td>Substance use and abuse</td>
</tr>
<tr>
<td><strong>Involvmt w/criminal justice sys</strong></td>
<td><strong>Involvmt w/criminal justice sys</strong></td>
<td><strong>Involvmt w/criminal justice sys</strong></td>
<td><strong>Involvmt w/criminal justice sys</strong></td>
</tr>
<tr>
<td>Attitudes, religion</td>
<td>Attitudes, religion</td>
<td>Work attitudes and chars, relig</td>
<td>Work attitudes and chars, relig</td>
</tr>
<tr>
<td>Economics, expectations</td>
<td>Economics, expectations</td>
<td>Economics, expectations</td>
<td>Economics, expectations</td>
</tr>
<tr>
<td>Psychological, personality</td>
<td>Psychological, personality</td>
<td><strong>Big 5 Personality, stressors</strong></td>
<td>Personality, Stressors</td>
</tr>
<tr>
<td><strong>Children and parenting</strong></td>
<td><strong>Children and parenting</strong></td>
<td><strong>Children and parenting</strong></td>
<td>Children and parenting</td>
</tr>
<tr>
<td>Civic participation</td>
<td>Civic participation</td>
<td>Civic participation</td>
<td>Civic participation</td>
</tr>
<tr>
<td>Gambling</td>
<td>Gambling</td>
<td><strong>Cognitive function</strong></td>
<td>Cognitive function</td>
</tr>
<tr>
<td>Mentoring</td>
<td>Mentoring</td>
<td>Psychosocial factors</td>
<td>Psychosocial factors</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Retrospective childhood health &amp; SES</strong></td>
</tr>
</tbody>
</table>
Physical and Social Contextual Data

- Census
- CDC (STD prevalence)
- FBI crime statistics
- National Center for Health Statistics
- National Council of Churches

- Questionnaire-based social context data
  - In-School Network Data
  - Adolescent Romantic Pair Data
  - In-School Friendship Nominations
  - In-Home Friendship Nominations
  - Family Context (Parent Interview, sibling information)
  - Wave III Married, cohabiting, and dating couples (“couple context”)
Physical and Social Contextual Data

• Ancillary studies
  – ONEdata - Obesity & Neighborhood Environment Database
    • Waves I, III, and IV
    • E.g., parks, street connectivity, sidewalks, presence of fast food restaurants, alcohol outlets
  – The Adolescent Health and Academic Achievement Study
    • Analysis of respondents’ high school transcripts
    • detailed measures of academic progress and high school curriculum
### Biological Data Across Waves

<table>
<thead>
<tr>
<th>Adolescence</th>
<th>Transition to Adulthood</th>
<th>Young Adulthood</th>
<th>Adulthood</th>
</tr>
</thead>
</table>

- **Embedded genetic sample of ~3,000 pairs**

#### Physical development

<table>
<thead>
<tr>
<th>Wave</th>
<th>Physical development</th>
</tr>
</thead>
<tbody>
<tr>
<td>I-II</td>
<td>Height, weight</td>
</tr>
<tr>
<td>III</td>
<td>Height, weight</td>
</tr>
<tr>
<td>IV</td>
<td>Height, weight, waist</td>
</tr>
<tr>
<td>V</td>
<td>Height, weight, waist</td>
</tr>
</tbody>
</table>

- **STI tests (urine)**
- **HIV test (saliva)**
- **Genetic (buccal cell DNA)**
- **Genetic (whole blood)**
- **Medications**
- **Renal**

- **Metabolic**
- **Immune function**
- **Inflammation**
- **Cardiovascular**
- **Genetic (buccal cell DNA)**
- **Genetic (whole blood)**
- **Medications**
- **Renal**
# Social and Biological Longitudinal Data in Add Health

## Adolescence vs. Adulthood

<table>
<thead>
<tr>
<th>Wave I-II</th>
<th>Wave III</th>
<th>Wave IV</th>
<th>Wave V</th>
</tr>
</thead>
<tbody>
<tr>
<td>(12-20)</td>
<td>(18-26)</td>
<td>(24-32)</td>
<td>(32-42)</td>
</tr>
</tbody>
</table>

### Social environmental data:
- School
- Family
- Romantic relationship
- Neighborhood
- Community
- Peer

### Biological data:
- Biological resemblance to siblings in household on 3,000 pairs
- Height
- Weight
- BMI
- STI test results
- HIV test results
- DNA
- DNA
- GWAS
- mRNA
- DNA
- Kidney disease
- mRNA

### Biological data:
- Height
- Weight
- BMI
- Blood pressure
- Pulse
- Immune
- Inflammation
- Diabetes
- DNA
- GWAS
- mRNA
Wave IV Biospecimen Participation

- 96% of respondents consented to provide saliva for DNA
- 94% consented to provide blood spots
- For each specimen, 81% of those consenting to collection also consented to archiving
- Approx 12,000 DNA samples available for further testing
- GWAS data will be available via dbGaP (N~10,000)
Wave V Overall Goals

• Re-interview Add Health cohort members to collect social, environmental, behavioral, and biological data with which to track the emergence of chronic disease.

• Build on the life course history of respondents by adding and refining early-life measures of their birth and childhood:
  – Retrospective questions about birth and early childhood in Wave V survey
  – Birth records of respondents born in a subset of states

• Bring these data together with existing longitudinal data to create a 40-year life course record to test hypotheses about developmental origins of health and disease.
Wave V Interview 2016-2018

- Wave I respondents who will be moving through their 4th decade of life (32-42 years)
- Collect the following data:
  - survey data reflecting longitudinal and new information
  - longitudinal and new biological data and specimens
  - geographic locations for longitudinal spatial data
  - State of birth to obtain birth records of Add Health respondents born in a subset of six states
- Main study fieldwork began June 6, 2016
Wave V Data Collection

• Mixed mode survey design:
  – Web/mail with in-person and phone non-response follow up

• Collect biological measures and specimens using separate biomarker subcontractor for in-home examination
  – Venous blood draw

• Continuous interviewing 2016-2018
Wave V Data Collection

• Conduct fieldwork on 3 nationally-representative samples sequentially during 2016-2018:
  – Sample 1: June 2016
  – Sample 2a: January 2017
  – Sample 2b: June 2017
    • subsample of 1,100 respondents interviewed in-person by interviewer to replicate Wave IV interview context to estimate mode effects
  – Sample 3: Oct 2017
Wave V Sampling Structure

All samples are nationally representative

Sample 1

Sample 2

Sample 2A (web & mail survey)

Sample 2B (in-home interview)

Sample 3

2016

2017

2018
Mixed Mode Wave V Survey

- Web survey: (50 mins total)
- Paper survey with same visual presentation of modules
- Obtain consent for biomarker collection in follow-up in-home exam
- Nonresponse follow-up
  - Sample nonrespondents and administer web survey in-person (CASI) on laptop or via telephone
  - Abbreviated telephone questionnaire for contact information and request consent for biomarkers.
## Sample 1 Experimental Treatments

<table>
<thead>
<tr>
<th>Factor A: Survey Protocol</th>
<th>Factor B: Propensity/Incentive</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model-directed Incentive Plan</td>
<td>Standard Incentive Plan</td>
</tr>
<tr>
<td>Modular Survey</td>
<td>Low Response Propensity/High Incentive</td>
<td>High Response Propensity/Low Incentive</td>
</tr>
<tr>
<td></td>
<td>$35/$30</td>
<td>$25/$20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$30/$25</td>
</tr>
<tr>
<td>Singular Survey</td>
<td>$65</td>
<td>$45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$55</td>
</tr>
</tbody>
</table>
Current Progress of Wave V

- Anticipated sample size ~12,000 (effective response rate of 80%)
- Completed 11,022 surveys to date
- 66% biomarker consent rate

<table>
<thead>
<tr>
<th>Survey Mode</th>
<th>N</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web</td>
<td>8,540</td>
<td>77.5</td>
</tr>
<tr>
<td>Paper</td>
<td>361</td>
<td>3.3</td>
</tr>
<tr>
<td>In-person</td>
<td>1,882</td>
<td>17.1</td>
</tr>
<tr>
<td>Phone</td>
<td>239</td>
<td>2.2</td>
</tr>
<tr>
<td>Total</td>
<td>11,022</td>
<td>100</td>
</tr>
</tbody>
</table>
Wave V: Biological data

- Repeat measures of biomarkers:
  - anthropometrics
  - blood pressure
  - Whole blood assays of:
    - Inflammation
    - Lipids
    - Glucose
    - Glycosylated hemoglobin
- New biomarkers of kidney disease:
  - Creatinine
  - Cystatin C
- Medications inventory
Wave V Biomarkers

- 66% consent rate for biomarker visit
- Consent rates among those who complete biomarker visit:
  - Venous blood    94%
  - Blood pressure  95%
  - Anthropometrics 99%
Vital Events Data

- Birth records data from respondents in 6 states
  - To date have received vital records approval for 3 targeted states
  - Release birthweight, gestational age, other characteristics of birth circumstances

- Death surveillance
  - To date, 364 deaths identified since Wave I; of which we have collected 344 death certificates
  - Assembling gold standard information on cause of death and circumstances of death
Wave V Data Release

Two data releases:

- Sample 1 data released in Nov 2017 (N~3800)
  - Biomarker data for Sample 1 in Sept 2018

- Full Wave V sample data in 2019
  - All data and samples combined with sampling weights

- Birth records data (with final Wave V release or subsequent)

- Death records data (with final Wave V release or subsequent)
Add Health Data Dissemination

Wave V

Add Health Parent Study Phase 2
Add Health Data Dissemination

• Data dissemination structure & security

• Restricted-use data contracts

• Data discovery tools and resources

• CPC Data Portal
Four tier data dissemination according to disclosure risks

- Public-use data
- Restricted-use data
- High-security restricted data – Romantic Pairs
- Secure data facility for analyzing high school transcript data and for using geocodes to link contextual data
- dbGaP access for GWAS data (currently deposited)
Public-Use Data can be downloaded from the following places:
Data Security

• User requirements to protect from deductive disclosure:
  – Pledge of confidentiality
  – Monitoring of data use
  – Store data securely
  – Deletion of temporary data analysis files every six months.
  – Security of printed information
  – Password protected screen saver, set to activate after **3 minutes** of idle time.
  – Access data only from approved locations.
Restricted-Use Data Contract

• Complete contract application – pdf
• Include IRB approval or waiver letter
• Include data processing payment (by check only)
• Make sure to download Attachment A – security plan form and include in application
Don’t forget….

• Your contract expires every 3 years!
• Email us if you think your contract might be expired or to find out what your expiration date is.

• Annual reports are due every year!
• Let us know if you make changes to research staff or data storage locations.
Please visit our Contracts Homepage!

Add Health Restricted-Use Data Contracts

New Contract
To apply for restricted-use data, please download and complete the Restricted-Use Data Contract using the CPC Data Portal. See our Security Plans page for more information on developing your Data Security Plan. See our Annual Reports and Publications page for annual report requirements.

Renew Contract
To renew your contract, please email Add Health Contracts for instructions. Add Health Restricted-Use Contracts must be renewed every three years. Your renewal date is included with your approved contract. Please note that no processing fee is required for renewal. See our Annual Reports and Publications page for annual report requirements.

Add an Additional Person to Contract
For information on how to add additional researchers, collaborators, orifices, or information technology staff to your Add Health Restricted-Use Contract, navigate to the Add an Additional Person to a Restricted-Use Data Contract page.

Transfer Contract
If you are transferring institutions or transferring your Add Health Restricted-Use Contract to another researcher at your university, you must first obtain Add Health's approval. Please see the Transfer Restricted-Use Data Contract page.

Terminate Contract
If you will not be renewing your contract, please complete the steps found on the Terminate Restricted-Use Data Contract page.

Request Additional Data
If you have a current Restricted-Use Contract and would like to request additional data sets, please complete the Additional Data Order form. This form can be scanned and emailed to Add Health Contracts.

Romantic Pairs Contract
For information on how to request the Romantic Pairs data, navigate to the Romantic Pairs Data Contract page.

Request Remote Access
Current researchers that are accessing Add Health data through a server and would like to gain access from another location (home, office, etc.) should download the Security Information for Remote Access form and email it to Add Health Contracts.

Contract Questions
For questions regarding contracts, please navigate to our FAQ contracts page or email Add Health Contracts.
• Interactive tool, developed in 2015 by the Add Health Team

• Developed in response to feedback that pdf codebooks were difficult to navigate and information on data collected across waves was nearly impossible to find

• Browse by topic or search by variable name, keyword, or phrase to discover the rich volume of data collected by Add Health
CPC Data Portal

- Restricted-Use data contract applications are now submitted through the CPC Data Portal [https://data.cpc.unc.edu/](https://data.cpc.unc.edu/)

- Expedite and standardize contract application and renewal process for restricted-use Add Health data
<table>
<thead>
<tr>
<th>Project</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Add Health</strong></td>
<td>The National Longitudinal Study of Adolescent to Adult Health (Add Health) is a longitudinal study of a nationally representative sample of adolescents in grades 7-12 in the United States during the 1994-95 school year.</td>
</tr>
<tr>
<td>China Health and Nutrition Survey</td>
<td>The China Health and Nutrition Survey is an ongoing international collaborative project that was designed to examine the effects of the health, nutrition, and family planning policies.</td>
</tr>
<tr>
<td>MEASURE Evaluation</td>
<td>The MEASURE Evaluation project focuses on strengthening capacity in developing countries to gather, interpret, and use data to improve health. The project creates tools, approaches, and data for rigorous evaluations, providing evidence to address health challenges.</td>
</tr>
<tr>
<td>Measurement, Learning &amp; Evaluation (MLE)</td>
<td>The Measurement, Learning &amp; Evaluation (MLE) Project is the evaluation component of the Urban Reproductive Health Initiative (URHI). URHI is a Bill &amp; Melinda Gates Foundation funded multi-country program in India, Kenya, Nigeria, and Senegal that aimed to improve the reproductive health of the urban poor by increasing the accessibility, quality, and use of family planning services. From 2010 – 2015, the projects have used different interventions, both demand and supply side, to increase use of and access to contraceptives.</td>
</tr>
<tr>
<td>Russia Longitudinal Monitoring Survey of HSE</td>
<td>The Russia Longitudinal Monitoring Survey is a series of nationally representative surveys designed to monitor the effects of Russian reforms on the health and economic welfare of households and individuals in the Russian Federation.</td>
</tr>
<tr>
<td>The Transfer Project</td>
<td>The Transfer Project is an innovative research initiative led by the University of North Carolina at Chapel Hill, UNICEF, and the Food and Agriculture Organization (FAO) to understand the broad impacts of government-led cash transfer programs in sub-Saharan Africa. National governments and local research organizations are key partners in all stages to assure research is utilized to inform the design and expansion of national social cash transfers (SCTs).</td>
</tr>
</tbody>
</table>
How to apply

- Go to Add Health page
- Select Restricted-Use Data (Group View)
- Click “Add to Cart” to the desire datasets.
- Click the shopping “Cart” in the header menu
- Under Shopping Cart click the button
  
  Proceed to checkout

- Log in using a Microsoft Account

  Sign in with Microsoft (Personal, work or school account)

- Email confirmation will be sent to the user and CPC Data Portal Administrator

https://data.cpc.unc.edu/projects/2/view
CPC Data Portal

Parent Study Files

Parent with Spouse/Partner Data

The Parent data file contains social, demographic, behavioral, and health data collected in 2015-2017 on a probability sample of the parents of Add Health sample members who were originally interviewed in 1995. Data for 2,013 Wave I parents, representing 2,245 Add Health sample members, are available. Additionally, 988 current spouse/partner interviews are available.
Add Health Co-Funders

- National Institute of Child Health and Human Development
- National Cancer Institute*
- National Center for Health Statistics, Centers for Disease Control and Prevention, DHHS
- National Center for Injury Prevention and Control, Centers for Disease Control and Prevention, DHHS*
- National Center for Minority Health and Health Disparities*
- National Institute of Allergy and Infectious Diseases*
- National Institute of Deafness and Other Communication Disorders*
- National Institute of General Medical Sciences
- National Institute of Mental Health
- National Institute of Nursing Research*
- National Institute on Aging*
- National Institute on Alcohol Abuse and Alcoholism*
- National Institute on Drug Abuse*
- National Science Foundation*
- Office of AIDS Research, NIH*
- Office of the Assistant Secretary for Planning and Evaluation, DHHS*
- Office of Behavioral and Social Sciences Research, NIH*
- Office of the Director, NIH
- Office of Minority Health, Centers for Disease Control and Prevention, DHHS
- Office of Minority Health, Office of Public Health and Science, DHHS
- Office of Population Affairs, DHHS*
- Office of Research on Women's Health, NIH*

*Wave IV co-funders
Wave V co-funders
• Questions or Feedback?

• Contract Inquires:
• addhealth_contracts@unc.edu