The National Longitudinal Study of Adolescent to Adult Health (Add Health): Brief History, Updates, & Exciting Wave VI Plans*

ROBERT A. HUMMER, DIRECTOR
HOWARD W. ODUM DISTINGUISHED PROFESSOR OF SOCIOLOGY
UNIVERSITY OF NORTH CAROLINA – CHAPEL HILL

* PRESENTATION PREPARED FOR THE ADD HEALTH USERS CONFERENCE, JULY 11-12, 2022 *
Add Health Origins & Evolution

Developed in response to a 1993 Congressional mandate to fund a longitudinal study of adolescent health; project began in 1994 with primary funding from NICHD

Originally two waves (Waves I and II) conducted in 1994-95 and 1995-96; evolved into multi-wave design in early 2000s

Originally called the National Longitudinal Study of Adolescent Health; renamed the National Longitudinal Study of Adolescent to Adult Health with Wave V in 2014

Recipient of 2016 Golden Goose Award

Most recently, Wave VI was funded primarily by NIA (2021-25)
Add Health Sample Design

School Sampling Frame = QED

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 ~ 150/Community

Ethnic Samples
- High Educ Black
- Puerto Rican
- Chinese
- Cuban

Identical Twins
Fraternal Twins
Full Sibs
Half Sibs
Unrelated Pairs in Same HH
Original Add Health Multi-Level Conceptualization
Innovative Features of the Initial Add Health Design

Genetic Sub-Sample
N = 3,699

- Twins: 784
- Full sibs: 662
- Half sibs: 442
- Non-related: 1,251
- Adopted: 560

<table>
<thead>
<tr>
<th>Family Structure, Wave I</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 Biological parents</td>
<td>10,339</td>
</tr>
<tr>
<td>2 Adoptive parents</td>
<td>403</td>
</tr>
<tr>
<td>Bio Mom/Step Dad</td>
<td>2,756</td>
</tr>
<tr>
<td>Bio Dad/Step Mom</td>
<td>591</td>
</tr>
<tr>
<td>Single Mom</td>
<td>4,520</td>
</tr>
<tr>
<td>Single Dad</td>
<td>637</td>
</tr>
<tr>
<td>Surrogate parent(s)</td>
<td>1,499</td>
</tr>
<tr>
<td>Total</td>
<td>20,745</td>
</tr>
</tbody>
</table>
### Add Health Innovative Design: Wave I

#### Race/Ethnicity*

<table>
<thead>
<tr>
<th>Race/Ethnicity</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latina/o or Hispanic</td>
<td>3,492</td>
<td>16.9</td>
</tr>
<tr>
<td>Black or African American</td>
<td>4,601</td>
<td>22.2</td>
</tr>
<tr>
<td>Asian American</td>
<td>1,585</td>
<td>7.6</td>
</tr>
<tr>
<td>American Indian / Alaska Native</td>
<td>248</td>
<td>1.2</td>
</tr>
<tr>
<td>NH White</td>
<td>10,760</td>
<td>52.0</td>
</tr>
<tr>
<td>Other/Unknown</td>
<td>59</td>
<td>0.2</td>
</tr>
<tr>
<td>Total</td>
<td>20,745</td>
<td>100.0</td>
</tr>
</tbody>
</table>

* One of many ways to measure race and/or ethnicity in Add Health.
Add Health Grows Up…
Add Health longitudinal, multi-level, intergenerational design
(Each wave also includes contextual data specific to various levels of geography)

<table>
<thead>
<tr>
<th>Wave</th>
<th>Administrative</th>
<th>Survey Administration</th>
<th>Biological Data Collection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wave I</td>
<td>Students N=90,118</td>
<td>Adolescents in grades 7-12 N = 20,745</td>
<td>Height, weight</td>
</tr>
<tr>
<td>1994-95</td>
<td>School Admin N = 144</td>
<td>Parent N=17,670</td>
<td></td>
</tr>
<tr>
<td>Wave II</td>
<td>School Admin N = 128</td>
<td>Adolescents in grades 8-12 N = 14,738</td>
<td>Height, weight</td>
</tr>
<tr>
<td>1996</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wave II</td>
<td>High school</td>
<td>Young adults Aged 18-26 N = 15,197</td>
<td>Height, weight, STI, HIV, genetic (buccal cell DNA)</td>
</tr>
<tr>
<td>2001-02</td>
<td>transcripts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wave IV</td>
<td></td>
<td>Adults Aged 24-32</td>
<td>Height, weight, waist, metabolic, immune, inflammation, cardiovascular, medications, candidate genes, GWAS</td>
</tr>
<tr>
<td>2008-09</td>
<td></td>
<td>N = 15,701</td>
<td></td>
</tr>
<tr>
<td>Wave V</td>
<td>Birth &amp; death</td>
<td>Adults Aged 33-43</td>
<td>Height, weight, waist, metabolic, immune, inflammation, cardiovascular, renal, medications, gene expression, epigenetic, microbiome</td>
</tr>
<tr>
<td>2016-18</td>
<td>records</td>
<td>N = 12,300</td>
<td></td>
</tr>
<tr>
<td>Wave VI</td>
<td>Birth &amp; death</td>
<td>Early midlife adults</td>
<td>Height, weight, waist, metabolic, immune, inflammation, cardiovascular, renal, medications, gene expression, epigenetic, microbiome, cognitive, physical functioning</td>
</tr>
<tr>
<td>2022-24</td>
<td>records</td>
<td>Aged 40-49</td>
<td></td>
</tr>
</tbody>
</table>
Add Health is exceptionally unique compared to many population-based studies of health:

- Prospectively collected data on social, economic, and health experiences since adolescence
- Biomarker, anthropometric, other biological data, and cognitive data before many health outcomes manifest
- Rich data from multiple levels … cellular, individual survey, & contextual data from many levels
- Includes parental data from adolescence and, for some parents, in older adulthood
- Facilitates analyses of health disparities: racial/ethnic diversity, SES distribution, focus on gender identity & sexual orientation, and geographic distribution
- Nationally representative
Sneak Peek into Wave VI
Add Health Wave VI: Overall Goal

Collect & disseminate high-quality data on life course determinants and trajectories of health, health behavior, and health disparities among a large, nationally representative cohort aging into midlife.

Add Health as an aging study; maintain longitudinal integrity while innovating:

* Disparities focus (in survey design & content; contextual data)
* Ages 40-49: Work; Caregiving; Stress; Discrimination
* Behavioral health focus
* Cardiometabolic health focus
* Cognitive & physical/sensory functioning focus
* Mortality surveillance
Wave VI Summary of Data Domains

Domains:
- Survey data (3/4 web; 1/4 in-person)
- Cognitive, sensory, and physical functioning
- Contextual data from multiple levels
- Biological data (biomarker, anthropometric, meds)
- Mortality surveillance; birth records

Data ETA: 2025
Wave VI Survey Design: 2 Sub-Samples

Total eligible
N=19,300

SAMPLE 1: WEB SURVEY & COGNITION;
IN-PERSON NON-RESPONSE FOLLOW-UP (NRFU)
ELIGIBLE = 15,300

Total Web/NRFU
Expected Completes = 10,000

SAMPLE 2: IN-PERSON SURVEY &
COGNITIVE/SENSORY FUNCTIONING
ELIGIBLE = 4,000

Total In-Person Survey
Expected Completes = 3,000

Total Wave VI
Expected Survey Completes = 13,000
Wave VI survey enhancements: many related to race/racism, midlife stresses, & health

- Multiple measures of race & self-reported skin tone
- Multiple measures of discrimination
- Chronic stressors
- Interactions/experiences with police & criminal justice system
- Trust in institutions
- Enhanced measures of mental & physical health
- Survey measures of cognitive & sensory functioning
Wave VI In-Home Health Exam Goals

Of ~13,000 survey respondents, our goals are to:

1) Collect ~10,000 consents for the home health examination

2) Conduct 7,500 home health examinations

3) Conduct 100 repeat home health examinations (assess intra-individual variation)

4) Receive 4,000 leave-behind microbiome collection kits via mail
### Add Health Bio/Clinical Content, Waves I–VI

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Anthropometric</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Omics</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Immune/Infectious</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Metabolic</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Medications</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Renal/ Liver</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Microbiome</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Neurocognitive (AD/ADRD Risk)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>
Current Add Health Funders

National Institute on Aging (NIA)
Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD)
National Institute on Minority Health and Health Disparities (NIMHD)
National Institute on Drug Abuse (NIDA)
NIH Office of Behavioral and Social Sciences Research (OBSSR)
NIH Office of Disease Prevention (ODP)
Acknowledgements

Wave VI of Add Health is supported by two grants from the National Institute on Aging (1U01AG071448, principal investigator Robert A. Hummer, and 1U01AG071450, principal investigators Allison E. Aiello and Robert A. Hummer) to the University of North Carolina at Chapel Hill. Co-funding for Wave VI is being provided by the Eunice Kennedy Shriver National Institute of Child Health and Human Development, the National Institute on Minority Health and Health Disparities, the National Institute on Drug Abuse, the NIH Office of Behavioral and Social Science Research, and the NIH Office of Disease Prevention. Waves I-V data are from the Add Health Program Project, grant P01 HD31921 (Kathleen Mullan Harris) from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. The content of this presentation is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health or the University of North Carolina at Chapel Hill.

Add Health was originally designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill. Add Health is currently directed by Robert A. Hummer; it was previously directed by Kathleen Mullan Harris (2004-2021) and J. Richard Udry (1994-2004).

Information on obtaining Add Health data is available on the project website (https://addhealth.cpc.unc.edu).
Extra Slides
## Survey Data 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, sibs, friends</td>
<td>Family, sibs, friends</td>
<td>Family, sibs, friends</td>
<td>Family, sibs, friends</td>
</tr>
<tr>
<td>Education, work</td>
<td>Education, work</td>
<td>Education, work</td>
<td>Education, work</td>
</tr>
<tr>
<td>Physical &amp; mental health</td>
<td>Physical &amp; mental health</td>
<td>Physical &amp; mental health</td>
<td>Physical &amp; mental health</td>
</tr>
<tr>
<td>Health Behaviors &amp; Care</td>
<td>Health Behaviors &amp; Care</td>
<td>Health Behaviors &amp; Care</td>
<td>Health Behaviors &amp; Care</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>Involvement w/criminal justice</td>
<td>Involvement w/criminal justice</td>
<td>Involvement w/criminal justice sys</td>
<td>Involvement w/criminal justice sys</td>
</tr>
<tr>
<td>Work attitudes and chars</td>
<td>Work attitudes and chars</td>
<td>Work attitudes and chars</td>
<td>Work attitudes and chars</td>
</tr>
<tr>
<td>Attitudes, religion</td>
<td>Attitudes, religion</td>
<td>Attitudes, religion</td>
<td>Attitudes, religion</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>Psychological, personality</td>
<td>Psychological, personality</td>
</tr>
<tr>
<td>Cognitive</td>
<td>Cognitive</td>
<td>Cognitive</td>
<td>Cognitive</td>
</tr>
<tr>
<td>Children and parenting</td>
<td>Children and parenting</td>
<td>Children and parenting</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>Gambling</td>
<td>Stressors</td>
</tr>
<tr>
<td>Mentoring</td>
<td>Mentoring</td>
<td>Mentoring</td>
<td>Stressors</td>
</tr>
<tr>
<td>Retrospective child health &amp; SES</td>
<td>Retrospective child health &amp; SES</td>
<td>Retrospective child health &amp; SES</td>
<td>Stressors</td>
</tr>
</tbody>
</table>
Examples of Contextual & Administrative Data

Social network data from respondents’ schools
School characteristics which respondents attended
High school transcripts; College contexts
Census and ACS-based data … tracts, counties, states
Neighborhood: parks, street connectivity, sidewalks, fast food, alcohol outlets, etc.
FBI crime statistics
NCHS mortality rates, life expectancy
Chetty’s social mobility contexts
Policy Contexts (e.g., ACA; sexual minority policy; school desegregation)
Pollutants and Climate
National Council of Churches
Birth certificate data (in progress)
Mortality surveillance data (in progress)
## Biological/Clinical Data Across Waves

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Adolescence</strong></td>
<td><strong>Transition to Adulthood</strong></td>
<td><strong>Young Adulthood</strong></td>
<td><strong>Adulthood</strong></td>
</tr>
<tr>
<td>Embedded genetic sample of 3,667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Physical development</strong></td>
<td><strong>Height, weight</strong></td>
<td><strong>Height, weight, waist</strong></td>
<td><strong>Height, weight, waist/arm</strong></td>
</tr>
<tr>
<td>Height, weight</td>
<td>Genomic Buccal cell candidate genes</td>
<td>Genomic Buccal cell GWAS</td>
<td>Genomic Whole blood DNAm, gene expression</td>
</tr>
<tr>
<td>Infection</td>
<td>CRP, IL-6, IL-8, IL-10, EBV, CMV, H.pylori, HSV</td>
<td>Immune/Infection CRP</td>
<td></td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>Blood pressure, pulse</td>
<td>Cardiovascular Blood pressure, pulse</td>
<td></td>
</tr>
<tr>
<td>Metabolic</td>
<td>Cholesterol, HbA1c, glucose</td>
<td>Metabolic Cholesterol, triglycerides, HbA1c, glucose</td>
<td></td>
</tr>
<tr>
<td>Medications</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Renal/Liver</td>
<td>Creatinine, cystatin C,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Microbiome</td>
<td>Fecal</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

[UNC CAROLINA POPULATION CENTER](https://www.popcenter.org/)

[Add Health](https://www.addhealth.org/) - The National Longitudinal Study of Adolescents to Adult Health
Add Health Selected Impacts to Date

GRANTS, USERS, PUBLICATIONS AND CITATIONS
Sampling of Key Contributions

Cardiovascular Health
- Preeminent dataset for tracking the obesity epidemic and cardio-metabolic risk among US young adults in recent years
- Identifying critical early life exposures and events associated with increased cardiovascular risk in adulthood, such as infant/child health, socioeconomic status, neighborhood effects, and family relationships and structure

Disparities and Aging/Life Course
- Documenting the formation and widening of gender, sexual minority, racial/ethnic, and socioeconomic disparities in key health indicators from adolescence into early adulthood
- Emphasizing the importance of socioeconomic status, social relationships and stress exposures as determinants of health and “weathering” throughout the life course

Genetics and Health
- Disentangling the contributions of genetic and socioenvironmental factors shaping adult health, especially for smoking, obesity, and metabolic risk

Emerging Trends at Midlife
- Providing a broader understanding of rising despair among a recent young adult cohort, spanning multiple indicators of poor mental health and – critically – not limited to rural areas and/or White adults
Grants

Add Health data have been utilized in ~225 NIH grants to date:

- NICHD: 111
- NIDA: 34
- NIAAA: 27
- NIMH: 25
- NINR: 6
- NIDDK: 5
- NIA: 4
- NCI: 4
- NIMHD: 4
- NHLBI: 3
- AHRQ: 2
- NIAID: 2

... and more ... NIH, other federal (e.g., NSF), and foundations.
Users

Public-Use Data
  - Access requirement: Data use agreement
  - ~50,000+ researchers

Restricted-Use Data
  - Access requirement: Approved contract (renewed every 3 yrs.)
  - ~20+ contracts received each month; recent release of Wave V data substantially increased demand
  - Users in 43 U.S. states (94%) & 26 additional countries (6%)
  - Users from:
    - R1 (74%), R2, R3, International, & Non-Research universities
    - Non-Profit/Private Firms, Medical Schools/Facilities, & Government Agencies
Publications and Citations

~4000 peer-reviewed publications with 300,000+ citations

Averaging **150+ publications a year** over past 2 decades
  - Almost one published article *per day* over last five years

Broad reach across **multiple disciplines**
  - **Medical**: JAMA, NEJM, BMJ, Preventive Medicine, PLoS One

More than 800 dissertations and theses to date
Will Continue to Solicit, Write, Review, and Help Ancillary Studies Get Funded

1) Merges secondary data sources (e.g., neighborhood, county, and state characteristics) onto Add Health respondent files using personal identifiers

2) Uses archived bio-specimens collected by the Add Health study to develop new measures
Add Health Data Availability
Obtaining & Using Add Health Data

1) Restricted Use:
   Through User Contracts
   GWAS Data … through dbGaP
   For info, go to:
   https://addhealth.cpc.unc.edu/data/
   https://www.cpc.unc.edu/projects/addhealth/documentation

2) Public Use: through ICPSR … go to:
   https://www.icpsr.umich.edu/web/DSDR/studies/21600
Other Major Add Health Data Sets

Omics: https://addhealth.cpc.unc.edu/about/omics/
- GWAS
- Candidate Genes
- Polygenic Risk Scores

Add Health Parent Study (PIs: Kathleen Mullan Harris & Joseph Hotz): https://addhealth.cpc.unc.edu/about/#studies-satellite

Sexual Orientation, Gender Identity, Socioeconomic Status, and Health Across the Life Course (PIs: Carolyn Tucker Halpern & Kerith Conron) … data coming soon: