

Add Health Research Design



UNC
CAROLINA
POPULATION
CENTER



Add Health

The National Longitudinal Study of Adolescent to Adult Health

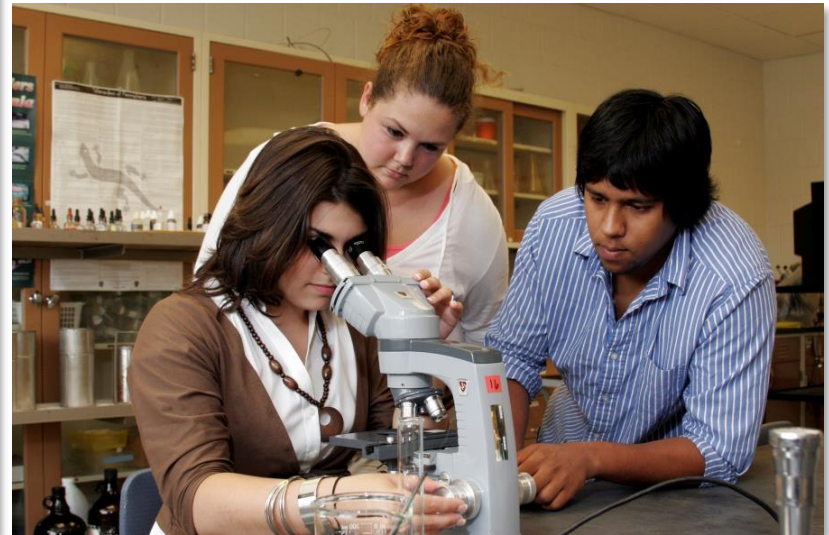
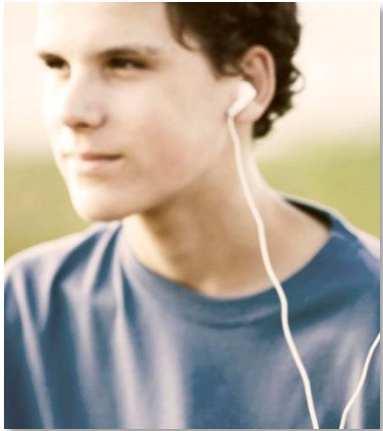
National Longitudinal Study of Adolescent to Adult Health

- 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.

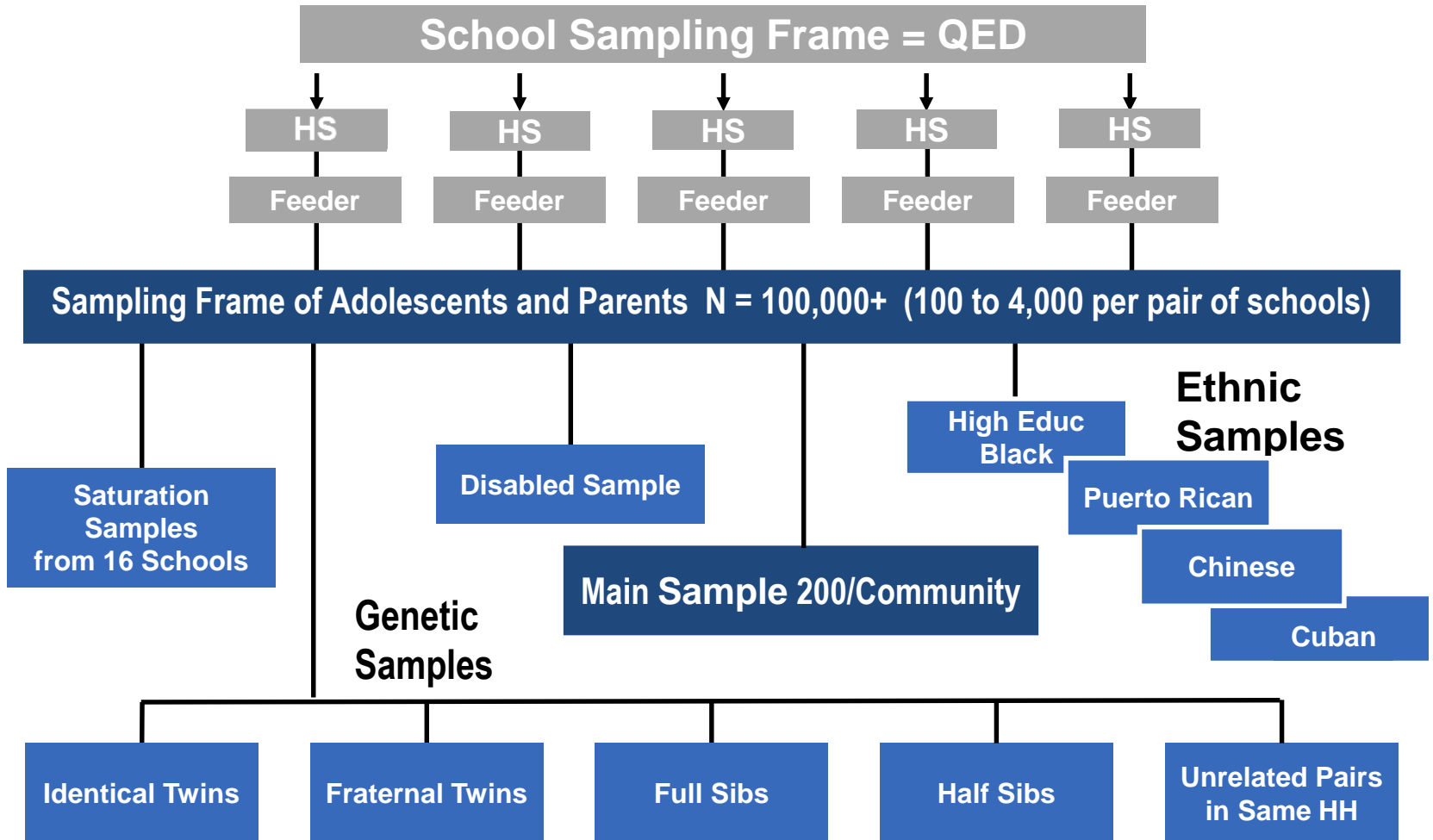
Key Features of Add Health

- 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

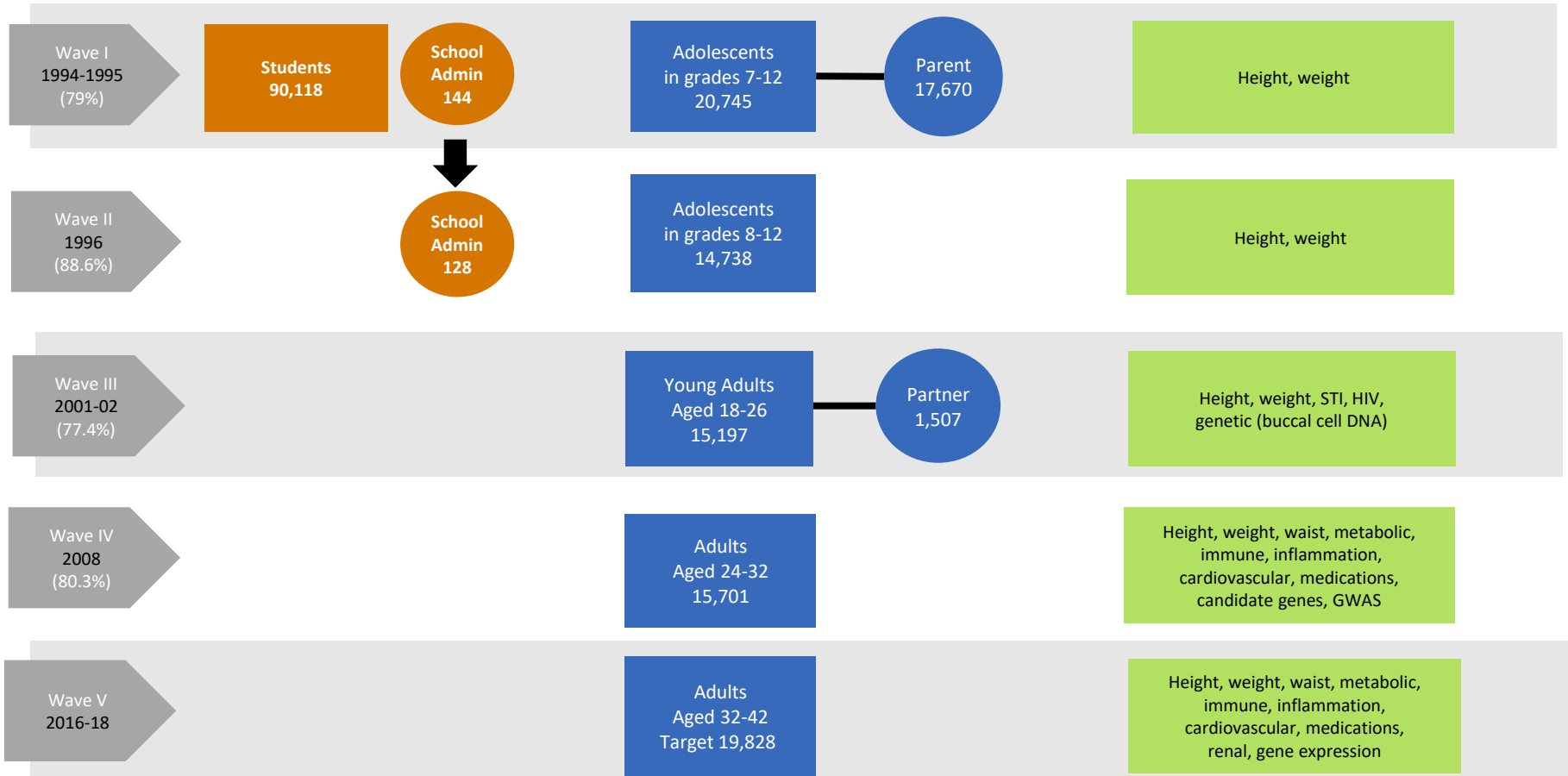


Longitudinal Design

In-School Administration

In-Home Administration

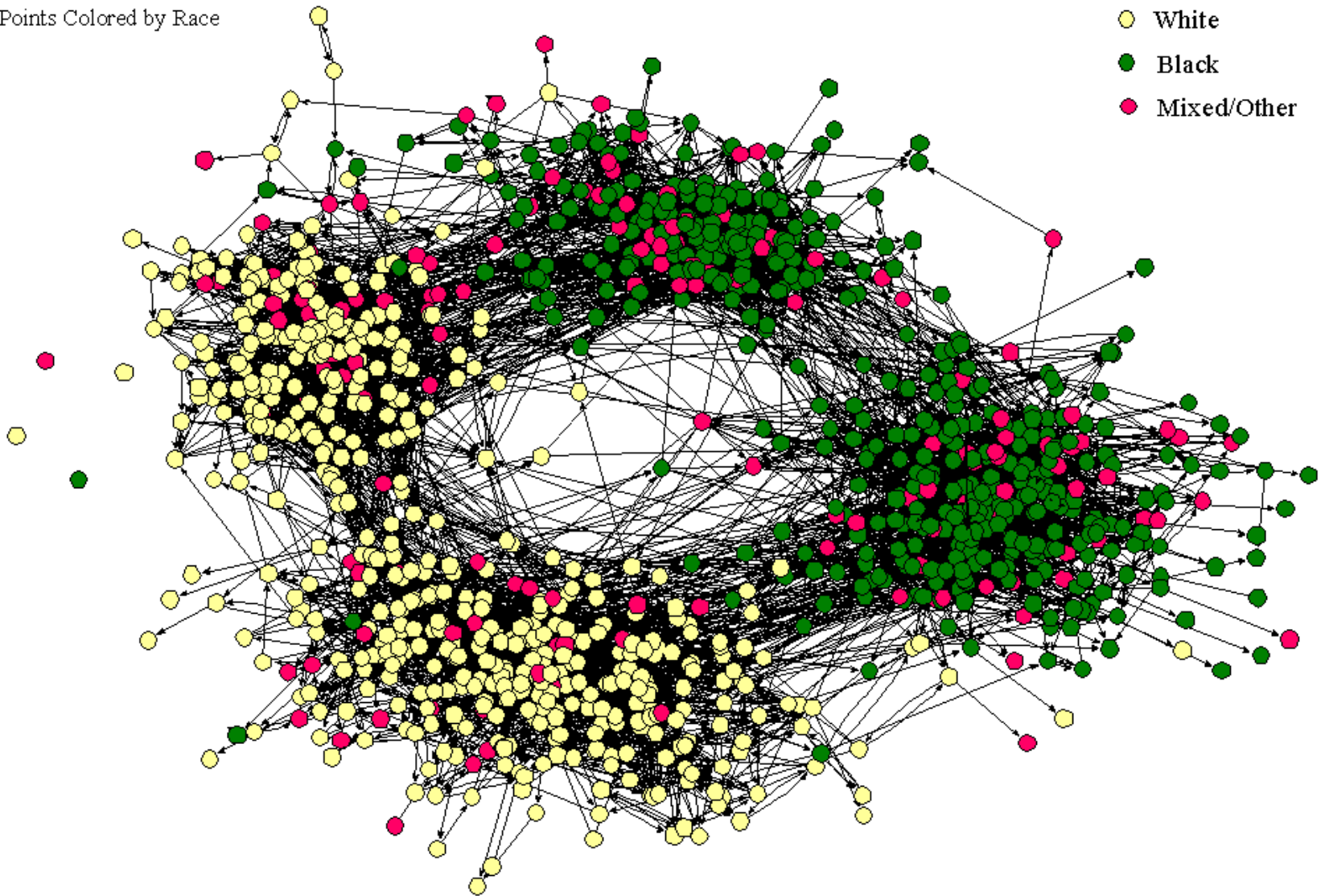
Biological Data Collection



The Social Structure of “Countryside” School District

Points Colored by Race

- White
- Black
- Mixed/Other



Source: Moody, 2001, *American Journal of Sociology* 107: 679-716

Wave I Sample Design

- 80 high schools were selected from a sample frame of 26,666.
- Prior to sampling, schools were sorted by:
 - size
 - school type
 - census region
 - level of urbanization
 - percent white

Replacement Schools

- Of the 80 selected high schools, 52 were eligible and agreed to participate.
- The remaining 28 schools were replaced by similar high schools. They were found by sorting the frame by eight variables:
 - school size
 - school type
 - level of urbanization
 - percent white
 - grade span
 - percent black
 - census region
 - census division

Replacement Schools

- Schools were sorted in a random order within each category.
- The replacement school was the school that followed the initial sample school on the sorted file.
- As a result, the replacement school matched the selection it was replacing with respect to the eight characteristics.

Replacement Schools

- If the first replacement school was ineligible or failed to cooperate, it was replaced by the next school on the list.
- Within some categories, there were not enough potential replacements, or the list of replacements was exhausted. In these cases, similar categories were combined and the file resorted.

Feeder Schools

- Participating high schools were asked to identify junior high or middle schools that were expected to provide at least 5 students to the entering class of the high school.
- A single feeder school was selected for each high school.

Feeder Schools

- The feeder's probability of selection was proportional to the percentage of the high school's entering class that came from that feeder.
- Four schools had no eligible feeder, as students came from a very large number of junior high or middle schools.
- Twenty high schools were their own feeder schools, as they had grade ranges that included 7th or 8th grades.

Participating High Schools

Region		Grade Span	
Northeast	17	K-12	7
South	27	7-12	10
Midwest	19	9-12	49
West	17	10-12	7
		Other	7
School Type		Metropolitan Status	
Public	71	Urban	24
Catholic	3	Suburban	42
Other Private	6	Rural	14



Parental Consent

- Parental consent was required to list student names in a directory and to allow students to participate in the study.
- Unless otherwise directed by the school, passive consent forms were used (it was assumed that a parent granted permission unless the form was returned with a signature that indicated otherwise).
- Some schools required active consent forms (the form had to be returned with a signature indicating that permission was granted).

Security System

- To protect the identities of participants, a rigorous security system prevents anyone from being able to link a respondent's answers to a name or other identity.
- Identification numbers used to collect data are never used for data distribution.
- At the same time, the security system allows researchers to link questionnaires across all components of the study

In-school Components

- 145 middle, junior high, and high schools participated.
- 90,118 students completed a 45-minute questionnaire.
- Each participating school was asked to complete a School Administrator questionnaire.

In-school Questionnaire

Gathered general descriptive information about the student:

- student's and parent's background
- his or her friends
- school life
- school work and school activities
- general health status and health-related behaviors

School Administrator Questionnaire

Gathered information about the educational setting and environment of the school.

- Content areas included:
 - general characteristics of the school and student body
 - curriculum, school services, and programs

In-Home Components

- Each school provided a roster of all students enrolled.
- From the rosters and the pool of participants in the in-school survey, adolescents in grades 7 to 12 were sampled to participate in the In-Home Interview.
- Adolescents were interviewed at two points in time, first at Wave I and then a year later, at Wave II.

In-Home Interview

- Written informed consent was obtained from the parent or legal guardian and the adolescent.
- A Computer-Assisted Personal Interview (CAPI)/Audio Computer-Assisted Self Interview (ACASI) was administered to the adolescent.
- Sections with more sensitive questions were asked in the self-administered portion of the interview.

Parent Interview

- A parent or guardian was interviewed during Wave I of the study.
- This interview provided further information about the family composition and the adolescent's health history.
- The questionnaire asked demographic and health-related information about the parent or guardian and general questions about the adolescent respondent.

In-Home Sample

- A core sample of adolescents derived from rosters supplied by the selected schools and from in-school participation
- A sample for a study of relationship patterns where all respondents were selected from two schools
- A genetic sample composed of siblings and twins
- A sample of unrelated adolescents who reside in the same household

In-Home Sample

- An oversample of black adolescents with college-educated parents
- An oversample of Cuban and Puerto Rican adolescents
- An oversample of Chinese adolescents
- An oversample of physically disabled adolescents (Wave I only)

Sample Selection: Core and PAIRS Samples

- The core sample consisted of roughly equal-sized samples drawn from 12 student-level strata.
- Strata were formed by cross-classifying students by their sex and grade.
- Overall sample-size targets were set for each stratum by dividing the total size of the core sample by the number of strata.

Sample Selection: Core and PAIRS Samples

- School-level targets were also set for each stratum by dividing the overall stratum target by the number of schools with at least one student in the stratum.
- The main frame for selecting the core sample was the set of rosters developed at the sample high schools and their linked feeders.
- At the two (purposively selected) PAIRS schools, all of the students were selected for data collection.

Sample Selection: Non-Genetic Supplements

Eligibility was determined by race/ethnicity and by disability status:

- **High education blacks** – included black students, either of whose parents was a college graduate
- **Cubans** – included students who were from Cuba or whose parents were from Cuba

Sample Selection: Non-Genetic Supplements

- **Puerto Ricans** – included students of Puerto Rican descent
- **Chinese** – included students of Chinese descent
- **Disabled** – included students who had difficulty using their limbs for the year prior to the survey, and, as a result, used a cane, wheelchair, orthopedic shoes, artificial limb, or some other mechanical aid (Wave I only)

Sample Selection: Genetic Supplements

- **Twins** – any student who identified himself as a twin (in the in-school questionnaire); previously unreported twins discovered during the In-Home Interview were added to the sample at that time.
- **Other siblings of twins** – the non-twin siblings of those in the twin sample; to be eligible they had to be enrolled in grades 7 to 12 at the time of sample selection.

Sample Selection: Genetic Supplements

- **Other full siblings** – full siblings where neither member was a twin and both were enrolled in grades 7 to12.
- **Half-siblings** – pairs of half-siblings of which both members were enrolled in grades 7 to12.
- **Non-related** – adolescents living in the same household who did not share the same biological mother or father and were enrolled in grades 7 to12.

Wave I In-Home Adolescent Participants

Overall	20,745	Twins	1,534
Core	12,105	Full-siblings	2,500
Disabled	957	Half-siblings	848
High ed blacks	1,547	Non-related	1,314
Cuban	538	Pairs	2,553
Puerto Rican	633		
Chinese	406		



Wave II Sample Selection

- The Wave II sample was drawn primarily from the pool of participants in Wave I.
- The majority of 12th-grade respondents were removed from the Wave II sample, as they exceeded the grade eligibility requirement.

Wave II Sample Selection

- Twelfth-graders who were part of a genetic pair were retained.
- Wave I disabled sample was not re-interviewed at Wave II.
- The Wave II sample contains a small number of adolescents who did not participate in the first wave.
- No parent interview was conducted during the second wave.

Wave III In-Home Components

- **Wave III respondents** **15,197**
 - Includes 15,170 original Wave I respondents and 27 Wave II special genetic respondents
 - 14,979 were interviewed during the main study and 218 were interviewed during the pretest
- **Romantic partners** **1,507**
 - Partners of Add Health respondents, new in Wave III

Wave III Interview

- Interviewer administered In-Home Interview (CAPI) with sections containing sensitive questions asked in the self-administered portion.
- Respondents who agreed to provide a urine and/or saliva sample were asked to read and sign an additional consent form.
- Respondents received an additional incentive for these samples.

Wave III Biomarkers

- Urine was analyzed for the presence of three sexually-transmitted diseases (STDs), chlamydia, gonorrhea, and trichomoniasis.
- Saliva was collected and analyzed for DNA from all full sibs and twins interviewed at Wave III.
- Additional saliva was analyzed for the presence of HIV antibodies.

Wave III Binge Drinking

- All respondents with Wave I and II data who were in the 7th or 8th grade at Wave I were pre-identified to answer questions about attitudes toward binge drinking.

Composition

- 700 attending two or four year college: 350 males, 350 females, never married.
- 700 not attending college: 350 males, 350 females, never married.

Wave III Couples

- 1500 from pre-selected group of 10,000
- final quota to include 1/3 married, 1/3 cohabiting, 1/3 dating
- heterosexual couples only
- 18 years or older
- current relationship
- relationship had lasted three months or more

Additional Wave III Components

- measured and self-reported height and weight
- latitude and longitude of current residence
- questions about contact with school friends
- information on relationships with siblings

Wave III Sample Selection

- The Wave III sample was drawn primarily from the pool of participants in Wave I.
- Respondents were required to be 18 years of age or older.
- Partners had to be 18 years old or older, opposite sex, in a current relationship that had lasted for at least three months.

Wave III Consent

- Respondent was asked to read and sign an informed consent form.
- Parental consent was not needed as respondents were 18 to 26 years old.
- All respondents who agreed to participate in the interview received an incentive payment.

Wave IV Sample Selection

- Follow-up of Wave I respondents.
- Data collection in 2008 when cohort members were 24-32 years old and completing transition to adulthood.
- 92.5% locate rate
- 80.3% response rate
- N=15,701

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 12,200 DNA samples available for further testing
- Currently conducting GWAS genotyping on archived samples
 - Will be deposited into dbGaP in 2018.

Design Features of Wave IV

- 90 minute computer-based survey instrument
- 30 minute biomarker collection
 - Anthropometrics
 - Blood pressure
 - Blood spots
 - Saliva
- IIV (Intra-Individual Variation) Study

Wave V Data Collection

- Wave I respondents who will be moving through their 4th decade of life.
- Mixed mode survey design:
 - Web/mail with in-person and phone non-response follow-up
- Continuous interviewing 2016-2018
- Collect biological measures and specimens
 - Venous blood draw



Web



Paper



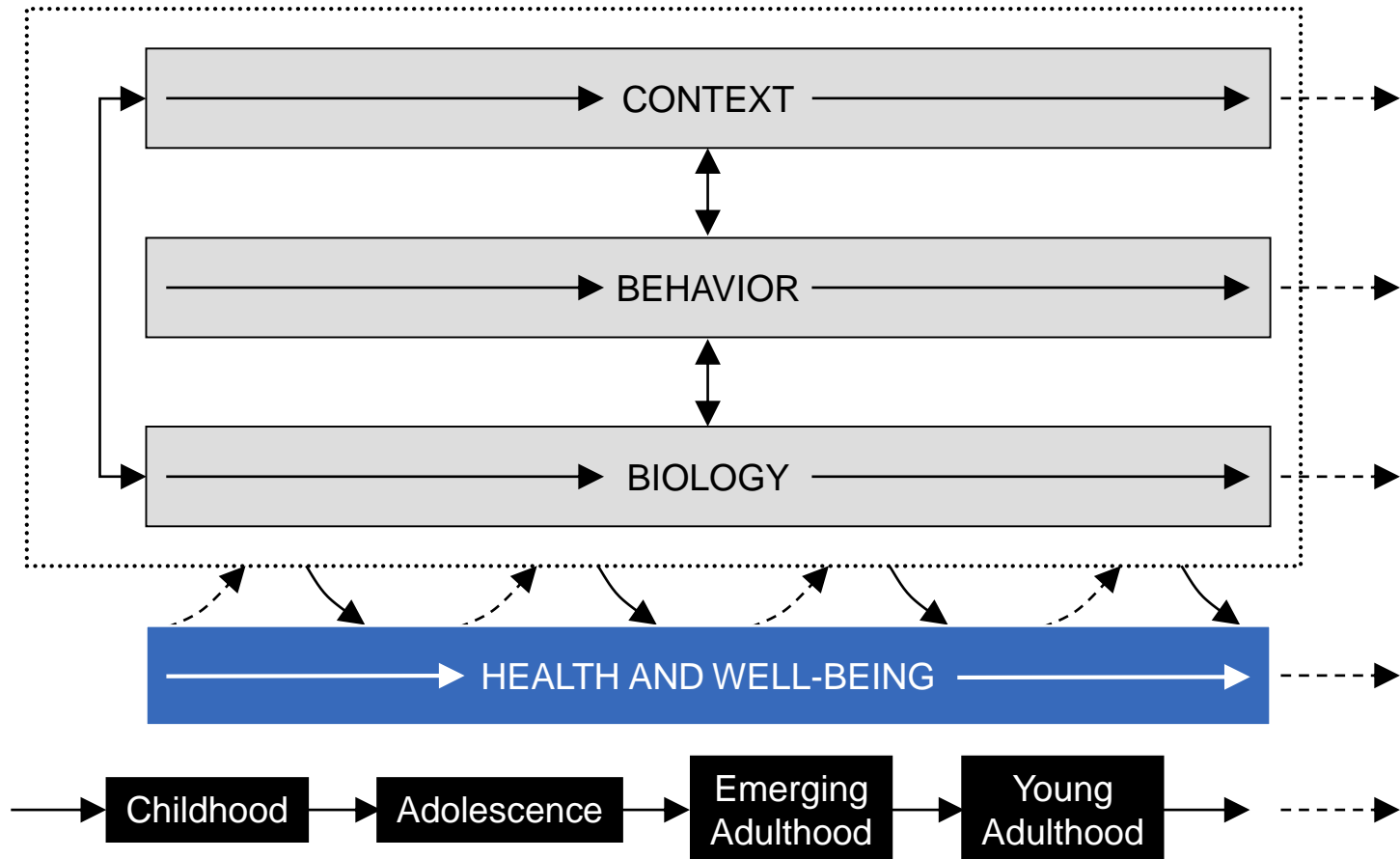
Biospecimen
Collection

Wave V mixed mode

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.

Integrative Life Course Theoretical Framework



Social and Biological Longitudinal Data in Add Health

Adolescence



Adulthood

Wave I-II
(12-20)

Wave III
(18-26)

Wave IV
(24-32)

Wave V
(31-42)

Social environmental data:

school

college

college

work

family

family

family

family

romantic rel

romantic rel

romantic rel

romantic rel

neighborhood

neighborhood

neighborhood

neighborhood

community

community

community

community

peer

peer

Biological data:

Biological resemblance to siblings in household on 3,000 pairs

height

height

ht, wt, waist, BMI

ht, wt, waist, BMI

weight

weight, BMI

BP, pulse

BP, pulse

BMI

STI test results

immune

immune

HIV test results

inflammation

inflammation

DNA

diabetes


diabetes

DNA

kidney disease

GWAS

Biological Data Across Waves

Adolescence	Transition to Adulthood	Young Adulthood	Adulthood
Wave I-II (Ages 12-20)	Wave III (Ages 18-26)	Wave IV (Ages 24-32)	Wave V (Ages 31-42)
Embedded genetic sample of ~3,000 pairs			
Physical development 			
Height, weight	Height, weight	Height, weight, waist	Height, weight, waist
	STI tests (urine)	Metabolic	Metabolic
	HIV test (saliva)	Immune function	Immune function
	Genetic (buccal cell DNA)	Inflammation	Inflammation
		Cardiovascular	Cardiovascular
		Genetic (buccal cell DNA)	Genetic (whole blood)
		Medications	Medications
		GWAS	Renal

Add Health Accomplishments

- Data made available to more than 10,000 investigators for analyses.
- Almost 700 grants awarded to analyze data.
- 3,000+ peer-reviewed publications (over 1,800 since 2006)
- 30 books, 100 reports, and 80 book chapters based on Add Health data
- 600 master's theses and doctoral dissertations

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 4 co-funders

^Wave 5 co-funders